Particulate Air Pollution Standards and Morbidity and Mortality: Case Study

Daniel S. Greenbaum,1 John D. Bachmann,2 Daniel Krewski,3 Jonathan M. Samet,4 Ronald White,5 and Ronald E. Wyngaard6

EPIDEMIOLOGY AT THE CENTER OF A HIGHLY-VISIBLE NATIONAL DEBATE

In July 1997, the US Environmental Protection Agency (EPA), under the auspices of the federal Clean Air Act, promulgated National Ambient Air Quality Standards for particulate matter, retaining the previously existing standards for PM10 (particles less than or equal to a nominal 10 micrometers in aerodynamic diameter) with minor revisions, and adding new annual and 24-hour standards for PM2.5 (1). This action was based primarily on epidemiologic studies finding an association between particulate matter and morbidity and mortality, studies which reasonably consistently found increases in mortality, hospitalization, and other morbidity endpoints associated with daily increases in particulate matter, as well as increased risk of premature death associated with particulate matter levels in larger, longer-term cohort studies (2). The EPA decision capped the most recent chapter in the more than 30-year discussion and debate on the epidemiologic evidence on particulate matter-related health effects, and on the public health policy actions to be implemented based on that evidence. This action, however, has not ended the issue. As required by the Clean Air Act, the EPA will review the standard again by July 2002 (3), and the continuing debate and interest has sparked substantial further investment in research (4).

This case study looks back over this decades-long history of the use of epidemiologic evidence in public policy, gleaned from that history and the most recent debate lessons for the design and conduct of epidemiologic studies and the dissemination of findings. This case study raises some unique issues, such as the use of epidemiologic evidence for policy in the absence of evidence of biologic plausibility from toxicology, and the response to calls in the debate for public access to and reanalysis of epidemiologic data. The case study also provides guidance for epidemiologists to recognize and balance the multiple roles they are often asked to play in such debates.

THE CASE STUDY PARTICIPANTS AND THE APPROACH

The team assembled to prepare this case study brought together people from a variety of disciplines who had played important and diverse roles in the debate, and who offered substantial historical, technical, and policy knowledge of the science and issues. The authors first reviewed the major documents produced over the course of the debate that summarized the existing scientific evidence and attempted to draw from that evidence recommendations for policy. The authors then reviewed the history of the development of the epidemiologic science, and of the national air quality standards for particulate matter. The authors then addressed the following questions:

1. How has the research on particulate matter been initiated and funded?
2. How have the findings been synthesized and evaluated?
3. How has the issue of biologic plausibility been considered?
4. What has been the role of science supported by opponents and proponents in the discussion and the interpretation of the science?
5. What has been the role of replication and reanalysis of epidemiologic studies?
6. What has been the role of the epidemiologist?

HISTORY OF DEVELOPMENT AND REVIEW OF US PARTICULATE MATTER STANDARDS

While regulatory response to the health effects of combustion products can be traced back to the 14th Century (5), concerns in the 20th century arose from dramatic air pollution episodes with excess mortality in the Meuse Valley, Belgium (6), Donora, Pennsylvania (7), and London, England (8). These episodes provided unequivocal evidence that high levels of air pollution, containing a mix of particulate and gaseous combustion products, produced adverse
effects on population health and led to calls for remediation and an interest in organized examination of the health effects of less severe pollution elevations (2).

Subsequently, during the 1950s and 1960s, observational studies on air pollution of primarily cross-sectional or ecological design were conducted. The results of these studies formed a major basis for the first federal government document describing the health effects of particulate matter and sulfur oxides as prepared by the EPA’s predecessors (a "criteria document") (9). The early scientific and policy interest stimulated by these pollution episodes framed subsequent examination of the issue in terms of a mixture of particles as measured by various devices and ensured the prominence of observational evidence in supporting policy decisions.

Legislative and policy context for particulate matter and other US air quality standards

The 1970 passage of the Clean Air Act prompted establishment of National Ambient Air Quality Standards for all pollutants for which "criteria" existed, including particulate matter (others are now sulfur dioxides, nitrogen dioxides, carbon monoxide, photochemical oxidants, and lead). Two sections of the Act govern the establishment, review, and revision of the National Ambient Air Quality Standards. Section 108 (42 USC 7408) directs the Administrator to identify pollutants which "may reasonably be anticipated to endanger public health and welfare" and to issue air quality criteria for them. These air quality criteria are to "accurately reflect the latest scientific knowledge useful in indicating the kind and extent of all identifiable effects on public health or welfare which may be expected from the presence of [a] pollutant in the ambient air...."

Section 109 (42 USC 7409) directs the Administrator to propose and promulgate “primary” and “secondary” National Ambient Air Quality Standards for pollutants identified under section 108. Section 109(b)(1) defines a primary standard as one “the attainment and maintenance of which, in the judgment of the Administrator, based on the criteria and allowing an adequate margin of safety, [are] requisite to protect the public health.” The EPA has maintained, and reviewing courts have upheld the position, that “[T]he statute and its legislative history make clear that economic considerations play no part in the promulgation of ambient air quality standards under Section 109” (10).

Section 109(d)(1) of the Act requires periodic review and, if appropriate, revision of existing air quality criteria and National Ambient Air Quality Standards. Section 109(d)(2) requires appointment by the EPA Administrator of an independent scientific review committee to review criteria and standards and recommend new standards or revisions of existing criteria and standards, as appropriate. The committee is known as the Clean Air Scientific Advisory Committee (CASAC), a standing committee of the EPA’s Science Advisory Board, and CASAC has generally included epidemiologists among its members.

