

## ORIGINAL ARTICLE

## Investigating the dose-response relation between air pollution and total mortality in the APHEA-2 multicity project

E Samoli, G Touloumi, A Zanobetti, A Le Tertre, Chr Schindler, R Atkinson, J Vonk, G Rossi, M Saez, D Rabczenko, J Schwartz, K Katsouyanni

*Occup Environ Med* 2003;**60**:977-982

See end of article for authors' affiliations

Correspondence to:  
Dr Katsouyanni,  
Department of Hygiene  
and Epidemiology,  
University of Athens,  
Medical School, 75 Mikras  
Asias Str, Athens 115 27,  
Greece; [kkatsouy@med.uoa.gr](mailto:kkatsouy@med.uoa.gr)

Accepted  
28 February 2003

**Background:** Several recent studies have reported significant health effects of air pollution even at low levels of air pollutants, but in most of these studies linear non-threshold relations were assumed.

**Aims:** To investigate the NO<sub>2</sub> mortality dose-response association in nine cities participating in the APHEA-2 project using two different methods: the meta-smooth and the cubic spline method.

**Methods:** The meta-smooth method developed by Schwartz and Zanobetti is based on combining individual city non-parametric smooth curves; the cubic spline method developed within the APHEA-2 project combines individual city estimates of cubic spline shaped dose-response relations. The meta-smooth method is easier and faster to implement, but the cubic spline method is more flexible for further investigation of possible heterogeneity in the dose-response curves among cities.

**Results:** In the range of the pollutant common to all cities the two methods gave similar and comparable curves. Using the cubic spline method it was found that smoking prevalence acts as an effect modifier with larger NO<sub>2</sub> effects on mortality at lower smoking prevalence.

**Conclusions:** The NO<sub>2</sub>-mortality association in the cities included in the present analysis, could be adequately estimated using the linear model. However, investigation of the city specific dose-response curves should precede the application of linear models.

Adverse short term effects of specific air pollutants on health outcomes have been documented in several epidemiological studies in recent years.<sup>1-5</sup> The pollutants implicated were mainly ambient particles<sup>4, 6, 7</sup> but gaseous pollutants such as NO<sub>2</sub>, O<sub>3</sub>, and CO have been shown to have adverse effects on mortality and morbidity.<sup>3, 8</sup> The results from epidemiological studies have led to revisions of air quality guidelines and standards and scheduled dates for regular revisions in the future.<sup>9-11</sup> These, in turn, increase the interest in better understanding the specific air pollutant-health outcome associations and reducing uncertainties. One issue that requires further investigation is the dose-response relation, especially for gaseous pollutants, which have received less attention so far. Dose-response functions are being used extensively for health impact assessment.<sup>12, 13</sup> Whether or not there is a threshold makes a large difference to the estimate of attributable deaths, and the linearity or otherwise of the dose-response association is important for predicting the benefits of regulation. Combined estimates are often used in health impact assessment, rather than individual city estimates, and thus it is important to investigate methods of combining dose-response functions across cities.

Most of the published studies on short term effects of air pollutants analyse daily time-series data and have been based on the assumption that the underlying associations are linear. A variety of analytical methods have been applied to analyse single city data, while most of the recent reports have used generalised additive Poisson regression models.

Recently, multicity national or international programmes have provided results based on data from several cities. Combined evidence was obtained using hierarchical models implemented in two stages by either classical or Bayesian statistical procedures.<sup>4, 5</sup> In either case, in the first stage data from each city were analysed separately, whereas in the second stage the city specific air pollution effects estimates were regressed on city specific covariates to obtain an overall

estimate and to explore the variation in the estimates across cities.

Two methods have been developed and proposed for the estimation and the combination of dose-response curves. The first, developed by Schwartz and Zanobetti<sup>14</sup> combines individual city specific non-parametric smooth curves (meta-smooth method); the second combines individual city estimates of natural cubic spline shaped dose-response relations (natural cubic spline method).<sup>15</sup> The latter method used Bayesian procedures for the estimation. Both methods have been applied to estimate the PM<sub>10</sub>-daily number of deaths relation in the US cities. In the present paper we applied the meta-smooth method developed by Schwartz and Zanobetti,<sup>14</sup> and a modification of the natural cubic spline method to investigate the dose-response relation between NO<sub>2</sub> and the daily number of deaths. Results from both methods were compared between them and with the results obtained from a linear non-threshold model. The natural cubic spline method developed by Daniels and colleagues<sup>15</sup> allows the data to determine the boundary knots, that are set to the minimum and maximum air pollution levels for each city. In the modified cubic spline method developed within the APHEA-2 project, unrestricted cubic splines (cubic spline method), rather than natural cubic splines, were used to fit individual city dose-response curves. In this way, the number and the location of the knots are pre-specified and the same for each city.

Very little work on the dose-response of the NO<sub>2</sub> effects on mortality has been produced so far. NO<sub>2</sub> is a major pollutant with accumulated evidence on its health effects<sup>16, 17</sup> and a good marker for traffic generated pollution. Because of its latter characteristic, it has also received attention as a potential effect modifier of the associations between other pollutants and health.<sup>5</sup> It is, therefore, important to obtain an understanding of the dose-response shape of its association with daily deaths.

