United States
Environmental
Protection Agency

EPA Science Advisory
Board (1400F)
Washington, DC

EPA-SAB-CASAC-05-007
June 2005
www.epa.gov/sab

# EPA's Review of the National Ambient Air Quality Standards for Particulate Matter (Second Draft PM Staff Paper, January 2005)

*A Review by the Particulate Matter Review Panel of the EPA Clean Air Scientific Advisory Committee*

**UNITED STATES ENVIRONMENTAL PROTECTION AGENCY**
**WASHINGTON D.C. 20460**

June 6, 2005

EPA-SAB-CASAC-05-007

**OFFICE OF THE ADMINISTRATOR**
**SCIENCE ADVISORY BOARD**

Honorable Stephen L. Johnson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Subject: Clean Air Scientific Advisory Committee (CASAC) Particulate Matter (PM) Review Panel's Peer Review of the Agency's *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information* (Second Draft PM Staff Paper, January 2005); and *Particulate Matter Health Risk Assessment for Selected Urban Areas: Second Draft Report* (Second Draft PM Risk Assessment, January 2005)

Dear Administrator Johnson:

EPA's Clean Air Scientific Advisory Committee (CASAC), supplemented by subject-matter-expert Panelists — collectively referred to as the CASAC Particulate Matter (PM) Review Panel ("Panel") — met in a public meeting held in Durham, NC, on April 5-6, 2005, to conduct a peer review of subject documents. The current Panel roster is found in Appendix A of this report.

This meeting was a continuation of the CASAC PM Review Panel's peer review of the *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information (*First Draft PM Staff Paper, August 2003*)* and a related draft technical report, Particulate Matter Health Risk Assessment for Selected Urban Areas (First Draft PM Risk Assessment, August 2003*)*. The previous draft of the PM Staff Paper was a preliminary version since the Panel has not yet finished its review of the Air Quality Criteria Document (AQCD) for PM (which was completed in October 2004). In addition, further risk analyses and analyses of alternative forms of the PM standards were included in the Second Draft PM Staff Paper and Second Draft PM Risk Assessment. The charge questions provided to the Panel by EPA are found in Appendix B to this report. Panelists' individual review comments are provided in Appendix C of this report.

In its peer review of the Second Draft of the PM Staff Paper, most of the members of the CASAC PM Review Panel found the document was generally well-written and scientifically well-reasoned for all but the short term primary $PM_{10-2.5}$ standard. A majority of the members of the Panel were in agreement with the following: the primary $PM_{2.5}$ 24-hour and annual PM national ambient air quality standards (NAAQS) should be modified to provide increased public health protection. Although the evidence for a standard for coarse-mode particles was weaker than for the $PM_{2.5}$, the Panel agreed that a 24-hour NAAQS for $PM_{10-2.5}$ was appropriate, especially in urban areas, with caveats to make exceptions for those types of rural dusts thought to have low toxicity. The Panel recommends that the Agency staff expand and strengthen the discussion of the exposure index (size-range plus composition and/or source) and the monitoring strategy to be used for the coarse-mode NAAQS, as well as the degree of public health protection against thoracic coarse PM expected relative to the protection afforded by the current $PM_{10}$ short-term NAAQS. As discussed below, the CASAC PM Review Panel will need to review the final version of the PM Staff Paper before providing a final opinion to EPA on the adequacy of a short-term $PM_{10-2.5}$ NAAQS.

The approach used to set secondary NAAQS to protect the environment was considered appropriate, but it was strongly recommended that, in the future, Agency staff also give serious consideration to a shift to the European approach of critical loads to protect vegetation and ecosystems in the U.S. In addition, most of the Panel supported Agency staff recommendations regarding a standard to address the issue of urban visibility impairment.

## 1. Background

The CASAC, comprised of seven members appointed by the EPA Administrator, was established under section 109(d)(2) of the Clean Air Act (CAA or “Act”) (42 U.S.C. § 7409) as an independent scientific advisory committee, in part to provide advice, information and recommendations on the scientific and technical aspects of issues related to air quality criteria and NAAQS under sections 108 and 109 of the Act. Section 109(d)(1) of the CAA requires that EPA carry out a periodic review and revision, where appropriate, of the air quality criteria and the NAAQS for “criteria” air pollutants such as PM. The CASAC, which is administratively located under EPA’s Science Advisory Board (SAB) Staff Office, is a Federal advisory committee chartered under the Federal Advisory Committee Act (FACA), as amended, 5 U.S.C., App. The CASAC PM Review Panel is comprised of the seven members of the chartered (statutory) Clean Air Scientific Advisory Committee, supplemented by fifteen technical experts.

Under section 108 of the CAA, the Agency is required to establish NAAQS for each pollutant for which EPA has issued criteria, including PM. Section 109(d) of the Act subsequently requires periodic review and, if appropriate, revision of existing air quality criteria to reflect advances in scientific knowledge on the effects of the pollutant on public health and welfare. EPA is also to revise the NAAQS, if appropriate, based on the revised criteria. The purpose of the Second Draft PM Staff Paper is to evaluate the policy implications of the key scientific and technical information contained in a related document, EPA’s revised PM AQCD (October 2004), and to identify critical elements that EPA believes should be considered in the review of the PM NAAQS. The Staff Paper for PM is intended to “bridge the gap” between the scientific review contained in the PM AQCD and the public health and welfare policy judgments required of the Administrator in reviewing the PM NAAQS.

This Second Draft PM Staff Paper is based on the information in the final PM AQCD, which had been the subject of review by the CASAC PM Review Panel since October 1999. (The report from the Panel's final meeting to review the PM AQCD, dated October 4, 2004, is posted on the SAB Web Site at: http://www.epa.gov/sab/pdf/casac05001.pdf. The Agency subsequently announced the availability of a final document, *Air Quality Criteria for Particulate Matter* (EPA/600/P-99/002aF, EPA/600/P-99/002bF) on October 29, 2004.) In addition, the Second Draft PM Staff Paper builds upon the First Draft PM Staff Paper, which was the subject of review by the CASAC PM Review Panel held on November 12-13, 2003. The report from the Panel's previous meeting to review these draft documents, dated February 18, 2004, is posted on the SAB Web Site at: http://www.epa.gov/sab/pdf/casac_04004.pdf. The Second Draft PM Staff Paper and the Second Draft PM Risk Assessment were made available for public review and comment on January 31, 2005 by EPA's Office of Air Quality Planning and Standards (OAQPS), within the Office of Air and Radiation (OAR). The Second Draft PM Risk Assessment, which builds upon the Agency's First Draft PM Risk Assessment, describes the methodology and presents the results from an updated PM health risk assessment for health risks associated with exposure to fine and thoracic coarse particles in a number of U.S. cities.

**2. CASAC PM Review Panel's Peer Review of the Second Draft PM Staff Paper and Second Draft PM Risk Assessment**

After reviewing the Second Draft PM Staff Paper and written comments from the public, and after hearing public comments at the meeting, a majority of the members of the CASAC PM Review Panel were in agreement with the following: the primary $PM_{2.5}$ 24-hour and annual NAAQS should be modified to provide increased public health protection. The evidence for a NAAQS for coarse mode particles is weaker than for $PM_{2.5}$. The Panel agreed, however, that a 24-hour NAAQS for $PM_{10-2.5}$ is appropriate, especially in urban areas and with caveats to make exceptions for those types of rural dusts thought to have low toxicity. Before the Panel renders its final recommendation concerning a daily $PM_{10-2.5}$ standard, the Panel recommends that the Agency staff expand and strengthen the discussion of the exposure index (size-range plus composition and/or source) and the monitoring strategy to be used for this standard, as well as the degree of public health protection expected relative to the protection against thoracic coarse PM afforded by the current $PM_{10}$ short-term NAAQS. Accordingly, after the Panel has reviewed the Final Staff Paper and Risk Assessment for Particulate Matter following its issuance on June 30, 2005, the Panel will meet again this summer via a public teleconference to consider the final Staff Paper's recommendations concerning the setting of a coarse PM standard. Subsequent to the Panel's teleconference meeting, we will send you a separate letter providing the Panel's recommendations concerning $PM_{10-2.5}$ as an indicator together with our views on the averaging time, statistical form, and level of any potential daily $PM_{10-2.5}$ standard.

The approach used to set secondary standards to protect the environment was considered appropriate, but it was strongly recommended that, in the future, Agency staff give serious consideration to the European approach of critical loads to protect vegetation and ecosystems in the U.S. In addition, most of the Panel supported Agency staff recommendations regarding a standard to address the issue of urban visibility impairment.

In its peer review of the Second Draft of the PM Staff Paper, most of the members of the CASAC PM Review Panel found the document was generally well-written and scientifically well-reasoned. The following represent summaries of advice and recommendations of the Panel in response to the charge questions provided by EPA, which are found in Appendix B to this report. More detailed responses are provided in the individual review comments of each member of the Panel included in Appendix C to this report.