In conformance with legislative requirements, the process for reviewing the scientific criteria and specifying air quality standards evolved to its present form by 1980. The full specification entails selecting a pollutant indicator (e.g., PM10 or PM2.5), averaging times (e.g., 24 hours, annual), form (e.g., 98th percentile), and levels (e.g., 15 µg/m³). As illustrated in figure 1, the process involves various offices of the EPA, the external scientific community, including CASAC, as well as various stakeholder groups and other members of the public. It begins with an examination of the scientific literature and the development of a revised criteria document by the Office of Research and Development. Drafts of the criteria document are reviewed by CASAC and the public. The EPA staff from the Office of Air and Radiation then develop a staff paper, which evaluates the policy implications of the key studies and scientific information contained in the criteria document. This assessment is intended to help bridge the gap between the scientific review contained in the criteria document and the judgments required of the Administrator in setting National Ambient Air Quality Standards. The staff paper is also reviewed by CASAC and the public.

After the reviews of both the criteria document and the staff paper, CASAC submits a “closure letter” to the Administrator closing their review, and finding that the document fairly and fully represents the current state of scientific knowledge. Based on this review, the Administrator decides provisionally whether revised or new standards are needed. If so, following review by other federal agencies under the auspices of the Office of Management and Budget, the EPA publishes such revisions or new standards in the Federal Register for public comment. Following a formal public comment period, which usually includes public meetings and hearings, the Administrator develops and publishes final decisions, including a summarization of and response to public comments.


Table 1 provides a general chronology of US particulate matter standards from the first standards for total suspended particles set in 1971, through the recent promulgation of standards for PM10 and PM2.5, to the ongoing criteria review. The progression of decisions on the standards has led to a sharper focus—both in research and regulation—on those particles most likely to be responsible for the observed adverse effects. As a first refinement, the review of the standard that began in 1979 and culminated in 1987 resulted in a decision to regulate only particles that penetrated to the thoracic region of the respiratory tract, using PM10 as the indicator, based largely on human dosimetry studies. In 1997, this indicator was amended by adding PM2.5 in recognition of the substantial differences between so-called fine and coarse fraction particles in physicochemical properties and potential health effects. The revision in 1997 also increased the protection afforded by the standards, adopting a substantially more stringent standard for PM2.5 than would have resulted from the regulations under the previous PM10 standard. Throughout this period, three separate scientific criteria documents and addenda, produced by the EPA together by three different CASAC panels, made recom-
recommendations to continue the regulation of particulate matter as a pollutant mixture (11–13).

The ambient levels, the particle size and characteristics of particulate matter-containing air pollution, and the approaches used in epidemiologic studies changed substantially over this period. The dramatic decrease in wintertime urban pollution levels in British and US cities in response to control programs led to the use of more sophisticated and sensitive study designs and to application of emerging, multivariate statistical techniques. By the early 1970s, air pollution control programs had lowered sulphur dioxide and particulate matter levels significantly for urban winters, but increasing regional emissions gave rise to concerns about elevated summertime sulfate and ozone levels in large regions of the United States. Whereas the early studies of acute episodes provided convincing evidence from simple data plots, further interpretation and analysis of multi-year data sets at lower ambient pollution concentrations were based in more sophisticated analytic techniques that tended to call into question the interpretations of original investigators regarding the existence of population thresholds for adverse effects at concentrations on the order of 500 to 1,000 µg/m³. The examination of lower concentrations and attempts to separate effects of sulfur oxides and particulate matter were major concerns during the first criteria and standards review. The 1982 criteria document adopted a set of criteria for evaluating epidemiology studies that would be most useful for “quantitative purposes” and narrowed the list of studies to a relatively small set that excluded some of the cross-sectional studies that had clearly formed the basis for the earlier total suspended particles standards. Because only total suspended particles exposure data were available, none of these studies had used the recommended PM₁₀ indicator, and the results had to be converted from the original metrics, contributing further uncertainty. To highlight the degree of uncertainty, EPA Administrator Ruckleshaus proposed a wide range of levels for the particulate matter standards in 1984, and ultimately chose to set the new National Ambient Air Quality Standards for PM₁₀ at a level comparable to the previous standard for total suspended particles.

In the 1996–1997 review of the standards, new time-series studies and reanalyses of these studies, together with new prospective cohort designs, raised concerns that health effects were occurring at levels allowed by the 1987 standards. Reanalyses of early findings by investigators at Johns Hopkins University selected and funded by the Health Effects Institute, a non-profit organization receiving funding equally from the EPA and the vehicle manufacturers, played a substantial role in resolving issues associated with validity of data in these studies and replicating results using a consistent modeling method (14). Thereafter, the debate became more focused on the issues of confounding, exposure misclassification, and related matters in the interpretation of the epidemiologic results (2, 15–18). The process culminated with promulgation of new standards in 1997, and the subsequent review began almost immediately, with the EPA now having begun a new review of criteria and standards to be completed by 2002 or shortly thereafter.

The period of these reviews also saw a shift in the way research into particulate matter was funded and overseen.

FIGURE 1. Overview of review process for National Ambient Air Quality Standards. *US Environmental Protection Agency. †Clean Air Scientific Advisory Committee.
The EPA’s Community Health and Environmental Surveillance System was, in the 1970s, suggesting potential effects associated with summertime sulfate aerosol and ozone, but concerns raised by scientists with the study design, exposure assessment, and other characteristics of that program called these observations into question (19). As a result of these questions, these studies were of limited use to policy-makers, and in 1977, Congress acted to prohibit use by the EPA of the original Community Health and Environmental Surveillance System studies in future policy deliberations (20). The political attention focused on the problems with the program, and the criticism of the EPA’s role in it, had the unintended consequence of discouraging the EPA from pursuing continued efforts to develop and implement an intramural epidemiology program for a number of years. At the same time, the desire for a stronger science base, and concerns during the energy crisis of the mid to late 1970s over potential increased use of coal, led to a shift to an increased role for the National Institutes for Health which, through its extramural, peer-reviewed grants program, initiated funding of the Harvard Six Cities Study. This study, which developed and followed a cohort of over 14,000 children (21) and 8,000 adults (22) in six cities in the eastern and midwestern