## Main messages

- The NO<sub>2</sub> total mortality association can be adequately estimated by the linear model, at concentration levels most commonly observed in Europe.
- Two methods for estimating and combining dose-response were evaluated. While the non-parametric “meta-smooth” method is simpler to apply, the parametric “cubic splines” method allows for further investigation of heterogeneity between individual location dose-response curves.
- The two methods give comparable results under some restrictions.

## METHODS

### Data

The APHEA-2 project is a multicentre study including 30 cities across Europe and associated regions (Istanbul and Tel Aviv) which studies health effects of air pollution. Data were collected on daily counts of all-cause mortality, excluding deaths from external causes (International Classification of Disease, ICD-9 >800). The data covered at least three consecutive years for each city within the years 1990–97. Details about the data have been published elsewhere.<sup>5</sup>

Daily air pollution measurements were provided by the monitoring networks established in each town participating in the APHEA-2 project. A monitor was included if certain completeness criteria were fulfilled.<sup>18</sup>

Time series data on daily temperature (°C, daily mean) and relative humidity (%) were used to control for the potential confounding effects of weather. External information on influenza epidemics or other unusual events (heatwaves, strikes, etc) were also collected, if available.<sup>5</sup>

In the present study we have analysed the average of lags 0 and 1 for NO<sub>2</sub> (maximum 1 hour value) for all cities, since there is evidence that the average of two days' pollution correlates better with mortality than a single day's exposure.<sup>19</sup>

Nine cities were selected on the basis of extensive pollutant ranges: Athens, Budapest, Madrid, Marseilles, Milan, Rome, Tel Aviv, Torino, and Valencia. Combining cities with wide and extensively overlapping NO<sub>2</sub> ranges was judged necessary for the application of both methods. Specifically, the selected cities had NO<sub>2</sub> levels with medians above 110 µg/m<sup>3</sup> and the third quartiles above 130 µg/m<sup>3</sup>.

The analysis was restricted to days with NO<sub>2</sub> levels below 300 µg/m<sup>3</sup>, since these few extreme values could influence disproportionately the results from both methods. Restricting analyses to days with less than 300 µg/m<sup>3</sup> resulted in excluding less than 2% of the available days in each city, except for Tel Aviv, for which 6% of the days were excluded.

### Methods

The pollution-mortality associations for each city were investigated using generalised additive Poisson regression models<sup>20</sup> allowing also for overdispersion. This approach allows inclusion of non-parametric smooth functions to model the potential non-linear dependence of daily deaths on weather and season. Assuming linear NO<sub>2</sub>-daily mortality association, the city specific model is of the form:

$$\log(E(Y^c)) = \beta^c NO_2 + \sum_{i=1}^q S_i^c(Z_i^c) \quad (1)$$

## Policy implications

- The health effects of air pollution have been documented in many studies and the results are used for regulatory decisions concerning air management. In assessing health risk, it is critical to have reliable information on dose-response relations.
- Dose-response curves assessing the association between NO<sub>2</sub> concentrations and total mortality were shown to be meaningfully combined to produce one curve based on multiple location data by using two methods: the non-parametric “meta-smooth” method and the parametric “cubic spline” method. The two methods give comparable results under certain conditions.
- Both methods indicate that there are adverse short term effects of NO<sub>2</sub> on mortality, at least in the range between 100 µg/m<sup>3</sup> and 200 µg/m<sup>3</sup>.
- The linear dose-response curve may be used to estimate the NO<sub>2</sub> effects on mortality at concentration levels commonly observed in Europe, but it is advised to investigate city specific curves before combining.

where  $Y^c$  is the daily count of deaths at city  $c$ ,  $E(Y^c)$  is the expected value of that count,  $\beta^c$  the NO<sub>2</sub> estimated log-relative rate at city  $c$ ,  $Z_i^c$  the non-pollutant covariates and smooth functions of these covariates. We chose loess,<sup>21</sup> a  $S_i^c$  moving regression smoother, to estimate the functions  $S_i$ . This has become a standard method in time series studies of mortality and morbidity.<sup>22</sup>

Following the APHEA-2 methodology, published in detail elsewhere,<sup>22a</sup> the first step in the city specific analyses is to adequately control for seasonal variation and long term trends. The principal issue in the use of non-parametric smoothers is the choice of the fraction of the data (smoothing parameter) that will be included in the running smooth. On this, we have 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 minimises Akaike's Information Criterion (AIC).<sup>23</sup> Time is used as a proxy for any outcome predictors not included in the model, which have long term trends or seasonal patterns. 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. It has been decided to choose the span that minimises partial autocorrelation in the residuals, but not to use a span of less than two months. This minimises the necessity for fitting autoregressive terms. Second, if serial correlation exists in the residuals, this may indicate that effects of an omitted time dependent covariate are still present in the data. Therefore, minimising this correlation seems a natural goal. To allow for city specific differences, the smoothing parameters for the covariates were optimised separately in each location. Day of the week effects, holidays, and epidemics were controlled for by using dummy variables. Robust regression<sup>24</sup> was used to reduce the effect of any extreme observations on the regression results. We then entered the pollutant variable in two ways according to the requirements of each method.

