Expert Judgment Assessment of the Mortality Impact of Changes in Ambient Fine Particulate Matter in the U.S.

HENRY A. ROMAN,*,† KATHERINE D. WALKER,† TYRA L. WALSH,† LISA CONNER,‡ HARVEY M. RICHMOND,‡ BRYAN J. HUBBELL,‡ AND PATRICK L. KINNEY§

Industrial Economics, Incorporated, Cambridge, Massachusetts 02140, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711, and Mailman School of Public Health, Columbia University, New York, New York 10032

Received June 11, 2007. Revised manuscript received November 16, 2007. Accepted December 16, 2007.

In this paper, we present findings from a multiyear expert judgment study that comprehensively characterizes uncertainty in estimates of mortality reductions associated with decreases in fine particulate matter ($PM_{2.5}$) in the U.S. Appropriate characterization of uncertainty is critical because mortality-related benefits represent up to 90% of the monetized benefits reported in the Environmental Protection Agency's (EPA's) analyses of proposed air regulations. Numerous epidemiological and toxicological studies have evaluated the $PM_{2.5}$—mortality association and investigated issues that may contribute to uncertainty in the concentration—response (*C*—*R*) function, such as exposure misclassification and potential confounding from other pollutant exposures. EPA's current uncertainty analysis methods rely largely on standard errors in published studies. However, no one study can capture the full suite of issues that arise in quantifying the *C*—*R* relationship. Therefore, EPA has applied state-of-the-art expert judgment elicitation techniques to develop probabilistic uncertainty distributions that reflect the broader array of uncertainties in the *C*—*R* relationship. These distributions, elicited from 12 of the world's leading experts on this issue, suggest both potentially larger central estimates of mortality reductions for decreases in long-term $PM_{2.5}$ exposure in the U.S. and a wider distribution of uncertainty than currently employed in EPA analyses.

Introduction

The U.S. Environmental Protection Agency's (EPA's) regulatory analyses of rules affecting air quality have historically estimated substantial benefits in the form of avoided premature mortality resulting from reduced exposure to fine particulate matter 2.5 μm in diameter and smaller ($PM_{2.5}$)—tens of thousands of adult deaths are avoided annually due to the entire Clean Air Act, for example (*1*). These benefits comprise up to 90% of EPA's total estimated monetized benefits in regulatory analyses (*1, 2*). A comprehensive characterization of uncertainty in the avoided mortality estimate is thus critical for well-informed decision making by regulators.

While EPA's avoided mortality estimates are based on an extensive epidemiological and toxicological literature base (*3*), uncertainties remain as to the true value of the mortality impact of changes in $PM_{2.5}$ exposure. Some of these uncertainties include questions about the accuracy of exposure characterization in the relevant epidemiologic studies, potential effect modification, the true shape of the concentration–response (*C*—*R*) function, and the strength of evidence for a causal mechanism. EPA has historically assessed these uncertainties through the use of sensitivity analyses and by applying the statistical error presented in selected studies for analysis (*2, 4*). However, no one epidemiological or toxicological study is able to capture the full suite of issues that arise in quantifying the *C*—*R* relationship between $PM_{2.5}$ and mortality.

With the encouragement of the National Research Council (NRC), in its 2002 review of EPA's benefits assessment methodology (*5*), EPA has been exploring methods for developing more comprehensive probabilistic uncertainty distributions for key inputs to benefits analyses, including the use of formal elicitation of expert judgments. Since 2003, the authors have supported EPA's efforts to apply expert judgment elicitation methods to characterize uncertainty in the $PM_{2.5}$—mortality relationship, first in a pilot study (*6*) and more recently in a full-scale study of 12 experts completed in 2006. This article describes the full-scale study and its results. Additional details can be found in the Supporting Information and on EPA's Web site (www.epa.gov/ttn/ecas/benefits.html).

Materials and Methods

This study followed standard best practices for expert elicitations based on the body of literature accumulated over the past 2 decades. As shown in Figure 1, these include explicit criteria for expert selection, a detailed interview protocol, briefing materials provided to experts in advance of the interview, and expert workshops prior to and following the elicitation (*7–12*).

Elicitation Protocol. An extensive written protocol was developed to provide a clear statement of the questions to be answered, document critical underlying assumptions, and establish a logical structure for the elicitation interview. The main goal of the protocol was to the answer the following question:

What is your estimate of the true percent change in annual, all-cause mortality in the adult U.S. population resulting from a permanent 1 $\mu g/m^3$ reduction in annual average ambient $PM_{2.5}$ across the U.S.? In formulating your answer, please consider mortality effects of reductions in both long-term and short-term exposures. To characterize your uncertainty in the *C*—*R* relationship, please provide the minimum, the fifth, 25th, 50th, 75th, and 95th percentiles and the maximum of the effect estimate.

The protocol also contained a comprehensive and detailed set of conditioning questions that established a foundation for the quantitative judgments.

Experts were asked to provide a national average *C*—*R* function for adults 18 and older exposed to annual average $PM_{2.5}$ levels between 4 and 30 $\mu g/m^3$ that could be applied

* Corresponding author phone: 617-354-0074; fax: 617-354-0463; e-mail: har@indecon.com.
† Industrial Economics, Inc.
‡ U.S. Environmental Protection Agency.
§ Columbia University.