**TABLE 1. Overview of US particulate matter standards chronology**

<table>
<thead>
<tr>
<th>Year(s)</th>
<th>Action</th>
<th>Observational data considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>1969</td>
<td>First particulate matter (PM&lt;sub&gt;*&lt;/sub&gt;) criteria document issued</td>
<td>Air pollution episodes (e.g., London fog) cross-sectional studies of morbidity and mortality</td>
</tr>
<tr>
<td>1970</td>
<td>Clean Air Act Amendments contain provisions for criteria, US National Ambient Air Quality Standards (NAAQS&lt;sup&gt;1&lt;/sup&gt;)</td>
<td>Based on 1969 criteria document. Rationale, key studies not clearly specified in Notice Final Rulemaking</td>
</tr>
<tr>
<td>1971</td>
<td>Promulgation of original standards for total suspended particles (260µg/m&lt;sup&gt;2&lt;/sup&gt; 24-hour; 1 exceedance/year; 75 µg/m&lt;sup&gt;3&lt;/sup&gt; annual geometric mean)</td>
<td>National Academy of Sciences/National Research Council report, “Airborne Particles&lt;sup&gt;2&lt;/sup&gt;”</td>
</tr>
<tr>
<td>1974–1977</td>
<td>Energy Research and Development Administration legislation proscribes use of US Environmental Protection Agency (EPA&lt;sup&gt;3&lt;/sup&gt;) Community Health and Environmental Surveillance System epidemiology studies for standards; Harvard Six Cities Study initiated by NIEHS&lt;sup&gt;4&lt;/sup&gt; in energy crisis</td>
<td>Preponderance of epidemiologic studies done in cities where sulfur oxides/PM both elevated, difficult to separate; recognition of transformation products of sulfur oxides (acid sulfates)</td>
</tr>
<tr>
<td>1976</td>
<td>EPA issues schedule for review of all six criteria documents and NAAQS; combined sulfur oxides/PM document scheduled last</td>
<td>EPA paper on inhalable and fine particle indicators based on dosimetry, air quality data from Inhalable Particles Network. External evaluations of epidemiology sponsored by Industry and EPA. Expanded analysis of 14 London winters, London bronchitis panel; prospective and cross-sectional epidemiologic studies. Criteria document criteria for quantitative epidemiology include need for PM measures to have direct relation to mass, other criteria. Concerns about cross-sectional methodology limit use of some studies</td>
</tr>
<tr>
<td>1977</td>
<td>Clean Air Act Amendments create Clean Air Scientific Advisory Committee (CASAC&lt;sup&gt;5&lt;/sup&gt;) mandate 5-year NAAQS reviews</td>
<td>1982 criteria document and staff paper, CASAC recommendations</td>
</tr>
<tr>
<td>1979–1982</td>
<td>First review of PM criteria and standards. American Iron and Steel Institute lawsuit heightens scrutiny of the process. First PM staff paper, CASAC review and recommendations for standards. Recommendations for new indicator (PM&lt;sub&gt;10&lt;/sub&gt;) for inhalable particles based on dosimetry; levels and daily/annual averaging times based on epidemiology</td>
<td>Reanalyses of London daily mortality-Black Smoke data with more advanced techniques. Studies of acute lung function change and initial data from Six Cities Study. No epidemiologic studies used PM&lt;sub&gt;10&lt;/sub&gt;; new standard resulted in initiation of expansive PM&lt;sub&gt;10&lt;/sub&gt; network</td>
</tr>
<tr>
<td>1984</td>
<td>EPA proposes replacing total suspended particles standards with PM&lt;sub&gt;10&lt;/sub&gt;. Wide range proposed to highlight scientific uncertainties, enhance discussion</td>
<td>1982 criteria document and staff paper, CASAC recommendations</td>
</tr>
<tr>
<td>1986–1987</td>
<td>Addenda to criteria document, staff paper. Staff paper and CASAC panel recommendations similar to 1982. Promulgation of PM&lt;sub&gt;10&lt;/sub&gt; standards (150 µg/m&lt;sup&gt;3&lt;/sup&gt; 24-hour, 1 expected exceedance per year–3 year test; 50µg/m&lt;sup&gt;3&lt;/sup&gt; annual mean–3 year test). Staff and CASAC find no basis for separate sulfate/acid standard. CASAC recommends research, acid aerosol issue paper. American Iron and Steel Institute raises data availability issue for Harvard Six Cities Study in lawsuit</td>
<td>Preprints for London daily mortality studies, Black Smoke data with more advanced techniques. Studies of acute lung function change and initial data from Six Cities Study. No epidemiologic studies used PM&lt;sub&gt;10&lt;/sub&gt;; new standard resulted in initiation of expansive PM&lt;sub&gt;10&lt;/sub&gt; network</td>
</tr>
</tbody>
</table>

Table continues
United States, produced findings that assumed increasing importance in standards reviews as results where reported from the 1980s to the present.

Controversy over the findings about particulate matter, and the EPA's reactions in response to them, also led to growing calls, especially from industry, for access to the underlying data for these studies. In litigation filed by the American Iron and Steel Institute to challenge the 1987 standards, the plaintiffs called into question whether the Harvard Six Cities Study, or any other study, could be used to support regulation unless all affected parties had access to the data underlying the studies and the opportunity to further analyze them. The court upheld the EPA's use of these studies (23), but the issue continued to raise substantial concern in the 1996–1997 review and led to increased use of reanalysis, as discussed below.

### THE KEY QUESTIONS ADDRESSED IN THE CASE STUDY

#### How has the research been initiated and funded?

The extent of research on an environmental health issue, such as particulate matter, depends upon the funding available. Generally, within a funding agency, the total funds available for health research on the criteria air pollutants have not varied considerably from year to year, and the same funding pool would generally support work by different disciplines and directed at other health issues as well. The regulatory and public interest in a given issue does, however, vary from year to year. Consequently, the proportion of the research funds available for a particular issue, such as particulate matter, can vary widely, often in cycles.
reflecting the level of regulatory concern and decision-making: sustained research may be difficult as a result. Over time, funds available for particulate matter research have fluctuated greatly. Ten years ago, ozone, sulphur dioxide, and even NO, were judged to be of greater public health and regulatory concern and better funded than particulate matter. Based on CASAC recommendations (24), the EPA and the National Institute of Environmental Health Sciences focused particulate matter research resources on monitoring and exposure to acid aerosols. Funds for epidemiologic research, especially within the EPA, were particularly limited after the problems with the Community Health and Environmental Surveillance System. Although, in general, funding lapsed, there have been exceptions: beginning in the 1970s, the National Institute of Environmental Health Sciences provided some of the only long-term funding for particulate matter studies with its 20-year support of the Harvard Six Cities Study. In the 1980s, the California Air Resources Board initiated a more than 10-year commitment to the Southern California Children’s Health Study.