Recently, Dominici and colleagues<sup>25</sup> identified that the application of GAM models in the SPLUS software with the default convergence criteria leads to biased parameters' estimates, while Ramsay and colleagues<sup>26</sup> found in addition that this function underestimated the parameters' variances. In response to these findings, we have reanalysed our data

using, as suggested, more stringent convergence criteria (that is, we set the maximum number of iterations to 1000 and the difference of two successive coefficients to  $10^{-14}$ ). In principle, this approach will eliminate the bias in the parameters' estimates. In order to deal with the underestimation of the parameters' variances we also applied the bootstrap method,<sup>27, 28</sup> which produces robust standard errors.

In the meta-smooth approach, developed by Schwartz and Zanobetti,<sup>14</sup> a smooth function of NO<sub>2</sub> with span equal to 0.5 is entered in the individual city model. More details on this have been published elsewhere.<sup>14</sup> In our data the span of 0.5 corresponded to a window of 546–1248 days, depending on the availability of the data from each city. According to the published methodology<sup>14</sup> and via the bootstrap method we computed the predicted values of the log relative risk of the daily number of deaths in each city for 5 µg/m<sup>3</sup> increments between 16 µg/m<sup>3</sup> and 300 µg/m<sup>3</sup>, along with their standard errors. The lower limit is the minimum value of the NO<sub>2</sub> across cities. These predicted values at each increment were then combined across cities using inverse variance weighting.

In the cubic spline method unrestricted cubic splines were used to estimate the exposure-response relation for each city. The knots were pre-specified and were the same for each city. This had the advantage that similar terms were pooled in the second stage of the analysis. The number and location of the knots were determined according to exploratory graphical analysis results, based on the previously described individual city models using loess. Three distinct patterns were evident across cities. Figure 1 shows the pollutant-mortality dose-response relation in the nine cities of the analysis, and

presents the three distinct patterns (linear and two parabolas) of the dose-response curves. We decided to use a cubic spline with two knots at 100 and 140 µg/m<sup>3</sup> to sufficiently capture the NO<sub>2</sub>-mortality association in our data. The (unrestricted) cubic spline function of a variable  $x$  is:

$$S(x) = \sum_{j=0}^3 \beta_{0j}x^j + \sum_{i=1}^k \beta_{i3}(x - t_i)_+^3 \quad (2)$$

where  $k$  is the number of knots, i.e. in our case  $k = 2$  and  $t_1 = 100, t_2 = 140$ , and using the “+” notation of Smith:<sup>29</sup>

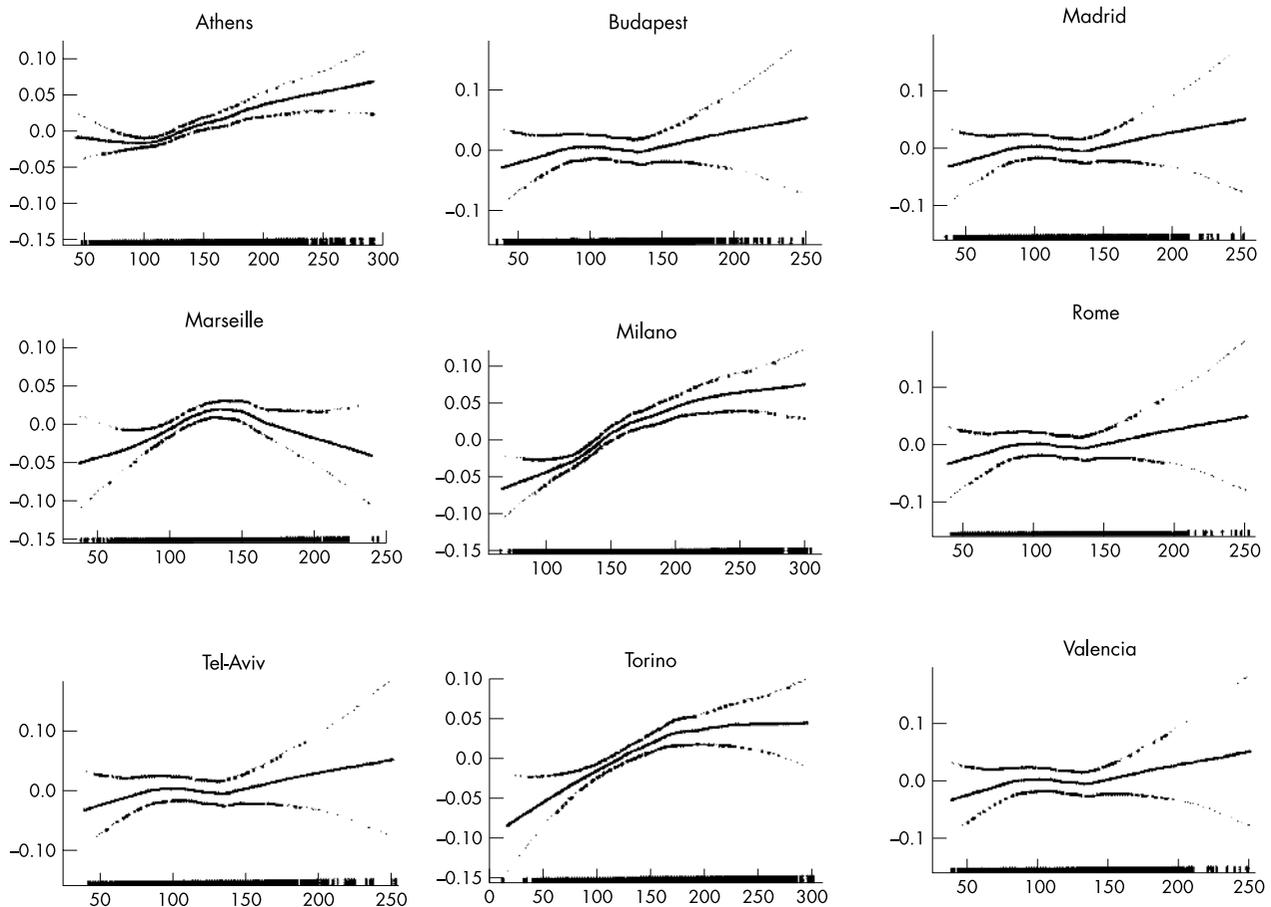
$$u_+ = u \text{ if } u > 0 \\ u_+ = 0 \text{ if } u \leq 0 \quad (3)$$