The CASAC PM Review Panel has reached agreement on the following synopsis of advice and recommendations for the Agency:

## AIR QUALITY

Chapter 2: **Characterization of Ambient PM**

Chapter 2 of the Second Draft PM Staff Paper was considered well-written, presenting an accurate and concise summary of Chapters 2, 3, and 5 of the PM Air Quality Criteria Document. The chapter was acceptable to the Panel reviewers as written, but some improvements were suggested in two areas. In the area of measurement methods, the Panel thought there should be more discussion of continuous PM monitoring methods in light of the recommended secondary fine particle standard based on 4- to 8-hour concentration averages and the likely availability of a continuous coarse particle monitor. A more quantitative characterization of PM mass measurement errors could be presented, especially for $PM_{10-2.5}$. Interest was expressed in a discussion of alternative PM indicators for future NAAQS considerations related to the source of the PM, especially for the potentially more toxic portion of coarse PM. In the area of health and visibility assessments, concern was expressed that spatial gradients near major arterials and other urban sources are not adequately addressed. It was suggested that spatial heterogeneity within a city might better be characterized in terms of departures of individual sites from the metropolitan average, in place of this draft's summary statistics of between-sites comparisons. Some members of the Panel expressed concerns about the policy-relevant background (PRB) estimates. The true background is not observable and is effectively unknowable. As indicated in the summary of Chapter 5 comments, alternative standards should be analyzed in ways that are insensitive to estimates of PRB.

## HEALTH-BASED STANDARDS

Chapter 3: **Policy-Relevant Assessment of Health Effects Evidence**

Chapter 3 addresses each of the health effects issues relevant to PM NAAQS reconsideration. Agency staff have adequately reviewed advances in understanding effects from studies conducted subsequent to the 1997 NAAQS, as summarized in the latest PM AQCD. Overall, EPA staff have done a reasonably good job of summarizing the health effects basis for considering revised or new PM standards. However, there are instances where the summary of findings and their interpretation are overstated (see individual Panel member review comments, particularly on pages C-82 and C-83). Specifically, there was confusion over strength of association versus strength of evidence, between confounding and effect modification, and between temporality and lag structure. There are some areas where Agency staff have either over-interpreted or overstated the extent to which the health data support a particular PM indicator variable. These problems can be addressed if EPA staff give heed to the individual comments of the CASAC PM Review Panel when revising the chapter. The discussion of the effect of co-pollutants in interpreting the results of PM health studies would benefit from a clearer discussion of EPA staff's approach to interpreting quantitative results from multi-pollutant studies.

Numerous epidemiological studies that are reviewed in this chapter have shown statistically significant associations between the concentrations of ambient air $PM_{2.5}$ and $PM_{10}$ (including levels that are lower than the current PM NAAQS) and excess mortality and morbidity. Furthermore, the evidence presented indicates that the effects appear to be robust, in that inclusion of other environmental variables in regression analyses has not been found to materially affect the associations of the adverse health effects with ambient PM concentrations. On the other hand, the scientific evidence included in the PM Air Quality Criteria Document and draft Staff Paper provides substantially less data derived from controlled exposure studies in humans or experimental animals; or from studies of biological mechanisms in animals *in vivo* or cells and tissues *in vitro*, to support the biological plausibility of the effects of the relatively low concentrations found in the human population studies. In the case of controlled human studies, this appears to be due to the inherent limitations of such studies, which are largely confined to young, healthy subjects exposed for short time periods and the examination of mild, reversible effects. In the case of animal studies, it may be due to not having adequate animal models of human disease processes and exposures to individual chemical agents rather than realistic mixtures. Both types of studies may be inadequate to represent the real-world situation of susceptible subpopulations of humans undergoing long-term exposures and occasional peak levels of complex mixtures of PM, associated surface coatings of reactive chemicals, and gaseous co-pollutants. This apparent discontinuity needs to be addressed in future research.

The current health effects data base for coarse mode particles ($PM_{10-2.5}$) is relatively weak. Few epidemiology studies have been conducted where $PM_{10-2.5}$ was measured directly as opposed to obtaining an estimate of this indicator variable by subtracting data from collocated $PM_{2.5}$ and $PM_{10}$ monitors. There is limited evidence that $PM_{10-2.5}$ may be related to cardiovascular mortality as well as to hospital admissions for respiratory diseases. The few controlled human studies that have been conducted with concentrated ambient particles have not shed any light on the morbidity findings from epidemiological studies. Moreover, animal toxicological studies using coarse mode particles are virtually nonexistent; they are difficult to perform because rodents are obligate nose breathers and thus few of these particles penetrate to the lungs. A further complication with current epidemiological studies of the health effects of $PM_{10-2.5}$ is that most have been conducted in urban areas, and because coarse mode particles from urban and rural areas may be markedly different, extrapolating these findings to rural settings may be difficult. Considerably more research with $PM_{10-2.5}$ is needed.

Chapter 4: **Characterization of Health Risks**

One major concern with the current version of the chapter is the clarity of presentation. Readers need to struggle through dense prose and jargon-ridden text to identify key aspects of the methods and findings. Key terms are sometimes used incorrectly or inconsistently across the chapter. The chapter could be substantially shortened, and redundancies need to be addressed. Figure 4-1 provides an overall framework for the risk assessment that could be used to shape the chapter. We suggest that the chapter refer to it repetitively as the risk assessment methods and findings are described. Subheadings, such as "assumptions" and "sensitivity analyses," might be more effectively used to guide the reader through the individual sections of the chapter.

A second concern is with methodological issues. The issue of the selection of concentration-response (C-R) relationships based on locally-derived coefficients needs more discussion. The Panel did not agree with EPA staff in calculating the burden of associated incidence in their risk assessment using either the predicted background or the lowest measured level (LML) in the utilized epidemiological analysis. The available epidemiological database on daily mortality and morbidity does not establish either the presence or absence of threshold concentrations for adverse health effects. Thus, in order to avoid emphasizing an approach that assumes effects that extend to either predicted background concentrations or LML, and to standardize the approach across cities, for the purpose of estimating public health impacts, the Panel favored the primary use of an assumed threshold of 10 $\mu g/m^3$. The original approach of using background or LML, as well as the other postulated thresholds, could still be used in a sensitivity analysis of threshold assumptions.

The analyses in this chapter highlight the impact of assumptions regarding thresholds, or lack of threshold, on the estimates of risk. The uncertainty associated with threshold or nonlinear models needs more thorough discussion. A major research need is for more work to determine the existence and level of any thresholds that may exist or the shape of nonlinear concentration-response curves at low levels of exposure that may exist, and to reduce uncertainty in estimated risks at the lowest PM concentrations.

Chapter 5: **Staff Conclusions and Recommendations on Primary PM NAAQS**

The Panel had the following advice and recommendations for the PM 2.5 standard:

The tack taken by EPA staff in recommending a suite of standards for $PM_{2.5}$ by using both an evidence-based and a risk-based approach, while necessarily *ad hoc*, was felt to be reasonable. Most Panel members favored continued use of the 98th percentile form because it is more robust than the 99th percentile form and therefore would provide more stability to prevent areas from bouncing in and out of attainment from year to year. Some concern was expressed as to whether EPA staff would exclude days on which natural phenomena such as forest fires distort the distribution. The Panel felt that such days should be eliminated before standard compliance is assessed. The link between the percentile form and the exposure level chosen is well-illustrated in the type of three-dimensional figures created by Dr. Miller at the April meeting (based on the data in Figure 5-2 in the 2nd draft PM Staff Paper), which were endorsed by the Panel and later provided in expanded form by OAQPS staff. The Panel endorses the inclusion of these types of figures in the Staff Paper. It would be helpful if reductions in risk associated with different regulatory options were expressed in the form of absolute numbers normalized to a fixed population size, in addition to those already expressed as percentage reductions.

In recommending revisions to the $PM_{2.5}$ NAAQS, changes to either the annual or the 24-hour standard, or both, could be recommended. Three arguments were made that support placing more emphasis on lowering the 24-hour NAAQS. First, the vast majority of studies indicating effects of short-term $PM_{2.5}$ exposures were carried out in settings in which $PM_{2.5}$ concentrations were largely below the current standard of 65 $\mu g/m^3$. Second, the amount of evidence on short-term effects, at least as reflected by the number of reported studies, is greater than for long-term effects. Third, toxicological findings largely relate to effects of short-term $PM_{2.5}$ exposures.

There was a consensus among the Panel members in agreement with the EPA staff recommendations that focused on decreasing $PM_{2.5}$ concentrations through lowering of the 24-hour PM standard, but the panel did not endorse the option of keeping the annual standard at its present level of 15 $\mu g/m^3$. It was appreciated that some cities have relatively high annual PM concentrations, but without much variation in concentrations from day to day. Such cities would only rarely exceed a 24-hour $PM_{2.5}$ standard, even if set at levels below the current standard. This observation indicates the desirability of lowering the level of the annual $PM_{2.5}$ standard as well.

Of the options presented by EPA staff for lowering the level of the PM standard, based on the above considerations and the predicted reductions in health impacts derived from the risk analyses, most Panel members favored the option of setting a 24-hour $PM_{2.5}$ NAAQS at concentrations in the range of 35 to 30 $\mu g/m^3$ with the $98^{th}$ percentile form, in concert with an annual NAAQS in the range of 14 to 13 $\mu g/m^3$. The justification for not moving to the lowest staff-recommended levels within these ranges is that these were generally associated with only small additional predicted reductions in risk. In addition, the uncertainties associated with concentration-response relationships increase greatly below these ranges, as reflected in substantial widening of the confidence limits for point estimates.