In the early 1990s, after several investigator-initiated studies found an association of particulate matter with adverse health effects at current ambient levels, availability of funds for particulate matter research increased and continued to do so as the issue became more visible and after the EPA took regulatory action in July 1997. For example, Congress appropriated to the EPA nearly $50 million for particulate matter research for fiscal years 1998 and 1999, funding which has continued in subsequent years. The National Institute of Environmental Health Sciences did not set aside specific funding for particulate matter, but continued to fund a substantial number of studies. Other major funding institutes, including the Health Effects Institute, the Electric Power Research Institute, and the California Air Resources Board, have also notably increased their funding for particulate matter research in recent years. Recently, a series of European studies has appeared in the literature, reflecting an increase in funding for particulate matter research by Directorate General Research of the European Commission driven by the emergence there of some of the same new analytic findings and issues found in the United States. The research in progress on particulate matter blends investigator-initiated research with more targeted, programmatic investigation, driven by agency planning.

The research needs for particulate matter research have historically been developed by the EPA, which issues statements of research needs to address uncertainties associated with an issue, such as particulate matter (25), with review by CASAC. The high profile of particulate matter research led the Congress to turn to the National Research Council to establish a committee, with a 5-year horizon, to set out a research agenda and to then monitor progress on that agenda. The committee’s first report, “Research Priorities for Airborne Particulate Matter” (4), offered a multi-year portfolio of particulate matter research needs and priorities, while its second report refined the portfolio and provided a plan for monitoring progress (26). The committee’s final report, due in November 2002, represents one of the first opportunities in the last 30 years of air pollution and health research to evaluate if a research agenda is actually met.

How have the findings been synthesized and evaluated?

In evaluating the scientific basis for standards for the criteria pollutants, the EPA assembles a comprehensive criteria document compiling all of the relevant evidence. The more recent criteria documents have tended to be encyclopedic, listing or summarizing all of the individual studies and other data relevant to the specific pollutant. The original motivation for this approach stemmed in part from past legal challenges to the documents for failing to include certain studies.

The history of particulate matter standards decisions, and of science more broadly, indicates that not all scientific studies carry equal weight. Publication or acceptance for publication in a peer-reviewed journal is a necessary, but not sufficient, condition for a study to be considered important. The EPA convenes expert workshops for the specific chapters to help determine the weight given to specific studies in succeeding drafts. Document review by EPA staff, CASAC, and the public may also influence the weight given to a specific study. Although there have been challenges in conducting this process, the past approach to development and review of criteria documents has proven to be adequate through a number of standard revisions.

Our review of particulate matter regulatory decisions, listed in table 1, shows that certain studies appear to have had particularly strong influence. These include the results of major cohort studies, like the Harvard Six Cities Studies, and of a series of dosimetry studies that helped determine the shift to a standard for the inhalable fraction of particulate matter. Studies that were solicited to address a specific issue (for example, replication of earlier studies and the weather and co-pollutant analyses organized and funded by the Health Effects Institute (14, 18), and the fine versus coarse fraction particulate matter mortality study (27)), were often found to be influential, most likely because they address key aspects of the decision regarding the indicator, averaging time, form, or level of the standards. While British studies formed a large part of the basis for the 1971 and 1987 decisions, international studies were generally not weighted as highly in the 1996–1997 review. This is likely the result of the greatly increased number of North American studies and the uncertainties inherent in extrapolation of international results, particularly those using different particulate matter indicators, to the US setting.

In preparing the criteria document and drawing out the implications for policy-making, the Agency faces the common challenge of synthesizing what is known from diverse lines of scientific investigation. Even summarizing and including the latest epidemiologic findings is a challenge (the 1996 criteria document for particulate matter, for example, included a 404-page chapter on epidemiologic studies with over 250 references), and one that is being faced in other applications of epidemiologic and clinical data, for example, in developing evidence-based guidelines for clinical practice.

The authors identified new approaches for synthesizing evidence that could benefit the decision-making process. In the 1996 criteria document (2), Chapter 13 provides an “inte-
marginal plausibility of the associations of particulate matter with morbidity and mortality has been a repeated point of discussion for promulgating a standard for fine particulate matter. Plausibility is among the criteria for causality. As with the other criteria, no specific guidelines exist for gauging the adequacy of the evidence on plausibility (i.e., there is no “plausibility court” of scientists who weigh the evidence). In fact, as we search for mechanisms of disease, we face an ever-deepening level of scientific inquiry that may make the target for plausibility even more remote. Epidemiologists may not be equipped to assess plausibility, and this determination probably best rests with appropriate multidisciplinary groups.

As shown in the example of particulate matter, plausibility may not need to be fully established before decisions can be made. The criteria for causality are not equivalent to criteria for decision-making, and the Administrator needs to promulgate standards that protect the public health “with an adequate margin of safety,” even in the face of uncertainty. Epidemiologists need to acknowledge uncertainty and provide findings in a fashion that facilitates characterization of the degree of uncertainty.

What has been the role of science supported by opponents and proponents in the discussion and interpretation of the science?

The setting of National Ambient Air Quality Standards is by its very nature a process that involves the intersection of science, public health policy, and advocacy. Implementation of the control measures that are required to attain the National Ambient Air Quality Standards typically requires the expenditure of substantial amounts of funds by industry. Historically, industry has invested substantial resources in supporting critiques of scientific studies viewed as supporting the need for tighter standards, as well as supporting replications and reanalyses of these studies. Public health and environmental organizations are, by their very mission, active advocates for precautionary approaches to protecting the health of the public. These organizations are involved in monitoring the standard-setting process and in translating for the public and policy-makers the scientific basis for tightening National Ambient Air Quality Standards when appropriate.