Hence our cubic spline dose-response model assumed that:

$$\log(E(Y)) = S(\text{NO}_2, \text{knots}(100,140)) + \text{confounders} \quad (4)$$

where  $S(\text{NO}_2, \text{knots}(100,140))$  represents the parameterisation of the spline function. Linear (2) and spline (4) models were fitted in Splus. For the latter the  $bs$  function was used.

In the second stage we regressed the city specific estimates produced from the first stage of the analysis ( $\beta^c$ ) on city specific covariates ( $\mathbf{X}^c$ ) to obtain the overall dose-response curve and to explore potential heterogeneity in the city specific curves. For the linear model  $\beta^c$  is the log-relative rate in city  $c$ , whereas for the spline model,  $\beta^c$  is the vector of the regression coefficients corresponding to the spline function (equation 4).



**Figure 1** Dose-response curves of NO<sub>2</sub> and total mortality in nine European cities. The x-axes are NO<sub>2</sub>-1h levels in µg/m<sup>3</sup>. The vertical scale can be interpreted as the relative rate of mortality as a function of NO<sub>2</sub>.

For the spline method, we fitted multivariate second stage regression models based on the method described by Berkey and colleagues.<sup>30</sup> More specifically, the models are of the form:

$$\beta^c = \mathbf{X}^c \alpha = \delta^c + \epsilon^c \tag{5}$$

where  $\beta^c$  is the (5×1) vector of the five spline estimates in each city  $c$  (the intercept term in equation (4) was ignored since only relative risks are considered);  $\mathbf{X}^c$  is a 5×5p matrix, where  $p$  is the number of city level covariates for city  $c$  (including the intercept). The rows of  $\mathbf{X}^c$  are divided into five blocks of  $p$  elements such that in the row corresponding to the  $i$ -th spline parameter the  $i$ -th block contains the  $p$  city specific covariate values whereas all other entries are 0;  $\alpha$  is the vector of regression coefficients to be estimated—it may include a separate intercept for each spline parameter and a separate slope for each spline parameter against each corresponding covariate;  $\delta^c$  is a vector of five random effects associated with city  $c$  representing, for each spline estimate, the city's deviation from the overall model and  $\epsilon^c$  (assumed independent from  $\delta^c$ ) is the vector of sampling errors within each city.

The 5×5 matrix  $\text{cov}(\delta^c) = \mathbf{D}$  represents the between cities covariances of the regression coefficients of the spline function that is unexplained by the regression.

It is assumed that:

$$\begin{aligned} \delta^c &\sim \text{MVN}(\mathbf{0}, \mathbf{D}) \\ \epsilon^c &\sim \text{MVN}(\mathbf{0}, \mathbf{S}^c) \\ \beta^c &\sim \text{MVN}(\mathbf{X}^c \alpha, \mathbf{D} + \mathbf{S}^c) \end{aligned} \tag{6}$$

where  $\mathbf{S}^c$  is the variance-covariance matrix, of the five regression coefficients of the spline function, in city  $c$  that is estimated in the first stage of the analysis, after applying the bootstrap method. When  $\mathbf{D} \approx \mathbf{0}$  we get the corresponding fixed effects estimates while when  $\mathbf{D} \neq \mathbf{0}$  we get the random effects estimates.