10.1021/es0713882 CCC: $40.75

Published on Web 02/27/2008

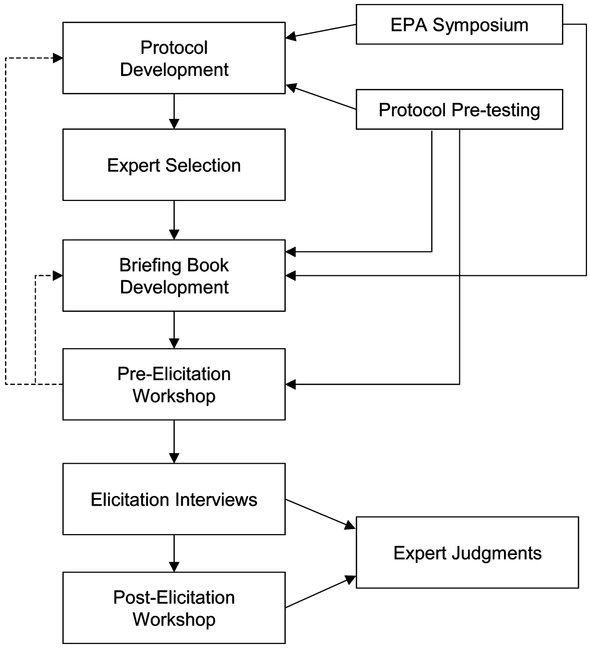


FIGURE 1. Overview of the expert judgment process.

throughout the U.S. in a benefits analysis. Other assumptions specified that the reduction in $PM_{2.5}$ resulting from regulatory action was immediate and permanent and that the action achieved proportional reductions in all PM components.

Figure 2 shows the three-part structure of the elicitation protocol. The first part promoted careful examination and discussion of the key quantitative question. The second part consisted of a systematic discussion of factors to consider when characterizing the relationship between $PM_{2.5}$ exposure and premature mortality. In the final part, experts answered the quantitative question by first specifying a functional form for the *C*−*R* function and then providing a distribution for the slope of that function (i.e., the percent change in mortality per unit change in annual average $PM_{2.5}$).

We elicited the distribution by asking experts to assign a value to the percentiles specified above. Experts were also given the option to provide fewer percentiles (e.g., fifth, 50th, and 95th), which were then fit to an expert-specified parametric distribution in real time using Crystal Ball (CB) software, a forecasting and risk analysis program that utilizes Monte Carlo simulation. To minimize experts' use of the "anchoring and adjustment" heuristic (*13*), the elicitation began with the tails of the distribution (minimum, maximum, fifth, and 95th percentiles) before eliciting central values. As shown in Figure 2, the protocol was iterative, revisiting the expert's responses to conditioning questions to ensure his distribution was consistent with his previous responses.

The protocol was reviewed by PM experts at an EPA symposium and by an interagency team consisting of EPA and Office of Management and Budget (OMB) employees. It was then pretested prior to the interviews with two EPA experts in PM-related mortality, a toxicologist and an epidemiologist.

Expert Selection. The expert selection process sought to create a panel of experts who collectively represent a reasonably balanced range of respected scientific expertise and opinion on the study topic. Other objectives included the use of an explicit and reproducible process that was cost-effective and straightforward to execute and that minimized the level of control of the researcher conducting the elicitation (*14*). Experts were selected in two phases, both of which relied on a peer nomination process.

In the first phase, nominators were identified through a literature search and publication count. The 32 authors with the greatest total number of publications as first, second, or last author were asked to provide nominations. To encourage nomination of a broad array of experts, we randomly assigned nominators to four groups, each of which sought experts with a distinct expertise. Nominators were provided with a set of group-specific criteria to assist them in addition to a set of general criteria common to all groups (see Table 1). Twenty-five sets of nominations were received (at least five from each group).

Experts were ranked by the number of peer nominations received within each group and overall. We selected nine experts—the top two nominees from each group, plus the next most highly nominated individual overall. Invited experts who were unwilling or unable to participate were replaced with the next most highly nominated candidate in that expert's group, provided they were nominated by at least half of the respondents. Otherwise, the expert was replaced with the next most highly nominated expert overall.

The first phase, which had an acceptance rate of 75%, yielded nine experts; however, the panel exhibited less diversity in expertise than expected (eight epidemiologists and one toxicologist). To increase representation of the biological, medical, and toxicological disciplines, we obtained additional expert nominations, using the same general nomination criteria, with assistance from the Health Effects Institute (HEI), a nonprofit organization funded by EPA and industry to study the health effects of air pollution. HEI provided 10 nominees, plus two alternates. HEI was not involved in the study design or execution, beyond providing the additional nominations. Nominees were contacted in random order, with a goal of inviting three additional experts. The acceptance rate for the second phase was 38%, with most who declined citing scheduling conflicts. The final list of experts included the following:

- **Douglas Dockery**, Professor of Environmental Epidemiology, Harvard School of Public Health;
- **Kazuhiko Ito**, Assistant Professor of Environmental Medicine, New York University (NYU) School of Medicine;
- **Daniel Krewski**, Professor of Epidemiology and Community Medicine, University of Ottawa;
- **Nino Künzli**, Associate Professor of Preventative Medicine, University of Southern California Keck School of Medicine;
- **Morton Lippmann**, Professor of Environmental Medicine, NYU School of Medicine;
- **Joe Mauderly**, Vice President and Senior Scientist, Lovelace Respiratory Research Institute;
- **Bart Ostro**, Chief of the Air Pollution Epidemiology Section, California Office of Environmental Health Hazard Assessment;
- **C. Arden Pope III**, Professor of Economics, Brigham Young University;
- **Richard Schlesinger**, Professor of Biology and Heath Sciences, Pace University;
- **Joel Schwartz**, Professor of Environmental Epidemiology, Harvard School of Public Health;
- **George Thurston**, Professor of Environmental Medicine, NYU School of Medicine; and
- **Mark Utell**, Professor of Environmental Medicine, University of Rochester School of Medicine and Dentistry.