Congress has also involved itself in the interpretation of the science with regard to the appropriate setting or revising of National Ambient Air Quality Standards. As noted above, controversy regarding the quality of the particulate matter/sulphur dioxide data and analysis of the Community Health and Environmental Surveillance System studies led to a legislative proscription against the use by the EPA of the Community Health and Environmental Surveillance System data. More recently, Congress held extensive hearings on the scientific basis for the revision of the particulate matter standards in 1997.

The history of reviews and revisions of the particulate matter standard (table 1) provides numerous examples of the influence on the regulatory process of science supported by interest groups. Efforts to influence the EPA’s late 1970s particulate matter standard review are exemplified by environmental groups such as the Natural Resources Defense Council in its 1977 report on the need for a fine particle air...
quality standard (28), as well as the 1979 review papers by Holland et al. (29) supported by the American Iron and Steel Institute, which suggested that change in the particulate matter standard was not warranted. EPA’s need to respond to the Holland et al. papers resulted in the Agency’s support of another review of the same evidence by Ware et al. (30) in 1981. Industry-supported efforts to provide either new analyses or reanalyses of previous studies occurred in the early-1980s (31) and more extensively in the mid-1990s (32, 33). The mid-1990s analyses commissioned by industry critics of tighter particulate matter standards typically had findings of weaker statistical strength or negative associations between particulate matter and adverse health effects. In contrast, the Natural Resources Defense Council’s 1996 report (34) calculating national particulate-related mortality provided health and environmental organizations with evidence packaged to advocate for tighter particulate matter standards. The target audiences for these studies and reports included the EPA and CASAC, Congress, and the media.

Interest group involvement in the particulate matter review process seems to differ between the pre- and post-proposal periods. In the pre-proposal period, the interest groups attempt to influence the scientific basis for the EPA’s decision on how to proceed with review of the standard. These efforts include providing comments to the EPA and CASAC on draft criteria documents and staff papers, as well as participating in EPA meetings for reviewing evidence. It is during this process that new peer-reviewed scientific data on particulate matter health effects are presented, important recent scientific results emphasized, and differences in interpretations of the scientific evidence are discussed.

During the post-promulgation period, the emphasis shifts to providing Congress, local elected officials, the media, and the public with “spin” on the science. The emphasis on discussion of peer-reviewed science is often replaced by discussion of commissioned analyses with results distilled to the “sound bite.”

Ultimately, science supported by interest groups can be used in one or more of the following ways: 1) to critique existing studies and raise questions regarding their reliability for use in the standard-setting process; 2) to provide new results that fill data gaps in the existing science; 3) to replicate and/or reanalyze existing studies that support a regulatory action to provide alternative interpretations of the study results; and 4) to extend study results to provide advocates at the national and local levels with data to present to the public, the media, and elected officials.

What has been the role of replication and reanalysis of epidemiologic studies?

Emphasis in the recent particulate matter debate on a set of analyses that applied newer techniques for time-series analysis, and on the findings of two cohort studies on mortality, sparked intense debate over the need for replication of the studies, and for all or some parties to the debate to have access to the underlying data for reanalysis. This pressure emerged first in 1994 and was then amplified in early 1997 as Congress debated the proposed National Ambient Air Quality Standards. The result was the development and implementation of approaches to formal reanalysis by third parties.

**Reanalysis of epidemiologic data.** Epidemiologic research is becoming increasingly sophisticated, with advanced statistical methods and models now commonly applied in the analysis and evaluation of complex data. While these advances have led to greater clarity in the interpretation of epidemiologic data, findings of analyses may be sensitive to the choice of methods. Reanalysis of original data by independent investigators using alternative analytic techniques can increase confidence that findings are robust to the choice of methods.

Reanalysis differs from replication in that new primary data are not generated. It can, however, be extended to include supplementary information to key covariates and effect modifiers that may not have been available to the original investigators. Reanalysis of data has played a big role in the recent consideration of the evidence on particulate matter and health.

**The role of reanalysis in the particulate matter debate.** Prior to the 1990s, time-series approaches had been used for assessing air pollution, including particulate matter, and health. Beginning in the early 1990s, Schwartz and others published a series of analyses of daily mortality counts and air pollution levels using techniques of time-series analysis that had not been applied previously to epidemiologic data. These analyses were facilitated by more powerful hardware and availability of software. The findings of the analyses, however, were quickly controversial, in part because of the “positive” findings and in part because of the suspicion of these new methods. A concern also arose that model specification and analysis might have led to selection of those models yielding significant associations, as alternatives were compared. The analyst selects, for example, among alternative lag structures, alternative approaches for removing underlying “nuisance” trends, and alternatives for controlling potential confounding.

A mounting number of reports in the early 1990s showed associations of particulate matter and other air pollutants with daily mortality. An independent reanalysis, however,
was called for and organized by the Health Effects Institute. A new reanalysis team validated several data sets and used alternative methods of analysis to confirm the findings of the original investigators (14). This formal reanalysis, which was extensively peer-reviewed, set aside general concerns about data integrity and the validity of analytic findings. The results were presented to CASAC and incorporated into the 1996 criteria document.

Data from prospective cohort studies have also been the focus of a reanalysis effort. Two large-scale studies, the Harvard Six Cities Study and the American Cancer Society's Cancer Prevention Study II, have figured centrally in the EPA's decision-making. Both showed higher risk for mortality in persons living in more polluted cities, a finding interpreted as evidence of significant life-shortening by particulate air pollution. Because of the weight given to these studies by the EPA, reanalysis was called for by many, including the Congress, and the Health Effects Institute commissioned an independent reanalysis of the data from these two key studies, led by a team at the University of Ottawa. The reanalysis was designed to first confirm the results obtained by the original investigators, and then explore the sensitivity of the original findings to alternative analytic techniques, including the method of covariate adjustment and specification of the form of the exposure-response model. Since the original workshop on which this paper is based took place, findings have been reported (35).