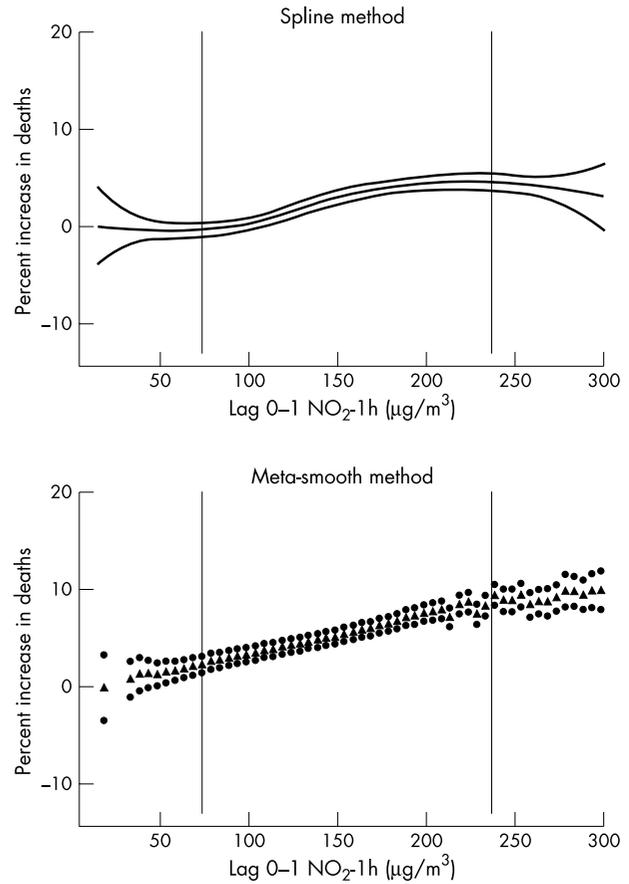
The iterative generalised least squares method was applied to estimate model parameters. That is:

$$\begin{aligned} \hat{\alpha} &= \left( \sum_c \mathbf{X}^{cT} \mathbf{V}^{c-1} \mathbf{X}^c \right)^{-1} \left( \sum_c \mathbf{X}^{cT} \mathbf{V}^{c-1} \beta^c \right) \\ \text{with } \text{Var}(\hat{\alpha}) &= \left( \sum_c \mathbf{X}^{cT} \mathbf{V}^{c-1} \mathbf{X}^c \right)^{-1} \end{aligned} \tag{7}$$

where  $\mathbf{V}^c = \mathbf{S}^c$  for the fixed effects estimates and  $\mathbf{V}^c = \mathbf{S}^c + \mathbf{D}$  for the random effects estimates. The parameters of the between cities covariance matrix  $\mathbf{D}$  are estimated by maximum likelihood.<sup>31</sup>

For the linear effects, model (5) collapses to a univariate one that expresses the usual meta-regression. In this case  $\mathbf{D}$  denotes the between cities variance in the effects estimates and can be estimated from the data using the maximum likelihood method described by Berkey and colleagues.<sup>31</sup>

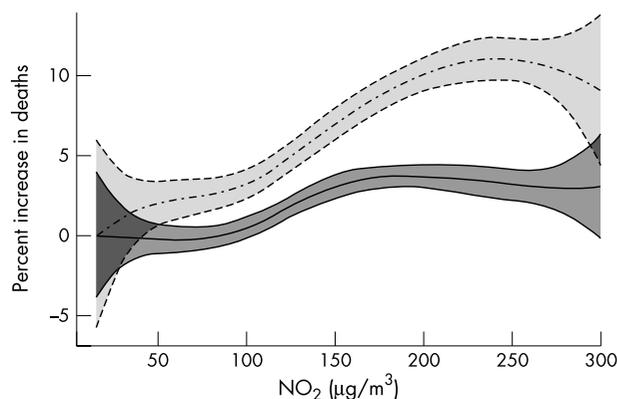
After obtaining an overall curve that draws information from all cities, we also compared the three types of models—the linear, the cubic spline, and the meta-smooth, within each city and over all cities to determine which best fits the data. We used the Akaike information criterion (AIC)<sup>23</sup> to compare these two methods to the linear dose-response model without threshold representing the standard approach in time series analyses estimating effects of air pollution on mortality or morbidity. For an overall comparison of the three different models, we computed the sum of the city specific AIC values.



**Figure 2** Average dose-response curves across nine cities and their 95% CIs provided by the cubic spline method (top) and the meta-smooth method (bottom). The vertical lines in both plots mark the common range of NO<sub>2</sub> values in all analysed cities.