<u>The Panel had the following advice and recommendation for the $PM_{10-2.5}$ standard</u>:

It was acknowledged that the scientific basis supporting a causal role of $PM_{10-2.5}$ in an array of adverse health effects is weaker than that of $PM_{2.5}$. Regardless, most of the Panel members felt that the evidence that exists supports a causal role for health effects for $PM_{10-2.5}$. Moreover, setting this NAAQS would allow continuation and expansion of the $PM_{10-2.5}$ monitoring network that would facilitate collection of data for future exposure assessment and epidemiology studies. Because the evidence for the toxicity of $PM_{10-2.5}$ comes from studies conducted primarily in urban areas and is related, in large part, to the re-entrainment of urban and suburban road dusts as well as primary combustion products, there is concern that the associations of adverse effects with $PM_{10-2.5}$ may not apply to rural areas where the $PM_{10-2.5}$ is largely composed of less-toxic components of windblown soil or products of agricultural operations for which there is either no or limited evidence of health issues.

Further, although there is some evidence that short-term changes in concentrations of $PM_{10-2.5}$ are associated with changes in mortality, particularly cardiovascular mortality, the evidence in support of effects on morbidity, especially respiratory morbidity, is stronger. Most Panel members therefore favored not including short-term mortality effects in the health impact predictions, in line with the approach taken by EPA staff. The Panel agreed with Agency staff's approach of not setting an annual NAAQS for $PM_{10-2.5}$ at this time.

One of the major reservations expressed by the Panel in recommending a 24-hour $PM_{10-2.5}$ NAAQS related to the non-specificity of the $PM_{10-2.5}$ mass metric. Given that most evidence indicates that the component of the coarse fraction in most rural areas has little or no toxicity at environmental concentrations, it was felt important to qualify the $PM_{10-2.5}$ standard by somehow allowing exceptions for regions where the coarse fraction was composed largely of material that was not contaminated by industrial- or motor vehicle traffic-associated sources. Options discussed by members of the Panel for attempting to achieve this approach included limiting the standard to cover "all" urban areas, the judicious siting of monitors with a focus on urban areas, or regulatory exceptions for regions where road dust is not an issue or where rural components dominate the source. No single option was favored.

The panel also agreed that there was a need for more research on the health effects of $PM_{10-2.5}$. Such research will require the continuation and expansion of the $PM_{10-2.5}$ monitoring network in both rural and urban areas. The Panel recommends that the Agency staff expand and strengthen the discussion of the exposure index (size-range plus composition and/or source) and the monitoring strategy to be used for this NAAQS as well as the degree of public health protection expected relative to the protection against thoracic coarse PM afforded by the current $PM_{10}$ short-term NAAQS.

## WELFARE-BASED STANDARDS

Chapters 6 & 7: **PM-Related Welfare Effects**

Overall, these chapters are well done. Comments are provided below regarding vegetation and ecosystem, materials soiling, and visibility.

Considering the effects of PM on vegetation and ecosystems, EPA staff are to be commended for a well-written and concise reflection of the key science as presented in the final PM AQCD. The ecological risk assessment is reasonable given the required "criteria pollutant" approach. That being said, the criteria pollutant approach in this case (*i.e.*, PM) has serious shortcomings when it comes to ensuring environmental protection of vegetation and ecosystems in the U.S. This is illustrated in the following discussion.

Scientific evidence presented in the PM Staff Paper and the PM AQCD indicates that forest ecosystems at a number of locations in the U.S. "are now showing severe symptoms of nitrogen saturation." The Staff Paper makes the point that this is the result of chronic long-term additions of reactive nitrogen (Nr) species that have been accumulating over time. The PM Staff Paper also makes the point that the issue of forest-ecosystem deterioration is broader and more complex than just Nr accumulation. The Staff Paper notes that, "The most significant PM-related ecosystem-level effects result from long-term cumulative deposition of a given chemical species (*e.g.*, nitrate) or mix (*e.g.*, acidic deposition) that exceeds the natural buffering or storage capacity of the ecosystem and/or affects the nutrient status of the ecosystem." A key point implied here and elaborated later in the PM Staff Paper text is that PM deposition is only partially-responsible for the observed ecosystem-level effects and that the extent of the role of PM deposition in these ecosystem-level effects needs to be determined. While this has scientific merit, the question must be asked as to whether knowing the role of PM alone will improve the protection of vegetation and ecosystems in the U.S.? The answer to this question is critical because forest ecosystems are responding to the cumulative total load that has resulted from the chronic long-term deposition of both PM as well as gases and not to PM alone.

While EPA staff have done a commendable job within the context of the criteria pollutant approach, it is strongly recommended that in the future that Agency staff give serious consideration to a philosophical shift from the criteria pollutant approach to the European approach of "critical loads" when it comes to ensuring protection of vegetation and ecosystems in the U.S. The critical load is defined in the criteria document and is a quantitative estimate of an exposure to one or more pollutants below which significant harmful effects on specified sensitive elements of the environment do not occur according to present knowledge. The current criteria pollutant approach is a significant limitation in the efforts of the Agency staff to address the cumulative load of all the pollutant stressors to which ecosystems are responding.

Considering soiling and materials effects, several of the Panel members specifically asked EPA to add some discussion of the welfare effects caused by soiling from coarse particles. This may lead to consideration of a secondary $PM_{10-2.5}$ standard intended to protect against adverse welfare effects.

Considering visibility effects, most Panel members strongly supported the EPA staff recommendation to establish a new, secondary $PM_{2.5}$ standard to protect urban visibility. Overall, the Second Draft Staff Paper visibility sections (Chapters 6 and 7 and the detailed technical appendix by Schmidt *et al.*, 2005) are well-conceived and clearly-written. Agency staff can also be commended for responsiveness to comments previously submitted by this Panel on the PM AQCD and the First Draft PM Staff Paper. The recommended new standard was considered by most Panel members to be a reasonable complement to the Regional Haze Rules that protect Class I areas. The dissenting view is provided in one Panel member's individual review comments (see pages C-101 and C-102).

The recommended range of secondary standards includes an indicator ($PM_{2.5}$ mass), averaging time (4 to 8 daylight hours), level (20 to 30 µg/$m^3$) and form (90$^{th}$ percentile "or slightly higher"). The sub-daily averaging time is an innovative approach that strengthens the quality of the $PM_{2.5}$ indicator by targeting the driest part of the day. An indirect but important benefit will come from the direct use of — and more intense scrutiny on the quality of —the hourly data from the widely deployed continuous $PM_{2.5}$ mass monitors. The net effect is a "responsive" standard that (for the first time) would directly link public perception of air pollution (predominantly due to visual effects of light scattering by fine particles in the ambient air) to a routinely measured pollutant indicator (*i.e.*, artificially-dried $PM_{2.5}$ mass).

The recommended level and form of the standard are more difficult to specify. The draft PM Staff Paper employs a "bounding" approach, suggesting a level that is below the "obviously adverse" level of the current secondary standard — under which extreme short-term concentrations exceeding 100 µg/$m^3$ have been observed on days when 24-hour concentrations do not exceed 65 µg/$m^3$. Some members of the Panel felt the recommended level (and form) of the standard were on the high side, but developing a more specific (and more protective) level in future standards would require updated and refined public visibility valuation studies. Agency staff are strongly encouraged to support such studies prior to the next round of NAAQS review, even as it moves forward with the currently-recommended standard.

Some felt the recommended 90th percentile form of the standard was the weakest element of the EPA staff recommendation and the least well-justified. The visual effects of fine particle pollution provide the most direct public perception of air pollution of any regulated (or unregulated) pollutant, and the adversity of the effect is greatest on the haziest days that the 90th percentile would discard. Some Panel members recommend considering a higher percentile (92nd to 98th), accompanied by a level toward the upper-end of the recommended range, and/or otherwise softened by an “exceptional events” policy to assure that secondary non-attainment is not driven by natural source influences such as dust storms and wild fires.

To determine the degree of non-attainment that will result from a secondary standard, Agency staff should include — for different combinations of 4-hour and 24-hour levels and upper percentiles — estimates of concentrations and locations that would be expected to exceed a recommended secondary standard. EPA staff should also add some discussion of estimated “background” $PM_{2.5}$ conditions for the 4-hour daylight period.

In conclusion, the CASAC PM Review Panel encourages EPA in its efforts to protect the public health and our environment from the adverse effects of ambient air PM in the most effective manner possible. The Panel will continue to offer its advice and recommendations to help the Agency in meeting the mandates of the Clean Air Act and will review the final version of the staff paper with respect to EPA staff’s approach to setting a $PM_{10-2.5}$ standard. As always, the CASAC PM Review Panel wishes the Agency well in this important endeavor.