Other reanalyses have been undertaken as well in recent years. Under Electric Power Research Institute sponsorship, Klemm et al. (36) undertook a reanalysis of a key study by Schwartz et al (27).

Issues relating to the provision of data for reanalysis. The recent trend towards reanalysis of key epidemiologic data sets has brought forward a series of difficult issues raised by giving access to data for reanalysis. Original investigators have certain proprietary rights to their data and should clearly be afforded the first opportunities to analyze the data and publish their research findings. At some point, however, data that are relevant to policy considerations should be made available for reanalysis. Although sharing of complex large data sets seems particularly appropriate, the scientific community is not now necessarily disposed to support such reanalysis initiatives. Scientists conduct their work in a competitive environment, with peer review generally involved in both the funding of epidemiologic research and in the publication of research results. Scientists are naturally reluctant to devote additional efforts to making their data available for reanalysis, preferring instead to devote their valuable time to original research.

Recent US legislation, however, enacted as part of the Omnibus Budget Reconciliation Act for Federal Fiscal Year 1999, has required that research funded by federal grants be subject to the Freedom of Information Act when that research has been cited in federal policies or regulations (37). This legislation, commonly known as the Shelby Amendment, was proposed and adopted at the request of industries who had been unable to obtain direct access to the data underlying particulate matter studies, including the Six Cities Study. As a federal law, it applies only to research projects receiving any amount of federal funding and conducted at “institutions of higher education, hospitals and other non-profit organizations.” The lack of any comparable requirement for privately-funded research may cause an imbalance in future years in the scrutiny to which research funded by different sources is put. The Shelby Amendment may have far-reaching implications in all branches of science, but its exact impact is not yet clear since the US Office of Management and Budget has only recently completed the implementation guidelines, and these guidelines, which tend to include safeguards for investigators and confidentiality, are already being tested through litigation. Nevertheless, the past approach to development and review of criteria documents has proven to be adequate through a number of standard revisions.

Reanalysis initiatives may raise concerns about the potential for charges of fraud should errors in the original analysis be uncovered. At the same time, the scientific community is likely to be generally supportive of reanalysis since this would serve to strengthen the credibility of the original conclusions. The challenge for future reanalysis efforts will be to find mechanisms that support the sharing of data for reanalysis purposes, without compromising the integrity of the original investigators or their data.

Conclusions. Replication of empirical research results is a well-established scientific principle designed to confirm initial findings and to evaluate the generalizability of such findings to broader conditions. Replication of epidemiologic research under different conditions serves to strengthen the scientific basis for risk management policy decisions made on the basis of epidemiologic findings. During the particulate matter debate, reanalysis of complex data sets has been conducted to bolster confidence in findings that are critical to decision-making. The approach followed by the Health Effects Institute has proved effective, although it requires independent funding and peer review. Difficult issues around broader data sharing, which is likely to increase as a result of new law, will need to be resolved.

What has been the role of the epidemiologists?

Epidemiologists have figured prominently in the recent promulgation of the new particulate matter National Ambient Air Quality Standards for fine particulate matter. New epidemiologic evidence prompted the review and was the basis of the lawsuit by the American Lung Association against the EPA that drove the schedule for review of the existing standard and promulgation of the new standard. Three collaborating investigators (Schwartz, Dockery, and Pope) conducted the majority of the first wave of time-series mortality studies that prompted the review. In addition, some epidemiologists involved in the debate have not shied away from publicly addressing the policy implications of the findings. Epidemiologists have also contributed to the development of the criteria document and have been prominent members of CASAC. Others have served as consultants to interested parties, conducting analyses, commenting on methodological issues, and offering testimony at Congressional and public hearings.

The authors identified five distinct roles played by epidemiologists in the particulate matter issue:
1. Conducting research (and seeking funds to do so);  
2. Serving as analysts for the criteria document process;  
3. Reviewing criteria documents;  
4. Serving as advisory committee members (CASAC);  
5. Acting as consultants or advocates; and  
6. Serving as members of panels setting research priorities.

The group of epidemiologists who have expertise relevant to the particulate matter problem is small and some have fulfilled several of the above roles in the recent standard-setting process. Individual epidemiologists conducted relevant research, reviewed the documents, and served on CASAC. Others were funded to conduct research by parties with an interest in the outcome of the process and also served as consultants to these parties. This situation created at least the perception of conflict of interest, particularly for those considered to have advocated for or against particular risk management strategies.

Of the various scientific disciplines involved in generating and evaluating the evidence on particulate matter, epidemiologists have the strongest links to preventive public health policies and perhaps greater willingness than other scientists to support population health policies based on epidemiologic data alone. Consequently, the views of epidemiologists figured prominently in the discussions of the data on particulate matter and some epidemiologists on CASAC voiced their views separately about the strength of the evidence. Patterns of views by scientific discipline were also evident in the opinions of the members of CASAC with regard to the need for short-term and long-term PM$_{2.5}$ research, as summarized in the closure letter to the Administrator on the staff paper (38). Toxicologists were less likely to see a need for new standards, perhaps reflecting their view of the need to better define underlying mechanisms of effect of particles on health and concern about the uncertainties in observational data on particles.

The particulate matter case also has implications for the role of epidemiologists as advocates. Several epidemiologists called for a standard for fine particulate matter, even as the EPA’s review was in progress. There was also a vigorous public debate between public health scientists supporting and not supporting a standard. The advocacy role reflected directly on the interpretation of the research conducted by those involved in putting forward positions. Although their research was ultimately considered in the standard-setting process, views of these epidemiologists as advocates for a particular point of view have continued.

CONCLUSIONS AND RECOMMENDATIONS

Based on its review of the evolution and use of epidemiology in the public policy debate concerning the health effects of particulate matter, the authors formulated the following conclusions and recommendations:

The initiation and funding of science

Because of the episodic nature of public policy debates about air pollution, the availability of funding for, and the level of interest in, research in particulate matter health effects has been cyclical and often driven by a need to use available data and techniques in an immediate fashion rather than having the sustained support needed for additional epidemiologic research. The Harvard Six Cities Study, which originated from concerns during the early 1970s as to the health effects of fossil fuels, is an exception demonstrating the benefits of sustained investment in research.