To maintain confidentiality, each expert was assigned a randomized letter between A and L to identify his judgments in this paper. We compensated experts using a uniform competitive academic consulting rate, plus travel expenses.

Briefing Book. Each expert was sent a "briefing book" at least 2 weeks before his interview. It included the elicitation protocol, a CD of relevant papers and compendia, recent U.S. data on air quality, health status, population demo-

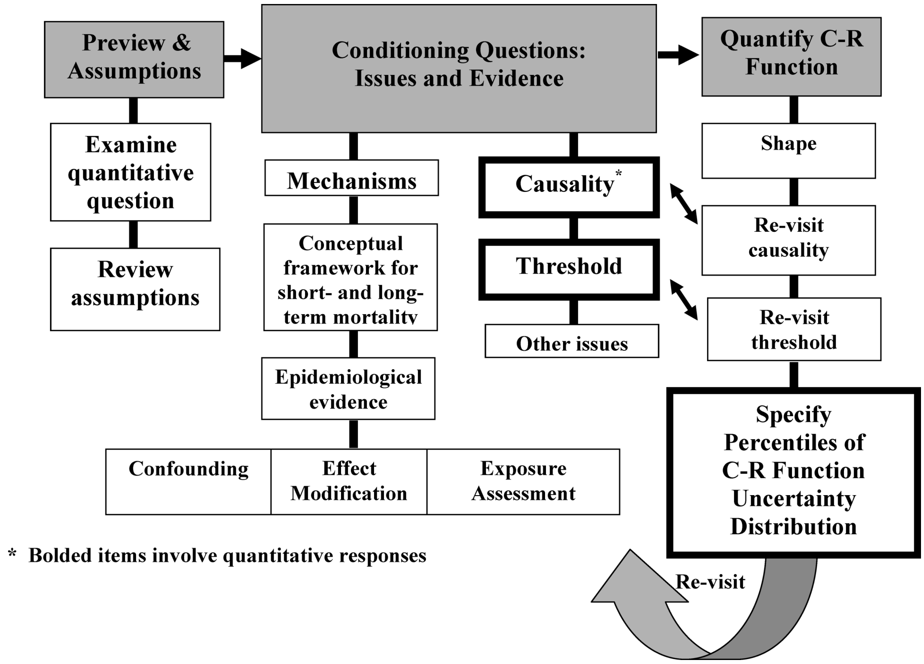


FIGURE 2. Structure of the elicitation protocol.

graphics, and other topics that may factor into the experts' probabilistic judgments, and a document describing factors to consider when providing probability judgments in order to avoid potential sources of bias.

Pre-elicitation Workshop. The experts were invited to a 1 day workshop in January 2006 designed to introduce the project, familiarize them with expert judgment and the elicitation process, and foster critical discussions of key evidence relevant to the questions posed by the study. Promoting consensus was not a goal of the workshop. Nine of the 12 experts participated. A workshop summary, copies of presentations, and papers cited at the workshop were sent to all 12 experts.

Elicitation Interview. The study team elicited judgments from each expert individually during a personal interview. The elicitations were conducted over a 4 month period in the spring of 2006. The elicitation team included two interviewers, one experienced in the elicitation of expert judgments (Dr. Walker) and one with expertise in PM health effects and exposure (Dr. Kinney).

During the day-long interviews, experts were asked to think systematically about, and cite, the evidence in support of their responses. The team encouraged experts to evaluate the robustness of their judgments by considering evidence that might support an opposing or alternative position, as well as sources of uncertainty, error, or bias that might challenge their interpretation of the evidence.

Each expert was also given the opportunity to participate in Internet-based conferencing with Industrial Economics representatives (Mr. Roman and/or Ms. Walsh) who provided real-time graphical and quantitative feedback regarding the expert's $C-R$ function distribution and responses to protocol questions. The expert could visualize his distribution, compare it to those from key epidemiological studies, and compare the mortality implications of his distribution with data on other major causes of death in the U.S.

Following each interview, the expert was sent a summary of his qualitative and quantitative judgments for review, adjustment, and/or confirmation of his responses.

Post-Elicitation Workshop. The study team held a final 4 hour workshop with the experts in June 2006 to share results anonymously with the group, highlight areas where expert opinion varied, clarify points of confusion, allow experts to raise issues for discussion, and encourage each expert to critically review his judgments. The purpose was not to promote consensus. Eleven of the 12 experts participated at least partially in the workshop.

Following this workshop, a meeting summary was sent to the experts, and they were provided an opportunity to revise their qualitative and/or quantitative judgments privately using a standardized form. Four experts chose to make changes.