Sincerely,

***/Signed/***

Dr. Rogene Henderson, Chair
Clean Air Scientific Advisory Committee

Appendix A – Roster of the CASAC Particulate Matter Review Panel

Appendix B – Charge to the CASAC Particulate Matter Review Panel

Appendix C – Review Comments from Individual CASAC Particulate Matter Review Panelists

# Appendix A – Roster of the CASAC Particulate Matter Review Panel

---

## U.S. Environmental Protection Agency
## Science Advisory Board (SAB) Staff Office
## Clean Air Scientific Advisory Committee (CASAC)
## CASAC Particulate Matter Review Panel*

**CHAIR**

**Dr. Rogene Henderson***, Scientist Emeritus, Lovelace Respiratory Research Institute, Albuquerque, NM

**MEMBERS**

**Dr. Ellis Cowling***, University Distinguished Professor-at-Large, North Carolina State University, Colleges of Natural Resources and Agriculture and Life Sciences, North Carolina State University, Raleigh, NC

**Dr. James D. Crapo***, Professor, Department of Medicine, Biomedical Research and Patient Care, National Jewish Medical and Research Center, Denver, CO

**Dr. Philip Hopke****, Bayard D. Clarkson Distinguished Professor, Department of Chemical Engineering, Clarkson University, Potsdam, NY

**Dr. Jane Q. Koenig**, Professor, Department of Environmental Health, School of Public Health and Community Medicine, University of Washington, Seattle, WA

**Dr. Petros Koutrakis**, Professor of Environmental Science, Environmental Health , School of Public Health, Harvard University (HSPH), Boston, MA

**Dr. Allan Legge**, President, Biosphere Solutions, Calgary, Alberta

**Dr. Paul J. Lioy**, Associate Director and Professor, Environmental and Occupational Health Sciences Institute, UMDNJ - Robert Wood Johnson Medical School, NJ

**Dr. Morton Lippmann**, Professor, Nelson Institute of Environmental Medicine, New York University School of Medicine, Tuxedo, NY

**Dr. Joe Mauderly**, Vice President, Senior Scientist, and Director, National Environmental Respiratory Center, Lovelace Respiratory Research Institute, Albuquerque, NM

**Dr. Roger O. McClellan**, Consultant, Albuquerque, NM

**Dr. Frederick J. Miller***, Consultant, Cary, NC

**Dr. Gunter Oberdorster**, Professor of Toxicology, Department of Environmental Medicine, School of Medicine and Dentistry, University of Rochester, Rochester, NY

**Mr. Richard L. Poirot***, Environmental Analyst, Air Pollution Control Division, Department of Environmental Conservation, Vermont Agency of Natural Resources, Waterbury, VT

**Dr. Robert D. Rowe**, President, Stratus Consulting, Inc., Boulder, CO

**Dr. Jonathan M. Samet**, Professor and Chair, Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

**Dr. Frank Speizer***, Edward Kass Professor of Medicine, Channing Laboratory, Harvard Medical School, Boston, MA

**Dr. Sverre Vedal**, Professor of Medicine, School of Public Health and Community Medicine University of Washington, Seattle, WA

**Mr. Ronald White**, Research Scientist, Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

**Dr. Warren H. White**, Visiting Professor, Crocker Nuclear Laboratory, University of California - Davis, Davis, CA

**Dr. George T. Wolff**, Principal Scientist, General Motors Corporation, Detroit, MI

**Dr. Barbara Zielinska***, Research Professor, Division of Atmospheric Science, Desert Research Institute, Reno, NV

## SCIENCE ADVISORY BOARD STAFF

**Mr. Fred Butterfield**, CASAC Designated Federal Officer, 1200 Pennsylvania Avenue, N.W., Washington, DC, 20460, Phone: 202-343-9994, Fax: 202-233-0643 (butterfield.fred@epa.gov) (Physical/Courier/FedEx Address: Fred A. Butterfield, III, EPA Science Advisory Board Staff Office (Mail Code 1400F), Woodies Building, 1025 F Street, N.W., Room 3604, Washington, DC 20004, Telephone: 202-343-9994)

* Members of the statutory Clean Air Scientific Advisory Committee (CASAC) appointed by the EPA Administrator

**Immediate past CASAC Chair

# Appendix B – Charge to the CASAC Particulate Matter Review Panel

---

PM air quality information and analyses (Chapter 2):

1. To what extent are the air quality characterizations and analyses clearly communicated, appropriately characterized, and relevant to the review of the primary and secondary PM NAAQS?

2. To what extent have appropriate distinctions been made between fine and coarse-fraction particles with regard to properties of ambient PM, spatial and temporal patterns of ambient PM, and relationships between ambient PM and human exposure?

3. Does the information in Chapter 2 provide a sufficient air quality-related basis for the human health and visibility assessments presented in later chapters?

PM-related health effects, risk assessment, and health-based standards (Chapters 3, 4, and 5):

1. To what extent is the presentation of evidence from the health studies assessed in the PM AQCD and the integration of information from across the various health-related research areas drawn from the PM AQCD technically sound, appropriately balanced, and clearly communicated?

2. What are the views of the Panel on the appropriateness of staff's discussion and conclusions in Chapter 3 on key issues related to quantitative interpretation of epidemiologic study results, including, for example, exposure error, the influence of alternative model specification, potential confounding or effect modification by co-pollutants, and lag structure?

3. What are the views of the Panel on the adequacy and clarity of staff discussions on the potential existence of thresholds in concentration-response relationships in Chapters 3, 4 and 5? In particular, to what extent are hypothetical thresholds addressed appropriately in the sensitivity analyses conducted as part of health risk assessment?

4. To what extent is the assessment, interpretation, and presentation of the results of the revised PM health risk assessment (as presented in Chapter 4 of the draft Staff Paper and in the draft Risk Assessment technical support document) technically sound, appropriately balanced, and clearly communicated?

   a. In general, is the set of health endpoints, epidemiologic studies, and concentration-response functions used in the assessment appropriate for both $PM_{2.5}$ and $PM_{10-2.5}$?

   b. In particular, what are the views of the Panel on the staff's approach of not including mortality associated with short-term exposure to $PM_{10-2.5}$ levels in the quantitative risk assessment given the overall weight of evidence for this effect?

c. To what extent are the uncertainties associated with the risk assessment clearly and appropriately characterized in both the draft Staff Paper and draft Risk Assessment technical support document?

d. What are the views of the Panel on the adequacy of the various sensitivity analyses conducted to evaluate the influence of uncertainties in the risk analyses?

5. What are the views of the Panel on the broader approach taken by staff (as discussed in Chapter 5) of using both evidence-based and quantitative risk-based considerations in reaching conclusions and recommendations as to alternative suites of standards to protect against health effects associated with long- and short-term exposures for consideration in this review of the PM NAAQS?

a. Does the Panel generally agree with the emphasis given to the quantitative risk assessment results for $PM_{2.5}$, including consideration of risk estimates from base case and hypothetical threshold analyses, in reaching conclusions and recommendations for alternative suites of annual and 24-hour $PM_{2.5}$ standards?

b. Does the Panel generally agree with placing less reliance on the $PM_{10-2.5}$ risk assessment results and giving more emphasis to the available evidence from health studies in reaching conclusions and recommendations for alternative $PM_{10-2.5}$ standards?

6. Does the Panel generally agree that the alternative suites of primary standards for fine particles (including indicator, averaging times, forms, and ranges of levels) recommended by staff are generally consistent with the available scientific information and are appropriate for consideration by the Administrator?

7. Does the Panel generally agree that the alternative standards for thoracic coarse particles (including indicator, averaging time(s), forms, and ranges of levels for a 24-hour standard) recommended by staff are generally consistent with the available scientific information and are appropriate for consideration by the Administrator?

<u>PM-related welfare effects and welfare-based standards (Chapters 6 and 7)</u>:

1. To what extent is the presentation of evidence drawn from the PM AQCD related to the various welfare effects considered in this review technically sound, appropriately balanced, and clearly communicated?

2. To what extent is the characterization of the relationship between ambient PM and visibility impairment in urban areas scientifically sound and clearly communicated? In particular, what are the views of the Panel as to the methodology used to relate ambient $PM_{2.5}$ levels with reconstructed light extinction in urban areas across the U.S.?

3. Does the Panel generally agree that the local and state visibility standards and programs discussed in Chapter 6 are appropriate to help inform judgments as to the acceptability of varying levels of visibility impairment primarily in urban areas for the purpose of setting national standards?

4. Does the Panel generally agree that it is appropriate to consider using a fine particle mass indicator, specifically $PM_{2.5}$, as a basis for national standards intended to provide protection of visual air quality primarily in urban areas? Further, does the Panel generally agree that the alternative averaging times, forms, and range of levels recommended by staff for such standards are generally consistent with the available scientific information and are appropriate for consideration by the Administrator, in conjunction with the Regional Haze Program that is focused on protecting Class I areas from all man-made visibility impairment?

5. What are the views of the Panel as to the manner in which a risk-based framework has been used to organize the information presented in Chapter 6 on PM-related effects on vegetation and ecosystems?

6. What are the views of the Panel on the scientific soundness and usefulness of the discussion of the "critical loads" concept as a way to focus future research on the characterization, assessment, and protection of sensitive ecosystems?

# Appendix C – Review Comments from Individual CASAC Particulate Matter Review Panelists

This appendix contains the preliminary and/or final written review comments of the individual members of the Clean Air Scientific Advisory Committee (CASAC) Particulate Matter (PM) Review Panel who submitted such comments electronically. The comments are included here to provide both a full perspective and a range of individual views expressed by Panel members during the review process. These comments do not represent the views of the CASAC PM Review Panel, the CASAC, the EPA Science Advisory Board, or the EPA itself. The views of the CASAC PM Review Panel and the CASAC as a whole are contained in the text of the report to which this appendix is attached. Panelists providing review comments are listed on the next page, and their individual comments follow.

## Dr. Ellis Cowling

**Ellis Cowling** comments and recommendations for improvement of chapters 6 and 7 in the EPA-OAQPS document titled:

> "Review of the National Ambient Air Quality Standards for Particulate Matter, Policy Assessment of Scientific and Technical Information, OAQPS Staff Paper Second Draft, January, 2005."

**Chapter 6: Policy-Relevant Assessment of PM-Related Welfare Effects**

In general, this chapter is well done and includes a wealth of information about the principal welfare effects of PM in the lower atmosphere of the earth – visibility and regional haze, vegetation and ecosystems, damage to materials, and climate change processes.