**RESULTS**

Figure 2 shows the estimated combined overall dose-response curves between NO<sub>2</sub> and total mortality, and their 95% confidence intervals (CIs), that we get from the second stage models based on the cubic spline method and the meta-smooth method in our nine study locations. The meta-smooth method gives a more linear shape than the cubic spline method, which reflects the S shape observed in some cities. Note that not all cities have values for the pollutant at both ends of the distribution, which is obvious from the wide CIs in the end points of the data. For example the 10th centile of NO<sub>2</sub> in Rome (111.6 µg/m<sup>3</sup>) is almost the median value in Valencia (114.6 µg/m<sup>3</sup>). The cubic spline method suggests a slight S shaped pattern with the effect of NO<sub>2</sub> fading above the levels of 150 µg/m<sup>3</sup>, whereas the meta-smooth suggests a more linear association up to 200 µg/m<sup>3</sup>. If we restrict the dose-response curves to the range from about 74 to 237 µg/m<sup>3</sup> where the majority of the data are, we can see that both methods give similar shapes that can be roughly approximated by a linear association. Outside that range, the summary estimates start to involve extrapolated values. The curves of the cubic spline and the meta-smooth method indicate that an increase in NO<sub>2</sub> from 100 µg/m<sup>3</sup> to 150 µg/m<sup>3</sup> is associated with a 2.6% and 1.8% increase in daily deaths, respectively. This is consistent with the results from regressions assuming a linear relation and giving an estimated increase in deaths of about 2% for a 50 µg/m<sup>3</sup> increment in NO<sub>2</sub>. There appears to be little increase in risk until the NO<sub>2</sub> concentration exceeds about 80 µg/m<sup>3</sup>.



**Figure 3** Dose-response curves and their 95% CI of NO<sub>2</sub> and total mortality in the 25th (dotted line) and in the 75th (solid line) centile of the smoking prevalence distribution.

We more formally examined the hypothesis of linearity in the NO<sub>2</sub>-mortality relation by comparing the AIC values obtained under the linear, cubic spline, and meta-smooth models. All methods gave very similar fits, while the meta-smooth method gave a slightly better overall fit (0.1% difference compared to the linear method). This small difference in favour of the meta-smooth method is mainly due to Athens and Marseilles, in which cities the dose-response curves are best captured by the meta-smooth model. In most of the other cities the linear models had a better fit.

To illustrate the use of the cubic spline method to explore sources of heterogeneity, we fit a second stage hierarchical model as described in the methods section. The effect modifier included here is the smoking prevalence in each city, since an exploratory analysis including linear terms for NO<sub>2</sub> indicated smoking prevalence as a statistically significant effect modifier, which reduced heterogeneity in the effect estimates by 23%. Potential effect modifiers used in the APHEA-2 analysis included variables describing the air pollution level and mix in each city, the health status of the population, the geographical area, and the climatic conditions.<sup>5</sup> The smoking prevalence in these nine cities ranged from 28.0% in Tel Aviv to 55.0% in Budapest. Figure 3 shows the resulting dose-response curves (and 95% CIs) for NO<sub>2</sub> for the 25th and the 75th centile of the distribution of smoking prevalence. Figure 3 shows that when the smoking prevalence is lower, the effect of the pollutant on mortality is greater. The curve corresponding to the 25th centile is steeper (that is, there is a stronger effect of the pollutant on mortality) up to about 200 µg/m<sup>3</sup>. The overall  $\chi^2$  test for heterogeneity was reduced by 50% with the inclusion of the smoking prevalence.

## DISCUSSION

Recently there has been growing interest in the methodology used to investigate the dose-response relation between air pollution and daily mortality. In this paper we have applied two methods for obtaining and combining dose-response curves—one method based on cubic spline models and one based on curves produced from non-parametric smooth models—and compared them to models based on the assumption of linearity.

When implementing the two methods it became obvious that the meta-smooth method is easier and faster to implement. On the other hand there are certain aspects of the method that need further investigation. As mentioned in the methods section we computed the predicted values of the log relative risk of daily death in each city for 5 µg/m<sup>3</sup> increments of NO<sub>2</sub>. These predicted values at each increment

were then combined across cities using inverse variance weighting. Hence in order to get the fixed effect estimate we use the available number of cities in each increment, but not all cities will have values in all intervals. Different numbers of city specific estimates are therefore added in each interval; that is not reflected in the variance estimates. The potential effect of not allowing for the different number of points contributing to the overall estimate in each increment has not been evaluated so far. Besides, this implies consistency of the curves between cities without observations in the high range and cities with available data in that range. Another weak point of the meta-smooth method is that it treats the predictions from the models as independent observations and does not take into account their covariances in the estimation of the pooled effect. This could lead to biased standard errors although the magnitude of the induced bias is not known. Finally, instead of using a span of 0.5 in each city it might be better to predefine a window of a fixed number of days and use, so that the dose-response curve would catch the exposure response relation in each city with the same accuracy. Another more adaptive approach towards modelling city specific curves would be to use the AIC criterion to determine the optimal window size, but again this method would not result in the same accuracy over cities.

Daniels and colleagues<sup>15</sup> investigated the dose-response relation between particulate matter and daily mortality by considering linear, cubic spline, and threshold models; the investigators used natural cubic splines. The comparison between the natural cubic splines (used by Daniels *et al*) and the unrestricted cubic splines (used in our methodology) is of particular interest and a subject for further investigation. The main difference between the two methods is that the natural cubic splines allow the data to determine the boundary knots that are set by default to the minimum and maximum air pollution level value for each city, which may vary between cities. In case the pollutant level across cities varies (as is most likely), the interpretation of the pooled results from the natural cubic splines method is unclear, since different regression coefficients are pooled together. Instead, in the unrestricted cubic splines method, where all the knots are pre-specified and the same across cities, similar regression coefficients are pooled in the second stage of the analysis, simplifying the interpretation of the results.