Results

The experts' responses exhibit substantial agreement regarding the nature and cause of mortality associated with $PM_{2.5}$, the likelihood of a causal connection between exposure and mortality, the shape of the $C-R$ function, and the central estimate of the mortality impact. We observed differences in the experts' responses about the relative importance and size of potential sources of uncertainty, and hence, in the spread of the distributions of the effect estimate.

Causes of Death and the Importance of Short-Term versus Long-Term Exposure. During the conditioning section, experts described how they conceptualized the "mortality" that was the focus of the main quantitative question. Ten of 12 experts believed most of the mortality resulted from long-term exposures, citing potential cumulative chronic cardiovascular and respiratory damage from PM exposure leading to increased risk of death. The experts thought these deaths would be primarily cardiovascular, with lesser amounts due to respiratory disease and lung cancer, on the basis of the strength of current evidence. They thought that the short-term mortality impacts that are not included in relative risk (RR) estimates from the cohort studies represented a very small percentage of the total mortality impact. Two experts disagreed. Expert D thought he had insufficient information to discuss relative magnitudes of the two types of impacts, and Expert K expressed greater confidence in evidence showing that mortality impacts reflect changes in short-term peak exposures. However, all experts ultimately relied on results from long-term studies when developing their effect estimates.

Likelihood of a Causal Relationship. After discussing the body of evidence, each expert was asked to estimate

TABLE 1. Expert Nomination Criteria

General Criteria

1. Ideal experts should possess the educational background and/or experience to display a thorough understanding of results from the epidemiological literature addressing the relationship between chronic exposure to $PM_{2.5}$ and mortality and to evaluate these results in the context of other evidence pertinent to the $PM_{2.5}$—mortality issue, such as relevant toxicological and physiological literature
2. Experts may include primary scientific researchers as well as prominent individuals from scientific panels, institutions, journal editorial boards, and other such groups who, through their educational background and experience, are in a position to carefully interpret the key evidence regarding $PM_{2.5}$ exposures and mortality
3. The overall set of experts nominated should be a balanced group that reflects the full range of respected scientific opinions concerning the strength of the evidence linking premature mortality with ambient $PM_{2.5}$ concentrations
4. The nominees should all be based in either the U.S. or Canada

Group Specific Criteria

group 1:
- are the most knowledgeable about the relationship between long-term $PM_{2.5}$ exposures and mortality; and/or
- have studied in-depth the uncertainties and methodological limitations of existing cohort studies on $PM_{2.5}$ and mortality

group 2:
- have made the most significant contributions to our understanding of the potential underlying biological mechanisms of the $PM_{2.5}$—mortality relationship; and/or
- have made the most significant contributions to our understanding of the likelihood of a causal relationship between $PM_{2.5}$ and mortality

group 3:
- display significant experience analyzing the relationship between $PM_{2.5}$ and mortality through participation in expert committees and workshops and/or publication of review articles; and/or
- display significant experience analyzing and applying the PM—mortality literature within a risk assessment and/or policy context

group 4:
- are conducting innovative, cutting-edge research investigating the relationship between $PM_{2.5}$ and mortality; and/or
- have made the most significant contribution to our understanding of the relationship between health effects and $PM_{2.5}$ exposures

quantitatively the probability that the relationships between short-term and/or long-term $PM_{2.5}$ exposures and mortality were causal. He was asked to provide a "best estimate" plus a range of probabilities to indicate his level of certainty.

The responses, presented at the bottom of Figure 3, show that 10 of the 12 experts placed similarly high likelihoods on a causal relationship. Two experts (G and K) were significantly less certain.

Shape of the Concentration–Response Function. Eight experts thought the true $C-R$ function relating mortality to changes in annual average $PM_{2.5}$ was log-linear across the entire study range ($\ln(\text{mortality}) = \beta \times \text{PM}$). Four experts (B, F, K, and L) specified a "piecewise" log-linear function, with different β coefficients for PM concentrations above and below an expert-specified break point. This approach allowed them to express increased uncertainty in mortality effects seen at lower concentrations in major epidemiological studies. Expert K thought the relationship would be log-linear above a threshold.

Quantification of the Concentration–Response and Associated Uncertainty. When developing their uncertainty distribution for the $C-R$ coefficient, seven experts (B, D, E, G, I, K, and L; "group 1") characterized uncertainty in the mortality effect conditional on the assumption that the relationship was causal, while the others (A, C, F, H, and J; "group 2") integrated their uncertainty about the likelihood of a causal relationship directly into their distributions. The uncertainty distributions for group 1 thus reflect sources of uncertainty other than those affecting the experts' judgments about whether or not the causal relationship exists. The experts in group 1 preferred that their causality judgments and $C-R$ distributions be presented separately and combined only in the context of a benefits analysis.

Four experts (B, F, H, and L) developed their $C-R$ distributions directly by providing the requested percentiles of their distributions. Seven of the remaining experts developed their distributions indirectly by specifying two or more percentiles, or a mean and standard deviation, and then fitting a parametric distribution with CB. Lastly, expert I quantified his uncertainty by choosing three studies, assigning subjective weights to each, combining their results via Monte Carlo simulation in CB, and fitting a final parametric distribution to the outcome. Of the eight (including expert I) who selected parametric distributions, six fit normal distributions (each fitting to different percentiles), one chose triangular, and one a Weibull.