I was especially pleased to see that the concept of critical loads for ecosystems was at least discussed in Section 6.3.6 on pages 6-55 through 6-58. But it also was very disappointing to see how very reluctant EPA Staff appear to be to engage in a serious and objective consideration of the advantages and limitations of this alternative system of place-based assessments of risks and benefits for ecosystems. I concur with the National Research Council 2004 recommendation that the concept of "Critical Loads" be considered more seriously (see footnote 4 on page 7-24). I also agree with Allen Legge's strong recommendation "that in the future, the 'Agency' give serious consideration to a philosophical shift from the 'criteria pollutant' approach to the European approach of 'critical loads' when it comes to ensuring protection of vegetation and ecosystems in the US."

Similarly, I recommend that the 'Agency' also give serious consideration to the "multiple-pollutant/multiple effects" approach that also is widely accepted within Europe.

Since my special competence is mainly in the realms of air-pollution effects on vegetation and ecosystems, most of my remarks and recommendations for improvement of this second draft of the PM Staff Paper are focused on various parts of Section 6.3 -- especially the effects of atmospheric deposition of PM-related reactive nitrogen (Section 6.3.3.1) and acidifying substances (Section 6.3.3.2).

**In Section 6.3.3.1**, it was especially good to see the Nitrogen Cascade diagram on page 6-34 and the generally very good discussion of reactive nitrogen (Nr) influences in both terrestrial and aquatic ecosystems on pages 6-32 through 6-41.

But some aspects of the discussion and the terminology used in the staff paper betray an incomplete understanding of the sometimes beneficial and sometimes detrimental effects of PM constituents on ecosystem processes. For example, the reactive forms of nitrogen include both reduced forms of this nutrient element ($NH_3$ and $NH_4^+$) and oxidized forms of this element (NO, $NO_2$, $NO_3^-$, $HNO_3$). Both reduced and oxidized forms of reactive nitrogen (Nr) [and sulfur] can have both growth-increasing and growth-decreasing effects on vegetation -- depending on the nutrient status of the ecosystem in which the PM is deposited. It is important to recognize that both reduced and oxidized forms of Nr will be taken up and used as nutrients that sustain the growth and development of all the plants, animals, microorganisms, and insects that inhabit ecosystems. Thus it is not always true that PM deposition of Nr will cause "stress" in ecosystems -- as implied by the titles and the discussions in Section 6.3.1, 6.3.2, 6.3.3, and 6.3.5. Better titles for these sections would be “6.3.1 Major Ecosystem Effects of PM, “6.3.2 Direct Effects of PM Deposition,” “6.3.3 Ecosystem Effects of PM Deposition,” and “6.3.5 Ecosystem Exposure to PM-Related Atmospheric Deposition.” Also, the effects of the added Nr [or sulfur constituents] of PM will not be the same on all the different organisms that make up the many different ecosystems of the world – some organisms and groups of organisms will be stimulated, others will be inhibited, and others will not be affected by the deposited PM.

The authors of this Second Draft Staff Paper seems to be much more cognizant of the effects of oxidized forms of Nr than those of reduced forms of Nr – this in spite of the fact that emissions of reduced forms of Nr from food production are generally 2-4 times larger than emissions of oxidized forms of nitrogen from combustion of fossil fuels (See “The Nitrogen Cascade” on page 6-34, and the summary paper by Galloway and Cowling, “Reactive Nitrogen and the World: 200 Years of Change,” AMBIO 31:64-71). Also, the ecological effects of reduced forms of Nr, although not identical, are roughly similar in terms of their “per-mole of Nr” effects on ecosystem processes.

Perhaps, the Staff Paper’s preoccupation with oxidized forms of Nr, rather than a balanced perspective about the importance of both reduced and oxidized forms of Nr, is a reflection of the fact that nitrogen oxides are a “criteria pollutant” but ammonia is not (yet) a “criteria pollutant.”

Preoccupation with oxidized forms of Nr is perhaps also a consequence of EPA’s traditional regulatory concerns with emissions from industrial and transportation sources such as power plants and motor vehicles (which are more often oxidized forms of Nr), and that agricultural emissions (which are more often reduced forms of Nr) have only recently become joint concerns of both EPA and USDA.

Also, the constant use of the term “reduced” and “reduction” with regard to pollutants generally leads to an impression that EPA staff believe it is OK to continue to ignore the distinction between the chemical and numerical meanings of the words “reduced” and “reduction.” It is very hard indeed to “reduce” ammonia -- in the chemical sense of the word!. In the interest of clarity of communication, it would be far better generally to use the unambiguous words “decrease” and “decreased” -- which have only numerical meanings -- rather than to continue the constant use of the ambiguous words “reduce” and “reduction” with regard to both emissions and pollutant concentrations. This comment and recommendation applies not only to Chapter 6, but also to all the other chapters in this Second Staff Paper on PM -- and, indeed, to many other publications by EPA and other organizations.

**Section 6.3.3.2.** In this Section as well, there are several aspects of the discussion and the terminology used in the staff paper that betray an incomplete understanding the acidification effects of PM constituents on ecosystem processes. For example, it was good to find the term “acidifying compounds” in the first line on page 6-43. The last line of page 6-42 properly indicates that these substances are “composed of ions, gases, and particles derived from the precursor gaseous emissions of $SO_2$, $NO_x$, $NH_3$ and particulate emissions.” Unfortunately, however, this one line on page 6-43 is the only place in this whole Section 6.3.3.2 (pages 6-42 through 6-58) where the term “acidifying deposition” is used.

In fact, the term “acidic deposition” is used in dozens of places throughout Section 6.3.3.2 where the emphasis is on “acidification processes” that are induced by “acidifying substances” that include ammonium ions ($NH_4^+$) that are not "acidic" at all -- but, in fact are just the opposite – they are indeed "basic" ions. The often poorly understood facts (among some atmospheric scientists and engineers, but <u>not</u> among ecologists) are that uptake of ammonium ions by plants results in the release of hydrogen ions that lead to acidification of ecosystems. Thus, atmospheric deposition of ammonium sulfate is twice as acidifying as atmospheric deposition of pure sulfuric acid.

A much more appropriate title for Section 6.3.3.2 would be "Environmental Effects of PM-Related Acidic and Acidifying Deposition" or perhaps even more appropriately "Environmental Effect of PM-Related Acidifying Deposition." In the interest of clarity of understanding, it also would generally be preferable to use the term "acid deposition" rather than "acidic deposition" and to more frequently use the term "acidifying deposition" throughout this whole section.

Specific editorial suggestions for improvement of Chapter 6 include the following:

| Page | Line | Change |
|---|---|---|
| 6-31 | 2 | "acidic precipitation" to "acidifying deposition" |
| | 14 | "acidic precipitation" to "acid precipitation" |
| 6-32 | 2 | "acidic deposition" to "acidic and acidifying deposition" |
| 6-34 | 8 | "particulate nitrates" to "ammonium and nitrate compounds" |
| 6-35 | 19 | "nitrate deposition" to "ammonium and nitrate deposition" |
| | 23 | "nitrogen" to "Nr" |
| | 24 | "nitrogen" to "Nr" |
| 6-36 | 15 | "nitrates" to "Nr" |
| | 17 | "nitrogen" to "Nr" |
| 6-40 | 19 | "N" to "Nr" |
| 6-41 | 7 | "nitrogen" to "Nr" |
| | 14 | "levels" to "amounts" |
| 6-42 | 10 | "nitrogen" to "Nr" |
| | 15 | "nitrogen" to "Nr" |
| | 20 | "Acidic" to "Acidifying" |
| | 21 | "Acidic" to "Acid" |
| | 25 | "Acidic" to "Acid" |
| 6-44 | 1 | "Acidic" to "Acid" |
| | 10 | "acidic" to "acid" |
| | 16 | "acidic" to "acidic and acidifying" |
| | 20 | "rain alteration" to "deposition altering" |
| 6-45 | 1 | "Acidic" to "Acid" |
| 6-46 | 9 | "acidic" to "acid" |
| 6-48 | 22 | "reductions" to "decreases" |
| | 26 | "reductions" to "decreases" |
| 6-51 | 23 | "levels" to "amounts" |
| 6-52 | 25 | "acidic" to "acidifying" |
| | 25a | "acidic" to "acid" |
| 6-53 | 10 | "levels" to "amounts" and "levels" to "amounts" once again |
| 6-54 | 14 | "levels" to "amounts" |
| 6-55 | 13 | "acidic" to "acid" |

| | | |
|---|---|---|
| 6-57 | 29 | "levels" to "rates" |
| 6-58 | 1 | "acidic" to "acid" |

**Chapter 7: Staff Conclusions and Recommendations on Secondary PM NAAQS**

Since EPA Staff have so often recommended adoption of equivalent primary and secondary standards for most of the Criteria Pollutants, I recommend that Chapter 7 begin with a brief section that provides a general description of: 1) EPA's rationale for so often adopting identical primary and secondary standards, and 2) the criteria EPA staff would consider necessary for adoption of a secondary standard different in form from the primary standard. This general discussion would be a useful background for the present recommendations with regard to possible secondary standards for PM.

# Dr. James Crapo

COMMENTS ON CHAPTER 4, OAQPS STAFF PAPER ON PM NAAQS

Jonathan M. Samet
James D. Crapo

March 30, 2005

## INTRODUCTORY COMMENTS

This chapter provides the general methodology, findings, and sensitivity analyses for EPA's risk assessment of $PM_{2.5}$ and $PM_{10-2.5}$. It is supported by the full Technical Support Document and associated appendices. The methods used in these documents have undergone review by CASAC as well as public comments. The chapter considers the morbidity and mortality burden associated with PM and the benefits of attaining the current standards, as well as several scenarios of more stringent standards. The findings of the risk assessment figure centrally in the recommendations of the Staff Paper.