Ensuring the timely production of high-quality epidemiologic research to answer key questions requires the development of an overarching research strategy and the establishment of a mechanism to monitor and encourage the implementation of that strategy. The recent efforts of the National Research Council’s Committee on Research Priorities for Airborne Particulate Matter (4, 26, 39) may be a model for providing such a mechanism.

At the same time, the authors noted that part of the incentive for the current substantial investments in additional research has been the required review in 5 years of the air quality standards. However, these reviews have often been spread over much longer periods (10 to 15 years and more) with interest in, and funding for, research into a particular pollutant waning in the intervening years. Regular policy decision-making (e.g., at consistent 5-year intervals) would facilitate the development of continued interest in, and the implementation of, a longer-term research strategy.

Synthesis and evaluation of the science

An extraordinarily large number of epidemiologic studies were available for consideration in the most recent review of the science on the health effects of particulate matter. Although there is extensive effort invested in collecting and sorting through these studies, currently no systematic techniques exist for identifying the most important studies, and the process has made relatively little use of emerging formal techniques for combining evidence.

The authors identified the need to develop and apply tools for systematic review and evaluation of studies so that the process is transparent and lends itself to more thoughtful analysis of only those studies most significant to standard-setting. They also recommended the increased and appropriate use of emerging, improving tools for meta- and pooled analyses, so that the combined weight of the evidence could be assessed more effectively.

The issue of biologic plausibility

The large body of epidemiologic evidence available on particulate matter has not been matched by an analogous body of toxicologic data. In addition, many observers questioned the biologic plausibility of the increases in mortality found associated with quite small decreases in air pollution in the epidemiologic studies. Although plausibility is only one of Hill’s criteria (40) for assessing causality, the debate over the absence of plausibility evidence raised the question of whether there was sufficient evidence of causality to support regulatory action. In this regard, the authors found the criteria for causality a useful guide for evaluation, but not
sufficiently detailed to be used as an absolute standard. Even though plausibility was clearly not established in the case of particulate matter, public policy decision-makers found the consistency and coherence of the epidemiologic evidence compelling enough to act, even in the face of uncertainty. Looking to the future, epidemiologists’ efforts could benefit from efforts to: 1) assess consistency and coherence of the data carefully in such circumstances; 2) acknowledge and find ways to characterize the uncertainty in the findings; and 3) develop new collaborations with the toxicologic community to investigate possible biologic mechanisms in order to assess plausibility.

The role of science supported by advocates

As would be expected in a public policy debate with large potential consequences for public health and for the economy, advocates for and against more stringent national air quality standards for particulate matter have been involved in funding replication and reanalysis efforts, reviews of the literature, analyses of the public health consequences, and some original research. During the first phases of the review of the science and prior to proposal of a change in the air quality standards, this science is reviewed in front of a scientific panel (CASAC). The minimum requirement for research considered is publication in the peer-reviewed literature. The panel serves as a form of “referee,” reviewing the general contributions of the scientific community, and the specific contributions of advocate-supported investigators, and helping the EPA distinguish between the significant number of useful contributions from the advocates and the smaller number of more purely “attack” analyses.

Subsequent discussions take place in the more highly-charged atmosphere in Washington, DC, where there are no prior requirements for peer review and no scientific “referees.” Here, the advocates engage more in public relations than in science, with the opponents emphasizing the large amount of uncertainty as a reason to withhold action, and the proponents de-emphasizing uncertainty and promoting their estimates of the public health consequences if no action is taken.

This latter arena is a challenging one for epidemiologists (and other scientists as well) since the normal, careful phrasing in technical terms of what science knows and does not know is unwelcome. The experience of the particulate matter debate suggests that epidemiologists have to become increasingly adept in communicating, clearly and concisely, scientific conclusions and uncertainties. At the same time, most scientists must realize that at this primarily policy stage of the debate, science is only one input into what is inherently a more complex political decision-making process.

The role of replication and reanalysis

The debate over particulate matter air quality standards evoked strong calls from opponents for public access to, and reanalyses of, certain epidemiologic studies. The concept of replication is common in review and acceptance of toxicologic evidence. Although the ability to find consistent epidemiologic results in analyses of different cities was viewed as a form of “replication,” the process of replication per se is more difficult in epidemiology, particularly with unique cohort databases that have been developed over many years, leading to calls for more widespread sharing of underlying data. However, examples exist of past abuses when investigators have shared their data and been unfairly charged with fraud, and there can be resistance on the part of some investigators to sharing with analysts hired by advocates.

At the same time, the withholding of data can breed suspicion of the scientist, and there is at least the public perception that if data are produced with public funds, they should be accessible following a reasonable length of time for the investigator to publish from them. This was particularly true in the particulate matter case where a relatively large proportion of the studies under primary consideration were conducted by a relatively small group of investigators.

Given this background, the concept of independent third-party reanalysis emerged in the particulate matter debate, with the Health Effects Institute organizing first, in 1994–1995, a reanalysis of several time-series studies of particulate matter and mortality (14), and then, in 1998–1999, a reanalysis of the Harvard Six Cities Study and the American Cancer Society study of longer term exposure to particulate matter and mortality (35). Other organizations have also organized reanalyses.

The authors noted that, in general, with appropriate protections for confidentiality and for the investigator, access to data should be encouraged. They also noted that reanalysis can be useful in key circumstances, but that it is neither necessary nor desirable that every study under consideration be subjected to such reanalysis, in which case normal peer review and the CASAC process should suffice.

The role of the epidemiologist

In the case of the particulate matter debate, as in many other epidemiology/policy settings, a relatively small number of active and qualified epidemiologists are involved. As a result, the same epidemiologist may be called upon to play multiple roles—researcher (and seeker of research funds), reviewer/analyst on behalf of interested parties, advocate for public health protection, and member of an expert panel (e.g., CASAC).