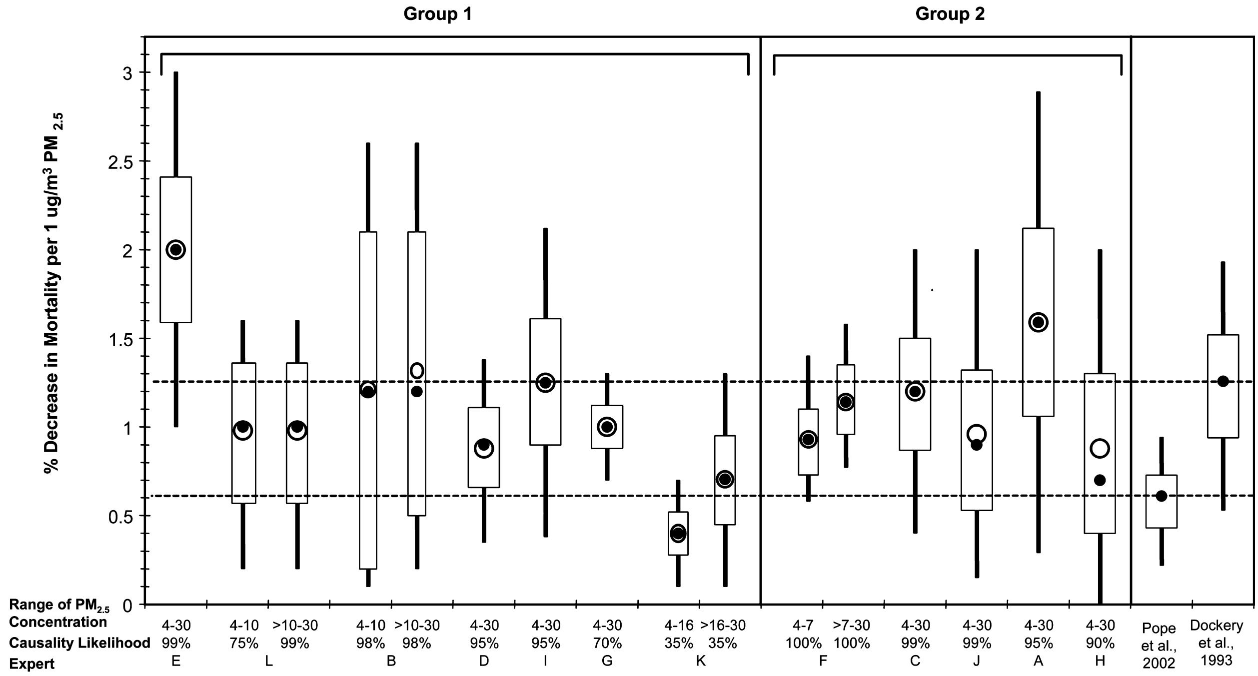
Figure 3 presents box plots of the experts' distributions for groups 1 and 2. In addition to the median (closed circle) and other percentiles specified in the protocol, the plots include a mean value (open circle) estimated from the elicited parameters using CB. The distributions within each group are arrayed in order of decreasing elicited likelihood of a causal relationship.

The results in each group show substantial variation in the amount of uncertainty expressed, even among individuals expressing similar views on the strength of the causal relationship. For example, in group 1, expert B specified a much wider range of values than expert L, although both specified similar causal likelihoods. In group 2, expert F was both highly certain of the likelihood of a causal relationship (100%) and of his ability to predict the magnitude of the relationship, specifying a much narrower range of values than expert A or H, whose causal likelihoods were also very high.

Figure 3 displays pairs of distributions for the four experts who specified two log-linear segments across the PM study range. The transition point between the segments ranged from 7 to 16 $\mu g/m^3$, with most experts basing their choice on the range of PM observed in the major cohort studies.

The differences in the $C-R$ slope and uncertainty expressed by the experts above and below their transition points were generally modest. Expert L's distributions differ only in their minima—zero for the lower versus 0.02 for the upper (not shown)—and in his expressed confidence in a causal relationship, which was higher in the upper range (99% versus 75%). Expert F reduced his median estimate for the lower range by 18% although the spread was similar. Expert K's segments differed most; he reduced his median 43% and decreased the spread of his lower distribution.

Expert K also applied a threshold, T, to his function, which he described probabilistically. He specified $P(T > 0) = 0.5$. Given $T > 0$, he indicated $P(T \leq 5\mu g/m^3) = 0.8$ and $P(5\mu g/m^3 < T \leq 10\mu g/m^3) = 0.2$. Figure 3 does not include the impact of applying expert K's threshold, as the size of the reduction in benefits will depend on the distribution of baseline PM levels in a benefits analysis.



Key: Closed circle = median; Open circle = mean; Box = interquartile range; Solid line = 90% credible interval

FIGURE 3. Uncertainty distributions for the $PM_{2.5}$—mortality *C*—*R* coefficient for annual average $PM_{2.5}$ concentrations of 4—30 μg/m^3 Note: Box plots represent distributions as provided by the experts to the elicitation team. Experts in group 1 preferred to give conditional distributions and keep their probabilistic judgment about the likelihood of a causal or noncausal relationship separate. Experts in group 2 preferred to give distributions that incorporate their likelihood that the $PM_{2.5}$—mortality association may be noncausal. Therefore, the expert distributions from these two groups are not directly comparable.