## GENERAL COMMENTS

- One major concern with the current version of the chapter is the clarity of presentation. Readers need to struggle through dense prose and jargon-ridden prose to identify key aspects of the methods and findings. Concern about the document's style is more than cosmetic, as the risk assessment needs to be clearly presented so that there is no ambiguity as to its findings. In this regard, key terms are sometimes used incorrectly or inconsistently across the chapter. The chapter could be substantially shortened.
- Figure 4-1 provides an overall framework for the risk assessment that could be used to shape the chapter. It shows where sensitivity analyses are carried out and even numbers them by subscript. This potentially valuable framework is not subsequently utilized, however. We suggest that the chapter refer to it repetitively as the risk assessment methods and findings are described. The various sensitivity analyses might be listed in expansions of the "diamonds" on the figures.
- Subheadings might be more effectively used to guide the reader through the individual sections of the chapters. For example, clearly listing "assumptions" and "sensitivity analyses" so that the distinctions are clear and uniformly worded across sections.

## METHODOLOGICAL CONCERNS

- The selection of C-R relationships is premised in the concept that locally-derived coefficients are likely to be most appropriate. The Staff Paper mentions the possibility that the suite of potential confounding and modifying factors may vary from location to location. Is there a basis for assuming substantial variation? Is effect modification anticipated on the relative risk scale on which the risk assessment is carried out? There is evidence that coefficients from single-city time-series analyses tend to be biased upwards, in comparison to those from multi-city analyses (Dominici et al, in press).

Additional variability is introduced by variations in methods from analyst to analyst. These issues need discussion.

- In calculating the burden of associated incidence, the risk assessment uses either the predicted background or the lowest measured level in the utilized epidemiological analysis for the counterfactual. We suggest that the background level be used throughout to eliminate a needless difference in approach across locations. While there may be some further uncertainty in extending the C-R relationship beyond the lowest measured level, the larger uncertainty comes with the reliance on a linear, non-threshold model.
- The analyses in this chapter highlight the impact of the assumption of a linear nonthreshold model in overestimating actual risk. The absence of data near the threshold does not imply the absence of a threshold. Threshold models should be emphasized in this risk assessment. A major research need is for more work to be done to determine the correct threshold.
- Uncertainty receives comment throughout the chapter. Its inherent asymmetry needs acknowledgment; i.e., uncertainty is greater for scenarios set at lower and lower concentrations.

**SPECIFIC COMMENTS**

Page 4-2, first paragraph: There are methods for characterizing uncertainty beyond probabilistic judgments of “health scientists.”

Page 4-2, line 8: Confused sentence conceptually; Is the reference to statistical variability or to population variation—quite distinct concepts?

Page 4-2, second paragraph: See comments above. Ideally, a multi-location analysis would be done, if the data were available. Reliance on single-city analyses by individual analysts suffers from both variation in methods and limited precision.

Page 4-3, line 8: “precise measures” should be “certain measures”, one of many examples of careless wording.

Pages 4-6 and 4-7: The discussion of causality remains muddled. As a first question, EPA should determine whether $PM_{10}$ or $PM_{10-2.5}$ is causally associated with injury and adverse health effects and then select epidemiological or population indicators of the injury to health for use in the risk assessment. The sentence concluding the first paragraph on page 4-7 is not clear. There is also inconsistency in the chapter’s discussion of the level of causation inferred for $PM_{10-2.5}$ which is given as “causally related” here but “suggestive” elsewhere (see page 3-67, line 1; page 4-40, line 23).

Page 4-8, line 14: should read: “”..intended to provide protection from health effects of ambient PM.”

Page 4-27, line 22: would not use the phrase “mortality incidences” here or elsewhere in the document. Consider “mortality events”.
Page 4-53, full paragraph: The discussion of the basis for selecting the “thresholds” should be expanded.

## RESPONSES TO EPA QUESTIONS

- Question 3, PM-related health effects, risk assessment, and health-based standards (Chapters 3, 4, and 5).

Chapter 3 offers a general review of the epidemiological literature on thresholds (Section 3.6.6). The focus on this topic is applauded and the consideration of a threshold represents the largest factor in subsequent quantitative risk assessment in Chapter 4. This discussion reviews some of the relevant epidemiological literature but has no grounding in relevant toxicologic or mechanistic considerations. It does not lend direct support to the thresholds picked for sensitivity analyses in Chapter 4. A figure should be used to explain the slope adjustment in the "hockey stick" models.

- Question 4a.

In general, the set of health endpoints selected is appropriate and supported by relevant studies. We are concerned by the reliance on single-city analyses as a precedent and urge that multi-city analyses, once available, be used in future risk assessments. In this instance, there is not great variability across the C-R relationships selected.

- Question 4b

With regard to inclusion of mortality associated with $PM_{10\text{-}2.5}$ in the risk assessment, we are in agreement with not including such estimates. The epidemiological literature is mixed and there are inherent limitations to their findings, including the problem of measurement error for this derivative PM indicator and the difficulty of estimating a possibly separate effect from that of $PM_{2.5}$.

- Question 4c

- Question 4d

With regard to the handling of uncertainty in the risk assessment, an overview of the model is supplied in Figure 4-1, and key sensitivity analyses are indicated. Pages 4-37 through 4-41 offer a descriptive summary of the findings of these analyses. This section might be strengthened by adding the quantitative findings of these analyses, rather than including very limited verbal descriptions. It is unfortunate that a more comprehensive, quantitative characterization of uncertainty has not been undertaken, even if it only took into account several sources of uncertainty simultaneously. The chapter acknowledges this limitation of the risk assessment. There is also likely to be directionality to the degree of uncertainty, with greater uncertainty around effects at lower, compared with higher PM levels. Overall, the chapter tends to understate uncertainty, both through style, (e.g., inclusion of numerically specific estimates,. e.g., "403" deaths rather than "400" or "about 400", and by not bringing together the individual sensitivity analyses.

- Question 5. We agree with the general views and approach taken by staff in Chapter 5. We agree with the emphasis on the quantitative risk results for $PM_{2.5}$ and with a general approach on the use of $PM_{10\text{-}2.5}$ risk assessment.

- Question 6. We agree generally with the proposed alternatives for primary standards for fine particles. The range of proposed standards are consistent with the available scientific information.

- Question 7. We agree with the proposed alternative standards for thoracic coarse particles. The proposals are generally consistent with the available scientific information.

# Dr. Frederick J. Miller

## Chapter 3.

General Comments – Fred J. Miller

Overall, staff have done a reasonable job of summarizing the health effects basis for considering revised or new PM standards. There are a few areas that staff have either over interpreted or overstated the extent to which the health data support a particular PM indicator variable. While these instances are noted in the Specific Comments section below, a few of them are worthy of note here. The Summary Section on page 3-31 states results only using PM as the indicator variable with no size association, and as such, the Summary is not useful in establishing the case for any specific indicator variable that would be used for standard setting. This should be clarified by staff. Throughout the chapter, staff tend to overstate the case for $PM_{10\text{-}2.5}$ being associated or causative for specific types of health effects. For example, staff on page 3-33 make the statement that suggestive evidence is present for $PM_{10\text{-}2.5}$ on mortality, and yet the preceding discussion clearly shows multiple occasions where no effects of $PM_{10\text{-}2.5}$ were seen for both long term and short term mortality. To this reviewer, staff have made the case for annual and short term $PM_{2.5}$ standards based upon health effects reasonably associated with this indicator variable. However, the effects data presented and the interpretation of these effects would, to this reviewer, imply that an annual average standard for $PM_{10\text{-}2.5}$ is not warranted and that a short term standard for this indicator variable would have a wide range of uncertainty associated with it.

General Comments – Morton Lippmann, Ph.D.

Chapter 3 provides a reasonably unbiased view of the exceptionally large scientific literature on the health effects of ambient air PM that has been summarized in the PM CD. It addresses each of the issues relevant to PM NAAQS reconsideration with a review of the bases for the 1997 NAAQS, the advances in understanding of newer literature as described in the latest PM CD, and the Staff's recommendations for the significance and application of this collective knowledge in the need for new and revised PM NAAQS. If there is a bias in the treatment of this literature, it appears to be a leaning toward a public health protective stance, which I believe to be a reasonable perspective for EPA.

Our collective knowledge on the adverse health effects attributable to $PM_{2.5}$ has been greatly advanced since CASAC closure on the 1996 versions of the PM CD and Staff paper, and the 1997 decision to establish $PM_{2.5}$ NAAQS looks quite good in retrospect. Similarly, the 1987 decision to replace the TSP NAAQS with a $PM_{10}$ NAAQS has proven to be a judicious choice. The research and AIRS database that ensued from these PM NAAQS revisions greatly helped the epidemiologists and toxicologists produce much of the informative new scientific literature discussed in this PM Staff Paper. I review this historical background here because of the thorny issues we now face with regard to the establishment of new NAAQS for thoracic coarse particles ($PM_{10\text{-}2.5}$). There is no question that the scientific evidence supporting any specific form of one or more $PM_{10\text{-}2.5}$ NAAQS is far less extensive than we would like to have, and that we are therefore required to rely on expert judgment as much as on solid scientific data. The situation is highly reminiscent of 1985 and 1996 when CASAC encouraged EPA to move forward on the establishment of $PM_{10}$ and $PM_{2.5}$ NAAQS on the basis of what was known rather than on what we would have liked to have known.