Because of the central role of epidemiology in the particulate matter debate, epidemiologists found themselves, comfortably or otherwise, at the center. Their multiple roles could, and in some instances did, lead to at least the perception of conflict of interest among the roles. It is critical that epidemiologists recognize both real and perceived conflicts inherent in such settings, and act to minimize them so as to maintain the credibility of epidemiology and of the larger scientific process as contributors to the public policy debate.

Conclusion

For over 30 years, epidemiology has played a major contributing role in shaping the scientific and public policy debate about the health effects of particulate matter. During that time, the quality of the epidemiology produced has improved; this
case study suggests a number of ways in which epidemiology and epidemiologists can further advance both the state of the art and their participation in the scientific and policy debate.

DISCLOSURES

Mr. Dan Greenbaum is President and Chief Executive Officer of the Health Effects Institute, a non-profit research institute supported jointly by US Environmental Protection Agency (EPA) and industry to provide public and private decision-makers with high quality, impartial, relevant, and credible science about the health effects of air pollution. Mr. Greenbaum chaired the EPA Blue Ribbon Panel on Oxygenates in Gasoline which issued its report in July 1999. Mr. Greenbaum also serves on the National Research Council Board of Environmental Studies and Toxicology, and its Committees for Research Priorities on Airborne Particulate Matter and Air Quality Management. Just prior to coming to the Health Effects Institute he served as Commissioner of the Massachusetts Department of Environmental Protection from 1988 to 1994, where he was responsible for the Commonwealth’s response to the Clean Air Act, as well as its efforts on water pollution and solid and hazardous waste.

Dr. John Bachmann is Associate Director for Science/Policy and New Program Initiatives at the Environmental Protection Agency’s (EPA) Office of Air Quality Planning and Standards (OAQPS) in Research Triangle Park, NC. He has worked in a number of air-related programs since 1974, including ambient air quality standards; air carcinogen policy; strategy assessment for regional sulfur and nitrogen oxides emissions and related acid deposition, visibility, health, and other effects. Dr. Bachmann developed the first Agency “Position Paper on Regulation of Atmospheric Sulfates” in 1975, and led the production of the particulate matter EPA staff papers and addenda for the initial (1978–1987) as well as subsequent reviews of the National Ambient Air Quality Standards for particulate matter. These multi-year reviews resulted in the 1987 PM$_{10}$ standards and the 1997 standards for fine (PM$_{2.5}$) and coarse (PM$_{10}$) particles. Dr. Bachmann works with the EPA, other Federal Agencies, as well as academic and stakeholder funded programs in coordinating scientific research planning for regulatory and policy needs for air programs. For several years, he participated on the EPA team working to develop and assess the Clean Air Act Amendments of 1990. In the 1990s he worked on developing market-based programs for ozone and other pollutants, coordinating international air programs for North America, and developing integrated multi-pollutant programs for power generation.

Dr. Daniel Krewski is Professor and Director of the R. Samuel McLaughlin Centre for Population Health Risk Assessment at the University of Ottawa, where he is involved in a number of activities in population health risk assessment within the new Institute of Population Health. Dr. Krewski has also served as Adjunct Research Professor of Statistics in the Department of Mathematics and Statistics at Carleton University since 1984. Prior to joining the Faculty of Medicine at the University of Ottawa in 1998, Dr. Krewski was Director of Risk Management in the Health Protection Branch of Health Canada. While with Health Canada, he also served as Acting Director of the Bureau of Chemical Hazards and as Chief of the Biostatistics Division in the Environmental Health Directorate. His professional interests include epidemiology, biostatistics, risk assessment, and risk management. Dr. Krewski is a Fellow of the American Statistical Association and the Society for Risk Analysis. He currently Chairs the US National Academy of Science’s Committee on Acute Exposure Guidance Levels for Highly Hazardous Substances.

Dr. Jonathan Samet is Professor and Chairman of the Department of Epidemiology of the Bloomberg School of Public Health, Johns Hopkins University. He is also Director of the Institute for Global Tobacco Control and Co-Director of the Risk Sciences and Public Policy Institute at the Bloomberg School of Public Health, Johns Hopkins University. His research has addressed the effects of inhaled pollutants in the general environment and in the workplace. He has written widely on the health effects of active and passive smoking and served as Consultant Editor and Senior Editor for Reports of the Surgeon General on Smoking and Health. He has served on the Science Advisory Board for the US Environmental Protection Agency and was Chairman of the Biological Effects of Ionizing Radiation Committee (BEIR VI) of the National Research Council. He is presently Chairman of the National Research Council’s Committee on Research Priorities for Airborne Particulate Matter.

Mr. Ronald White is the Assistant Executive Director for Education, Research and Community Affairs at the National Osteoporosis Foundation. At the time of the Symposium, he was the Assistant Vice President for National Policy at the National Office of the American Lung Association, where he was responsible for directing the Lung Association’s policy initiatives on national air quality standards. Mr. White has been involved with national air quality public policy issues for more than two decades. He continues to serve as a non-paid volunteer for the American Lung Association on air quality issues. Mr. White also serves as a consultant member of the US Environmental Protection Agency Clean Air Scientific Advisory Committee on the Particulate Matter Air Quality Standards Review, and is a member of the National Research Council Committee on Research Priorities for Airborne Particulate Matter. Mr. White also serves on the External Scientific Advisory Committee of the National Environmental Respiratory Center.

Dr. Ronald Wyzga is Technical Executive/Area Manager, Air Quality Health and Risk at the Electric Power Research Institute in Palo Alto, CA. Prior to joining the Electric Power Research Institute, he worked at the Organization for Economic Cooperation and Development in Paris, where he co-authored a book on economic evaluation of environmental damage. He serves on, and has chaired, several committees for the US Environmental Protection Agency’s Science Advisory Board Committees and the National Academy of Sciences, including the ongoing National Academy of Sciences committee that oversees the US Environmental Protection Agency’s particulate matter research program through 2002. He is a fellow of the American Statistical
Association. His research interests are environmental risk assessment and health effects of air pollution.

REFERENCES