Figure 3 also includes box plots for studies EPA has applied previously to generate primary benefits estimates (*4*) and to conduct sensitivity analysis (*15*). Most of the subjective distributions encompass the 90% confidence intervals from both studies. Most of the experts' central estimates fall at or above the 2002 American Cancer Society study (ACS) median (0.6% per μg/m^3) and below the original Six Cities median (1.2% per μg/m^3).

Numerous factors influenced the development of each expert's distribution, as described at length in their interview summaries. However, some general patterns emerged in the weight experts gave to particular studies and to adjustments they made to account for particular concerns about those studies.

Table 2 summarizes the studies that the experts cited when developing their distributions. The dotted circle [⊙] indicates those studies from which the experts drew in developing their central (median) effect estimates or which they used to support adjustments to those estimates. The open circle [○] indicates studies mentioned in association with one or more of the percentiles used to characterize uncertainty.

Median effect estimates drew heavily on four studies: the Pope et al. 2002 ACS follow-up (*4*), the Jerrett et al. 2005 analysis of the ACS cohort in Los Angeles (ACS LA, *16*), the 1993 Dockery et al. Six Cities study (*15*), and the 2006 Laden et al. Six Cities follow-up (*17*). Although several experts mentioned the Krewski et al. 2000 (*18*) reanalysis of the original ACS and Six Cities studies when discussing effect modification and/or confounding, experts' quantitative judgments tended to reflect the influence of the more recent follow-up studies with larger data sets.

Discussions during the conditioning phase of the interview raised questions about a number of issues, including the impacts of effect modification, exposure measurement, and confounding. Two factors, in particular, influenced how the experts weighed the findings from the two principal cohort studies (ACS and Six Cities). First, several experts thought that the original ACS cohort underrepresented individuals with less than a high school education with respect to the U.S. population at the time of the study (*18, 19*). Second, several experts believed the ACS LA analysis raised questions about the impact of potential exposure misclassification in the original ACS study resulting from the larger-scale spatial resolution of exposure. Both factors could support the argument that the ACS study results might underestimate the total adult U.S. PM—mortality effect.

Most experts adjusted their median estimates to reflect the impact of educational attainment, exposure misclassification, or both. Others saw the ACS LA findings as rationale for placing more confidence in the original Six Cities study and its extended analyses than they previously had, or to rely more extensively on the ACS LA study. Both approaches tended to produce similar central estimates (approximately 1% per μg/m^3). Median effect estimates were lower when experts argued against adjusting the original ACS estimates for one of these factors (e.g., expert H did not think it appropriate to adjust for educational attainment; expert K adjusted for neither).

Although the experts mentioned a number of potential confounders in the interviews, only three (D, F, and G) made direct adjustments to their quantitative judgments to account for confounding (by occupational exposure and/or copollutants). Most argued that the key studies discussed above had adequately controlled for potential confounders. They noted that the persistence of a $PM_{2.5}$—mortality effect across multiple studies using different statistical designs made it difficult to believe that it could be caused solely by confounding.

Short-term exposure studies (i.e., time series) were rarely cited in the development of experts' distributions, even among those experts who generally found the database more compelling for the effects of short-term exposures than for long-term exposures. They were most often cited to define the lower end of the distributions. This appears consistent with the views of those experts who argued that mortality from short-term exposures likely comprise a small proportion of total $PM_{2.5}$-related deaths.

TABLE 2. Epidemiologic Studies Relied upon by Experts in Creating Their *C−R* Coefficient Uncertainty Distributions[a]

	ACS (Pope et al. (*4*))	ACS LA anal (Jerrett et al. (*16*))	Six Cities (Dockery et al. (*15*))	Six Cities (Laden et al. (*18*); cross-sectional)	ACS (Pope et al., 1995)[b]	Netherlands cohort study (Hoek et al., 2002)[b]	Six Cities (Laden et al. (*18*); change estimate)	Mallick et al., 2002[b]	Willis et al., 2002[b]	NMMAPS (Samet et al., 2000)[b]	Women's Health Initiative[c]	AHSMOG (Abbey et al., 1991 and 1999; McDonnell et al., 2000; Chen et al., 2005)[b]
expert A	○	⊙	○	⊙			○			○		
expert B	⊙	⊙			⊙ ○			⊙				
expert C	⊙	⊙	⊙									
expert D	⊙		⊙									
expert E	⊙	⊙		⊙		⊙	⊙		⊙			
expert F	○	○	⊙	⊙								
expert G	⊙○		⊙ ○									
expert H	⊙	⊙ ○	○									
expert I	⊙	⊙		⊙								
expert J[d]	⊙	⊙ ○	⊙ ○			○				○	○	○
expert K	⊙		⊙ ○									
expert L	○	⊙ ○		○						○		
total ⊙:	9	8	6	4	1	1	1	1	1	0	0	0
total ○:	4	4	5	1	1	1	1	0	0	3	1	1

[a] ⊙ = Expert used the study to inform the median of his *C−R* coefficient distribution(s). ○ = Expert used the study to inform the uncertainty of his *C−R* coefficient distribution(s). [b] Full study citation can be found in the Supporting Information. [c] Study not yet published at the time of the interview. [d] Expert J also cited a number of additional short-term studies as support for his uncertainty; see the Supporting Information for more information.