I approach my recommendation on the Staff Paper's treatment of the $PM_{10-2.5}$ literature in the context of the limited options that Staff faces in consideration of the mandates of the Clean Air Act provisions on setting and reviewing NAAQS and the Supreme Court's directive not to use $PM_{10}$ as a means of controlling hazardous coarse mode particles. It seems to me that leaves EPA with no option other than a $PM_{10-2.5}$ NAAQS. While there is virtually no evidence that $PM_{10-2.5}$ is associated with annual mortality, and therefore no pressing need for an annual average NAAQS, I do believe that the weight of the evidence for adverse acute effects is sufficient to warrant public health protection against short-term peaks of $PM_{10-2.5}$ exposures. In this context, I believe that the Staff Paper has presented the available information in a quite reasonable fashion for CASAC and public comment at the April 6 & 7 review session.

<u>Specific Comments</u> – For the specific comments listed below, those page and line numbers in italics reflect the comments of Dr. Lippmann.

| | |
|---|---|
| *p. 3-2, l. 14-15* | Delete “though not the larger accumulation mode particles.” The statement is not true. Particles up to 2.5µm do have more alveolar zone deposition than T-B deposition. |
| *p. 3-5, l. 18* | Remove the hyphen in extra-thoracic. It should be extrathoracic. |
| p. 3-5, l. 25 | Change “patterns” to “fractions in these regions.” Ultrafine particles, depositing by diffusion will be deposited more uniformly within these regions than will coarse mode particles depositing in these regions by impaction. |
| p. 3-5, l. 26 | Using “removal of particles…” in this sentence is misleading. Typically, removal refers to the disposition of particles once they have deposited. For this sentence to be correct, the words “from the air” should be inserted after the word “removal,” or else the sentence should be rewritten to make clear that staff are talking about deposition of these particles. |
| p. 3-6, l. 3 | It should be “anatomical focus.” |
| p. 3-6, l. 4 | Strike the word recent. The phenomenon described in this sentence has been known for at least 20 years. |
| p. 3-6, l. 9 | Remove the comma. |
| *p. 3-6, l. 8* | Change “indicates” to “confirms.” |
| p. 3-6, l. 27 | This sentence provides circular reasoning since thoracic particles by definition are those that can penetrate to the thorax and therefore are |

| | |
|---|---|
| | available for tracheobronchial and alveolar deposition. I believe the authors are trying to make the point that particles penetrating to the thorax have a nonzero probability of also penetrating to the alveolar region. |
| p. 3-8, l. 8 | Remove the semicolons and replace with commas. |
| *p. 3-8, l. 17* | Add, at end of line "of particles attributable to facility (steel mill) operation." |
| *p. 3-8, l. 20* | Insert "residual oil fired" before "combustion." |
| p. 3-9, l. 5–8 | Remove the semicolons and replace with commas. |
| p. 3-9, l. 14 | Strike "on the heart" as this is redundant based upon the lead in of the sentence. |
| p. 3-10, l. 11 | Strike the comma. |
| p. 3-10, l. 14 | The sentence here needs clarification because hygroscopic particles greater than 0.5 μm in diameter grow in the respiratory tract while those less than this size will shrink. |
| p. 3-12, l. 22 | Strike the word "from" in this sentence. |
| p. 3-12, l. 24 | Strike the comma. |
| p. 3-18 | For the figure presented on this page and specifically for $PM_{10\text{-}2.5,}$ please find a way to indicate which of the $PM_{10\text{-}2.5}$ study estimates presented here are from studies where this indicator variable was specifically measured as opposed to derived by subtraction from monitoring $PM_{10}$ and $PM_{2.5}$. In addition, for any study that measured $PM_{10}$ and $PM_{2.5}$, it would be useful to compare the variance estimate for $PM_{10\text{-}2.5}$ compared to the variance estimate obtained from studies where $PM_{10\text{-}2.5}$ was measured directly. |
| p. 3-20, ¶ 1 | The staff here have quoted the expanded body of evidence on short term exposure to thoracic particles in mortality as being especially strong, but the statements are made relative to PM not to a specific PM indicator variable. The staff should specifically state the body of evidence and its relative strength for $PM_{10}$, $PM_{10\text{-}2.5}$, and $PM_{2.5}$. Otherwise, to this reviewer, the implications of the paragraph can be misleading. |
| p. 3-25, l. 21 | The first sentence of this paragraph is somewhat misleading in that there are only three studies in the figure for the overall category of respiratory diseases and they are indeed significant. However, the sentence implies that all of the hospital admission and emergency department visit studies |

| | |
|---|---|
| | in the figure are statistically significant, which is not the case. Clarification is needed. |
| p. 3-26, l. 9 | "Staff observes…" is a sentence that illustrates the need for consistency in the document of treating staff as either a singular collective noun or as a pleural and ensuring that the verb tense agrees with the interpretation of this. Basically, throughout the staff paper, the decision needs to be made of treating this in one of the two categories and then making all of the verb tenses consistent with that decision. |
| *p. 3-28, l. 25* | Change "cohorts, a cohort" to "cohort studies, cohorts." There were multiple cohorts within the southern California study. |
| *p. 3-28, l. 30* | Delete the last "the." |
| *p. 3-29, l. 1* | Change "cohort" to "cohorts."<br><br>The study by Gauderman that is described here is the kind of study for which the estimate taking the difference between $PM_{10}$ and $PM_{2.5}$ and computing the variance of the resulting estimate from a statistical approach would be worthwhile. |
| *p. 3-29, l. 1* | Change "cohort" to "cohorts". |
| *p. 3-29, l. 3* | Change "group" to "cohort" and "cohort" to study." |
| p. 3-31, l. 7 | In this summary section, results are stated only using PM as the indicator variable with no size association. To be useful, staff need to be specific as to the strength of evidence for the various indicators ($PM_{10}$, $PM_{10-2.5}$, and $PM_{2.5}$). It is insufficient to simply allude to the consistency of results with PM when the discussion that precedes clearly shows a much greater strength of the data for effects associated with $PM_{2.5}$ and $PM_{10}$ as compared to any with $PM_{10-2.5}$. |
| *p. 3-33, l. 1* | Insert "short-term" before "mortality."<br><br>How do staff justify the statement of suggestive evidence of effects of $PM_{10-2.5}$ on mortality when the preceding discussion clearly states on multiple occasions that there were no effects of $PM_{10-2.5}$ seen for long term mortality as well as short term mortality? |
| p. 3-35, l. 13 | Insert the word "had" between the words "there" and "been." |
| p. 3-36, l. 23 | Insert a comma after the word "indicators." |

p. 3-37, l. 23 This sentence indicating that there are not long term or chronic studies to air pollution is not correct. There have been numerous diesel studies, ammonium sulfate, ammonium nitrate, and other pollutant studies for particles found in urban air.

*p. 3-37, l. 24* Insert “ambient” before “air” and “mixtures” after “pollution.” (This is to acknowledge that chronic exposure studies have been done with laboratory generated aerosols.)

p. 3-39, l. 12 This reviewer does not agree that the limited body of evidence for $PM_{10-2.5}$ is suggestive of causality between short term exposures and mortality effects. Looking at Figure 3-1 on page 3-18, it is difficult for this reviewer to see how staff have come to the conclusion they state.

*p. 3-41, l. 11* Add (at end of the line) “with the exception of annual mortality, where there is strong evidence for an association with $PM_{2.5}$, and consistent evidence for its absence with $PM_{10-2.5}$.

p. 3-41, l. 28 The way this sentence is worded is confusing. The current version seems to imply that the incidence would decrease in the order of the types of endpoints listed since mortality is listed first.

*p. 3-42, l. 16* Change “PM” to “$PM_{2.5}$,” and add “Furthermore, the ACS cohort is more highly educated than the U.S. population as a whole and, adjusting for the lesser effects in the more highly educated component in that population, the longevity reduction for the U.S. population would increase.”

p. 3-44, l. 1 In this paragraph, the topic of transference of apparent causality is discussed. The staff dismissed this as being unlikely to exist in current studies. However, $SO_2$ and PM would seem to qualify for this condition since they are highly correlated and most likely collinear. Staff should reexamine this potential situation to see if the conclusions in this paragraph should be modified.

p. 3-45, l. 5 Staff discuss here how increased errors in $PM_{10}$ monitoring methods would likely have an impact of making it more difficult for epidemiological studies to have statistically significant associations detected between $PM_{10-2.5}$ and a specific health outcome. This paragraph is making the case that staff are going to rely on patterns of effects rather than requiring statistical significance for any of the individual study estimates. This logic represents the case for abandoning statistical probabilities and going forth with, for lack of a better expression, “whatever we feel like using.” This reviewer understands the need to be conservative in the protection of public health when uncertainties exist in the database. However, the extent of these uncertainties can be made more explicit. For example, in the studies where $PM_{10-2.5}$ is determined by subtraction, did investigators form this difference and then simply take the variance of that variable or did they treat both as random variables, which is the correct statistical procedure, and compute the variance of the random variable $z = x - y$ where x and y are the appropriate PM indicators? Doing so would result in taking into account not only the variance of the individual variables but the covariance between them. Certainly the strength of the database for any indicator variable for which the staff are wanting to propose new standards or changes to existing standards should be robust enough that there is consensus there is a clear signal, particularly since compliance with any PM regulations will be billions of dollars, and there is pressure on "not getting it wrong."

p. 3-46, l. 9 — In this section on exposure, staff discuss the ability of fine versus coarse particles to penetrate into indoor environments. Staff make the conclusion that studies indicate exposure measurements from central site monitors likely result in an underestimation of the effects of PM exposures on health. However, staff fail to acknowledge that the indoor pollution environment extends to many compounds not found in the ambient air — dust mites and other household organisms, sidestream cigarette smoke, etc. Most people spend 90% of their time indoors, and we are in the process of evaluating ambient exposure standards without treating the whole environment to which individuals are exposed. At a minimum, staff should at least acknowledge that the indoor environment complicates the interpretation of assigning some of the effects to outdoor measurements of PM.