In a more recent paper, Dominici and colleagues<sup>32</sup> presented an improvement in the natural cubic spline method where the parameters of the spline dose-response curves and the number and location of the knots (assumed to be the same in each city) were estimated jointly.

Combining the individual city dose-response curves is an important issue for multi-city projects.<sup>4, 5</sup> Since heterogeneity has been observed in effect estimates based on linear models and specific effect modifiers have been identified, it is important to investigate whether differences in the shape of the dose-response curves are also due to specific city characteristics which act as effect modifiers. A method to study this issue has been developed in the present study. It has been shown that the cubic spline method is flexible for running second stage hierarchical models, and helps to investigate factors influencing the form of the dose-response relation, and thus might provide some explanations for the different shapes observed (for example, whether these are associated with geographical or socioeconomic differences across cities). In this study the role of smoking prevalence in changing the dose-response shape has been explored and a significant reduction in heterogeneity resulted. It was found that when smoking prevalence is lower, the effect of NO<sub>2</sub> on mortality is larger. A likely explanation may be that smoking acts as a competing risk, harvesting the population of susceptible individuals, but this issue needs further investigation.

A potential drawback in the second stage analysis presented in this paper is that the 5×5 between cities covariance matrix *D* is estimated only from the nine cities selected. This may result in unstable estimates, but this can be partly overcome by analysing a greater number of cities.

In conclusion, we have investigated the dose-response association between NO<sub>2</sub> exposure and the daily number of deaths using two flexible methods and compared them to the model based on the linear dose-response assumption. We used nine cities with overlapping NO<sub>2</sub> concentration ranges. The meta-smooth method is the easiest to apply and is useful for initial investigation of the dose-response shape. The spline method has the advantage of allowing the investigation of effect modification on the dose-response shape, which is very useful given the heterogeneity observed in between city curves. The NO<sub>2</sub>-mortality association in the cities included in the present analysis could be adequately estimated using the linear model. However, it became evident that the linear model should not be applied without investigating the city specific dose-response curves first.

## ACKNOWLEDGEMENTS

The APHEA-2 study is supported by the European Commission (EC) Environment and Climate 1994-98 Programme (contract number ENV4-CT97-0534). The Swedish group did not receive funding from the EC.

## Authors' affiliations

**E Samoli, G Touloumi, K Katsouyanni**, Department of Hygiene and Epidemiology, University of Athens, Medical School, Athens 115 27, Greece

**A Zanobetti, J Schwartz**, Harvard School of Public Health, Environmental Epidemiology Program, Boston, USA

**A Le Tertre**, Institut de Veille Sanitaire, Paris, France

**Chr Schindler**, Universität Basel, Institut für Sozial- und Präventivmedizin, Switzerland

**R Atkinson**, Department of Public Health Sciences, St George's Hospital Medical School, University of London, UK

**J Vonk**, Department of Epidemiology and Biostatistics, Faculty of Medicine, University of Groningen, Groningen, Netherlands

**G Rossi**, Institute of Clinical Physiology, CNR, Department of Epidemiology and Biostatistics, Pisa, Italy

**M Saez**, Research Group on Statistics, Applied Economics and Health (GRECS), Department of Economics, Universitat de Girona, Spain

**D Rabczenko**, National Institute of Hygiene, Department of Medical Statistics, Population Studies Laboratory, Warsaw, Poland