Discussion

This study created a structured approach through which individual scientists could characterize uncertainty in the $PM_{2.5}$−mortality relationship in a probabilistic form. They were encouraged to take into account numerous factors that to date have not been incorporated into uncertainty estimates from previous EPA benefits assessments, such as the potential for nonlinearities in the relationship over the range of $PM_{2.5}$ concentrations found in the U.S. They were also encouraged to consider the likelihood of a causal relationship, on the basis of the body of scientific evidence from epidemiology, toxicology, clinical medicine, and other relevant fields. Finally, the experts were able to account for the impact of limitations in individual studies and in the overall body of scientific evidence when developing their uncertainty distributions.

The study had a number of strengths that set it apart from previous studies of its kind. It included a large expert panel representing a range of respected scientific opinion and encouraged extensive interexpert communication through the use of pre- and post-elicitation workshops that reviewed project goals, the elicitation process, and key data. The use of a structured protocol, which underwent extensive external review, allowed experts flexibility in eliciting both the shape of the *C−R* function and values for the uncertainty distributions. Finally, experts were able to use Internet teleconferencing coupled with spreadsheet models and CB probabilistic modeling software during the interview to assist them in visualizing and evaluating their responses.

The study produced a set of judgments that individually and collectively provide a more comprehensive view of scientific uncertainty in the $PM_{2.5}$−mortality relationship. Most of the experts' distributions were broader than those defined by the statistical confidence intervals reported in the 2002 ACS study. The central estimates from most experts also exceeded the 2002 ACS study estimate, reflecting the quantitative influence of new research and analysis (*16, 19*).

We explored the sensitivity of our results to changes in the expert pool and in the elicitation process, using a simplified benefits analysis for a change in annual average $PM_{2.5}$ concentrations in the U.S. from 12 to 11 $\mu g/m^3$. We used CB to generate a large sample (10000 trials) of benefit estimates from each expert's effect distribution, combined the samples, and estimated the percentiles of the resulting data set. Our results were not substantially changed by any of the factors evaluated, including removal of individual experts, the use by experts of parametric or nonparametric approaches to characterize uncertainty, participation in the pre-elicitation workshop, and the decision by some experts to change their judgments following the post-elicitation workshop. Individual experts, rather than any process differences, had the greatest influence; removal of expert E shifted the mean pooled benefits downward by 8%, and the removal of expert K shifted the mean upward by 8%.

Despite the robustness of its findings, this study has limitations that are important to consider in any application of its results. For instance, the quantitative question answered by the experts was premised on a set of specific assumptions specified in the protocol regarding issues such as baseline and historical exposure conditions, $PM_{2.5}$ composition, and the characteristics of the U.S. population, potentially limiting the generalizability of the results to other contexts. In addition, the expert selection process yielded a group of experts representing a range of respected scientific opinions on the PM−mortality issue, not a statistical sample of expert opinion. Therefore, these results should not be used to assess the prevalence of individual opinions in the scientific community.

Defining the requisite balance of expertise necessary to address a complex question like this one is controversial. Characterizing uncertainty ideally requires integrating elements of multiple disciplines, and our panel was more heavily weighted toward epidemiological expertise. While it is unclear a priori how an expert's background would influence their judgment, diversity may have provided additional perspectives. However, in our experience, the experts who participated exhibited a high degree of knowledge not only in their specific field but also with regard to issues of mechanism, exposure, and statistics as they relate to the PM–mortality relationship. Therefore, we think that the lack of diversity in the expert panel does not degrade the results of this study. An option for addressing this concern in future studies would be to disaggregate key judgments by discipline and employ self-assigned or peer-assigned weights for each judgments, such as those used in Evans et al. (*10*), to provide perspective on the expertise and contribution of individual experts to the overall question.

The study team opted to keep the experts' judgments anonymous. Confidentiality is typically offered in elicitations to allow experts the freedom to express candid, independent opinions (*8–11*, *14*). While anonymity might increase the likelihood of motivationally biased answers, we believe that the benefits of anonymity outweigh the potential for motivational bias. In addition, the elicitation team extensively queried each expert during his interview to establish clear foundations for each of his judgments and to highlight potential inconsistencies in his judgments, in an effort to limit motivationally biased responses.

The integration of each expert's judgment about the likelihood of a causal relationship with his mortality effect distribution proved more difficult than expected. The experts found it particularly challenging to cognitively integrate (or separate) some of the issues that factor into characterizing uncertainties in $PM_{2.5}$ effects on mortality from judgments about the likelihood of a causal relationship. In a few cases, this led to apparent inconsistencies in expert's distributions (e.g., expert A's nonzero fifth percentile in a distribution that incorporates 95% confidence in a causal relationship). We believe that a discussion about the likelihood of a causal relationship is an important factor in assessing uncertainty about *C–R* functions, though one that merits further research regarding how best to elicit joint or conditional distributions of effect.

Overall, this expert judgment study of the mortality impacts of $PM_{2.5}$ represents a major advance for EPA in its efforts to more fully characterize uncertainty in the benefits associated with its air quality regulations. It represents the most comprehensive effort to date to describe in probabilistic terms the uncertainty in the impacts of long-term PM exposures on adult mortality in the U.S., based on an in-depth review and critique of theory and literature by some of the world's leading experts on this issue. The distributions provided by the experts and their associated rationales offer, in quantitative terms, an unusual glimpse of both the common ground and the deep divisions that exist among competent scientists on this extensively researched issue. We expect this uncertainty assessment will serve as an example to EPA and other agencies seeking to improve benefits analysis and will enrich future regulatory debate and decision-making.

Acknowledgments

Support for this work was provided by the EPA's Office of Air Quality Planning and Standards. We thank members of the EPA project team, including Robert Hetes, Mary Ross, and Zachary Pekar as well as staff of the Office of Management and Budget for their suggestions. In addition, we thank Jim Neumann of Industrial Economics and Dr. Andrew Wilson, formerly of the Harvard Center for Risk Analysis, for their contributions. Disclaimer: Although the information described in this article has been funded wholly or in part by the United States Environmental Protection Agency under Contract EP-D-04-006 to Industrial Economics, Inc., it does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

Supporting Information Available

Full detailed report outlining the study methodology and results, a copy of the protocol, example briefing book materials, a description of the sensitivity analysis, and summaries of each expert interview. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) U.S. Environmental Protection Agency. *The Benefits and Costs of the Clean Air Act 1990–2010*; EPA: Washington, D.C., 1999.
(2) U.S. Environmental Protection Agency. *Regulatory Impact Analysis for the Final Clean Air Interstate Rule*; EPA: Washington, D.C., 2005.
(3) U.S. Environmental Protection Agency. *Air Quality Criteria for Particulate Matter*; EPA: Washington, D.C., 2004.
(4) Pope, C. A., III; Burnett, R. T.; Thun, M. J.; Calle, E. E.; Krewski, D.; Ito, K.; Thurston, G. D. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA, J. Am. Med. Assoc.* **2002**, *287*, 1132–1141.
(5) National Research Council. *Estimating the Public Health Benefits of Proposed Air Pollution Regulations*. The National Academies Press: Washington, D.C., 2002.
(6) Industrial Economics, Inc. *An Expert Judgment Assessment of the Concentration–Response Relationship between $PM_{2.5}$ Exposure and Mortality*; EPA: Research Triangle Park, NC, 2004.
(7) Cooke, R. M. *Experts in Uncertainty: Opinion and Subjective Probability in Science*; Oxford University Press: New York, NY, 1991.
(8) Morgan, M. G.; Morris, S. C.; Henrion, M; Amaral, D. A. L.; Rish, W. R. Technical uncertainty in quantitative policy analysis—Sulfur air pollution example. *Risk Anal.* **1984**, *4*, 201–216.
(9) Morgan, M. G.; Pitelka, L. F.; Shevliakova, E. Elicitation of expert judgments of climate change impacts on forest ecosystems. *Clim. Change* **2001**, *49*, 279–307.
(10) Evans, J. S.; Gray, G. M.; Sielken, R. L.; Smith, A. E.; Valdez-Flores, C.; Graham, J. D. Use of probabilistic expert judgment in uncertainty analysis of carcinogenic potency. *Regul. Toxicol. Pharmacol.* **1994**, *20*, 15–36.
(11) Walker, K. D.; Evans, J. S.; MacIntosh, D. Use of expert judgment in exposure assessment. Part I. Characterization of personal exposure to benzene. *J. Exposure Anal. Environ. Epidemiol.* **2001**, *11*, 308–322.
(12) Walker, K. D.; Catalano, P.; Hammitt, J. K.; Evans, J. S. Use of expert judgment in exposure assessment. Part 2. Calibration of expert judgments about personal exposures to benzene. *J. Exposure Anal. Environ. Epidemiol.* **2003**, *13*, 1–16.
(13) Kahneman, D.; Slovic, P.; Tversky, A. *Judgment under Uncertainty: Heuristics and Biases*; Cambridge University Press: Cambridge, U.K., 1982.
(14) Hawkins, N. C.; Graham, J. D. Expert scientific judgment and cancer risk assessment: A pilot study of pharmacokinetic data. *Risk Anal.* **1990**, *8*, 615–625.
(15) Dockery, D. W.; Pope, C. A., III; Xu, X.; Spengler, J. D.; Ware, J. H.; Fay, M. E.; Ferris, B. J.; Speizer, F. E. An association between air pollution and mortality in six U.S. cities. *N. Engl. J. Med.* **1993**, *329*, 1753–1759.
(16) Jerrett, M.; Burnett, R. T.; Ma, R.; Pope, C. A., III; Krewski, D.; Newbold, K. B.; Thurston, G.; Shi, Y.; Finkelstein, N.; Calle, E. E.; Thun, M. J. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* **2005**, *16*, 1–10.
(17) Laden, F.; Schwartz, J.; Speizer, F. E.; Dockery, D. W. Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. *Am. J. Respir. Crit. Care Med.* **2006**, *173*, 667–672.
(18) Krewski, D.; Burnett, R. T.; Goldberg, M. S.; Hoover, K.; Siemiatycki, J.; Jerrett, M.; Abrahamowicz, M.; White, W. H. *Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality*. Health Effects Institute: Boston, MA, 2000.
(19) Pope, C. A., III; Dockery, D. W. Health effects of fine particulate air pollution: lines that connect. *J. Air Waste Manage. Assoc.* **2006**, *56*, 709–742.

ES0713882