## REFERENCES

- Schwartz J, Dockery DW. Increased mortality in Philadelphia associated with daily air pollution concentrations. *Am Rev Respir Dis* 1992;**145**:600-4.
- Pope CA, Dockery DW, Schwartz J. Review of epidemiologic evidence of health effects of particulate air pollution. *Inhal Toxicol* 1995;**7**:1-18.
- Touloumi G, Katsouyanni K, Zmirou D, et al. Short-term effects of ambient oxidant exposure on mortality: a combined analysis within the APHEA project. *Am J Epidemiol* 1997;**146**:177-84.
- Samet JM, Dominici F, Currier FC, et al. Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994. *N Engl J Med* 2000;**343**:1742-9.
- Katsouyanni K, Touloumi G, Samoli E, et al. Confounding and effect modification in the short-term effects of ambient particles on total mortality: results from 29 European cities within the APHEA2 project. *Epidemiology* 2001;**12**:521-31.
- Pope CA III, Dockery DW. Epidemiology of particle effects. In: Holgate ST, Samet JM, Koren HS, et al, eds. *Air pollution and health*. San Diego, CA: Academic Press, 1999:673-705.
- Katsouyanni K, Touloumi G, Spix C, et al. Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. *BMJ* 1997;**314**:1658-63.
- Schwartz J. Air pollution and hospital admissions for cardiovascular disease in Tucson. *Epidemiology* 1997;**8**:371-7.
- World Health Organization. *Air quality guidelines for Europe*, 2nd edn. WHO Regional Publications European Series No. 91. Copenhagen, 2000.
- US Environmental Protection Agency Office of Air Quality Planning and Standards. *Review of the national ambient air quality standards for particulate matter: policy assessment of scientific and technical information*. WAWPS Staff Paper. EPA-45/R-1996:96-013. Research Triangle Park, NC: US EPA Office of Air Quality Planning and Standards, 1996.
- Commission of the European Communities. Council Directive 1999/30/EC relating to limit values for sulphur dioxide, oxides of nitrogen, particulate matter and lead in ambient air. *Official Journal of the European Communities* 1999;**163**:41.29.6.1999.
- Department of Health Committee on the Medical Effects of Air Pollution. *Quantification of the effects of air pollution on health in the United Kingdom*. London: HMSO, 1998.
- Künzli N, Kaiser R, Medina S, et al. Public-health impact of outdoor and traffic-related air pollution: a European assessment. *Lancet* 2000;**356**:795-801.
- Schwartz J, Zanobetti A. Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. *Epidemiology* 2000;**11**:666-72.
- Daniels MJ, Dominici F, Samet JM, et al. Estimating particulate matter-mortality dose-response curves and threshold levels: an analysis of daily time-series for the 20 largest US cities. *Am J Epidemiol* 2000;**152**:397-405.
- Sunyer JSC, Quenel P, Ponce de Leon A, et al. Urban air pollution and emergency admissions for asthma in four European cities: the APHEA project. *Thorax* 1997;**52**:760-5.
- Hwang GS, Chang CC. Effects of air pollution on daily clinic visits for long respiratory tract illness. *Am J Epidemiol* 2002;**155**:1-10.
- Katsouyanni K, Schwartz J, Spix C, et al. Short-term effects of air pollution on health: a European approach using epidemiologic time-series data: the APHEA protocol. *J Epidemiol Community Health* 1996;**50**(suppl 1):S12-18.
- Schwartz J. The distributed lag between air pollution and daily deaths. *Epidemiology* 2000;**11**:320-6.
- Hastie T, Tibshirani R. *Generalized additive models*. London: Chapman and Hall, 1990.
- Cleveland WS, Devlin SJ. Robust locally-weighted regression and smoothing scatterplots. *J Am Stat Assoc* 1988;**74**:829-36.
- Schwartz J. Generalized additive models in epidemiology. *International Biometric Society, Invited Papers*. 17th International Biometric Conference, 1994:55-80.
- Touloumi G, Atkinson R, Le Tertre A, et al. Analysis of health outcome time series data: the APHEA-2 project. *Environmetrics* 2003;**14**:1-17.
- Akaike H. Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, eds. *Second International Symposium on Information Theory*. Budapest: Akademiai Kiado, 1973.
- Hampel FR, Ronchetti EM, Rousseeuw PJ, et al. *Robust statistics: the approach based on influence functions*. New York: John Wiley, 1986.
- Dominici F, McDermott A, Zeger S, et al. On the use of generalized additive models in time-series studies of air pollution and health. *Am J Epidemiol* 2002;**156**:193-203.
- Ramsay T, Burnett R, Krewski D. The effect of concurvity in generalized additive models linking mortality and ambient air pollution. *Epidemiology* 2003;**14**:11-12.
- Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Statist Sci* 1986;**1**:54-77.
- Efron B, Tibshirani R. *An introduction to the bootstrap*. New York: Chapman & Hall, 1993.
- Smith PL. Splines as a useful and convenient statistical tool. *The American Statistician* 1979;**33**:57-62.
- Berkey CS, Hoaglin DC, Antczak-Bouckoms A, et al. Meta-analysis of multiple outcomes by regression with random effects. *Statist Med* 1998;**17**:2537-50.
- Berkey CS, Hoaglin DC, Mosteller F, et al. A random-effects regression model for meta-analysis. *Statist Med* 1995;**14**:395-411.
- Dominici F, Daniels M, Zeger SL, et al. Air pollution and mortality: estimating regional and national dose-response relationships. *J Am Stat Assoc* 2002;**97**:100-11.



## Investigating the dose-response relation between air pollution and total mortality in the APHEA-2 multicity project

E Samoli, G Touloumi, A Zanobetti, A Le Tertre, Chr Schindler, R Atkinson, J Vonk, G Rossi, M Saez, D Rabczenko, J Schwartz and K Katsouyanni

*Occup Environ Med* 2003 60: 977-982  
doi: 10.1136/oem.60.12.977

---

Updated information and services can be found at:  
<http://oem.bmj.com/content/60/12/977>

---

*These include:*

### References

This article cites 17 articles, 6 of which you can access for free at:  
<http://oem.bmj.com/content/60/12/977#BIBL>

### Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

---

### Topic Collections

Articles on similar topics can be found in the following collections

[Air pollution, air quality](#) (207)  
[Other exposures](#) (1023)

---

### Notes

---

To request permissions go to:  
<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:  
<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:  
<http://group.bmj.com/subscribe/>