*p. 3-48, l. 11* — Change "lead" to "led."

p. 3-50, l. 16 — Staff indicate that models using more stringent GAM criteria likely provide the most representative effect estimate sizes. Staff chose to use results from GLM based analysis to show associations in figures that are contained in this chapter. Doing such would seem to reflect a bias by staff to over represent building the case for various PM indicator variables showing effects on health outcomes unless in the risk assessment portion of the staff paper risk estimates are provided for two scenarios: (1) risk estimates using the more stringent GAM criteria, and (2) any other risk estimates obtained from other models such as the GLM based ones.

*p. 3-50, l. 20-22* — Change "PM" to "$PM_{2.5}$," and add "Furthermore, the ACS cohort is more highly educated than the U.S. population as a whole and, adjusting for the lesser effects in the more highly educated component in that population, the longevity reduction for the U.S. population would increase."

p. 3-57, l. 7 — To this reviewer, the conclusion by staff that the CD suggests cardiovascular effects may be associated with acute exposure time periods on the order of an hour or so is not warranted at this time given the paucity of the number of studies and the frequency of monitoring.

| | |
|---|---|
| p. 3-63, l. 1 | Staff have accurately characterized the issues relative to exposure time periods and long term exposure studies. However, this reviewer would like to raise the question concerning mobility of the population and the reasonableness that the long term mortality studies are capturing correctly PM effects given this mobility. The basic question becomes how accurate the current studies are in comparison to the results if one had the ability to track segments of the population as they were exposed to PM in different cities over their lifetime. |
| p. 3-66, l. 4 | The staff have done a good job characterizing the issues related to identifying population thresholds and have proposed a reasonable way to handle the situation via identifying a level of 12−13 mg/$m^3$ as a cutoff in their risk calculations and addressing other aspects of thresholds via sensitivity analyses. If this reviewer is interpreting Figure 3-4 correctly, any log relative risks below 0 represent no effect of $PM_{2.5}$ on the specific health endpoint being examined. Is this indeed the case? |
| p. 3-66, l. 21 | The summary and conclusions statements that are provided in this section reasonably reflect the discussion of types of effects and appropriate indicator variables for short term exposures to PM. However, the summary and conclusion section needs to specifically address the long term exposures to $PM_{2.5}$ and $PM_{10-2.5}$. To this reviewer, staff have made the case that long term mortality effects can be established for $PM_{2.5}$ but not for $PM_{10-2.5}$. This would imply that the annual average standard for $PM_{2.5}$ should be continued but that an annual average standard for $PM_{10-2.5}$ is not warranted. Clearly, the staff need to articulate their conclusions relative to long term exposures to PM. |
| p. 3-68, l. 21 | The words “older children: could easily be misconstrued. Change to “children studied from fourth grade to eighth grade”. |
| p. 3-85, l. 6 & 7 | The words “very high” could be misunderstood. Most CAPs studies used concentrations ~ 10 x ambient and nowhere near as high as most prior toxicological studies. |
| p. 3-85, l. 19 | Insert “mass” before “indicators.” |

**Chapter 4.**

General Comments

Most of my general comments will be presented at the meeting next week. However, one point I would make now concerns staff frequently using a range of percentages for a given PM standard scenario involving annual averages, 98$^{th}$ or 99$^{th}$ percentile daily standards, spatial averaging versus maximum value monitor, etc. It is the opinion of this reviewer that providing these percentages is not sufficient and that the actual numbers of cases or percent incidence change should be provided as well. For example, a change from 5 to 7-1/2 expected incidences is a 50% increase but probably does not have the public health concern of going from 400 to 500 expected incidences of a particular health endpoint. Reporting percent increases or decreases can be significantly misleading when the absolute numbers being used to compute those percentages are small.

Specific Comments

| | |
|---|---|
| p. 4-9, l. 28 | A reference is made to Appendix 4A for the LML value for each study. When one goes to Appendix 4A, it is evident that the decimal point is left off the annual average entry value for most of the monitors. In addition, the appendix should describe how the composite value for the annual average was obtained. |
| p. 4-14, l. 22 | Either change “is” to “are” in this sentence or make it “datum”. The same comment applies to the next line. |
| p. 4-16, l. 10 | Staff describe here the risk assessment on $PM_{2.5}$ for long term exposure mortality. The assumption of this reviewer is that the areas described here were also included in the risk assessment for $PM_{2.5}$ effects on short term mortality. If this is indeed the case, staff should specifically so state this in this paragraph prior to discussing it in the next paragraph. In addition, it would help the reader if reference to Tables 4-1 and 4-2 were made much earlier in Section 4.2.2.2. |
| p. 4-21, Table 4 | Table 4 would be more useful if staff would provide the range of 24 hr 98$^{th}$ percentile values for the individual monitors that are used to aggregate for a single value as well as the range of annual averages for these individual monitors. Failure to provide these estimates of the potential values will tend to provide the Administrator with a false sense of accuracy in the values that are used in the risk assessment. This is particularly true given that the geographical range of the location of monitors could be associated with significantly different population groups, and, therefore, their potential for adverse PM effects. Some of this information will be in Appendix A, but this information is important enough that it should be presented in the main document. |

| | |
|---|---|
| p. 4-21, l. 1 | This paragraph does not clearly convey what was the basis of having four of the five urban areas under one modeling approach and the remaining one using a different value for the annual standard. |
| p. 4-24, l. 18 | Why isn't B reflecting the incidence of the health endpoint at background PM levels rather than at 0 PM levels? |
| p. 4-25, l. 19 | Isn't there a way to check the reasonableness of this assumption by taking a site for which there is complete data and then randomly deleting values to see if the distribution ends up being as the authors suggest? |
| p. 4-29 | In Tables 4-5 through 4-7, staff should clarify what is implied by the dashes in some of the table cells. |
| p. 4-39, l. 14 | If the LML represents the lowest measured level, why one would not always have an LML available? More importantly, calculating risk down to the background level when an LML was not available is extremely problematic and to this reviewer not appropriate. The risk estimation procedure is inherently invoking a low dose linear assumption and belies whether or not a threshold at some low level exposure does indeed exist. If the Agency continues to pursue this line of risk estimation, then at a minimum, they should present what fraction of overall risk or total number of cases that they report is due to the component between a LML and background. |
| p. 4-53, l. 3 | The exhibits in Section D of the TSD only go up to D.42. Where are exhibits D.84 and D.86 through D.89? In addition, the overall range for the two other $PM_{10-2.5}$ locations that were studied should be listed in the staff paper rather than sending the reviewer to the technical support document, particularly since this would only involve a couple of lines of text. |
| p. 4-53 | The section on hypothetical thresholds illustrates conclusively the point this reviewer has repeatedly made concerning the influence of threshold level on overall risk estimates. Staff have taken "the easy way out" by stating on line 25 "a more definitive evaluation of the effects of hypothetical thresholds and use of alternative non-linear approaches would require re-analysis of the original health and air quality data, which is beyond the scope of this risk assessment." Given the estimated billions of dollars per year in cost for compliance with PM standards, this reviewer finds this reasoning lacking in merit, particularly given the magnitude of analyses that have been conducted and depicted in the TSD. |
| p. 4-54, l. 26 | Staff examined via sensitivity analyses $PM_{2.5}$ concentrations that were 50% higher or 100% higher than those used in the original studies relative to long term exposure mortality risk. How was the use of these levels |

determined? Stated differently, over what period of time do staff believe is a relevant length of time for the long term mortality in the as is scenario to be influenced by prior exposures? We know that $PM_{10}$ trends decreased about 22% from 1988 to 1995, so are staff postulating a 15–20 year influence on currently mortality from prior long term exposures or what?

The percent reduction in total incidence is misleading. Where values of 10–45% reduction are stated, the real consideration is what does this represent in number of cases for the particular endpoint be it a mortality or a morbidity one. This was done for Figure 4-10, and the results are very informative to the reader showing that the number of long term mortality deaths from exposure to $PM_{2.5}$ can be expected to be reduced from 0 to about 35 cases based upon the point estimates and up to about 100 cases incorporating the upper bound of the 95% confidence interval. However, a more informative metric would be useful such as number of cases per hundred thousand in population or some similar metric.

p. 4-62 Something appears to be incorrect when comparing Figure 4-10 to Table 4-11 where Figure 4-10 is stated to be long term exposure mortality for $PM_{2.5}$ and Table 4-11 contains the total mortality for long term exposure. About 100 cases are implied from Figure 4-10 whereas about 400 are reduced in Detroit and by approximately 1100 cases in Los Angeles. Staff should clarify which of these is correct.

p. 4-67, l. 16 What is the rationale for the declining increments of 5 $mg^3$ in the $98^{th}$ percentile daily values as an annual standard drops from 15 to 12 in steps of 1 $mg^3$? A proportionality rollback would have each decrement being about 2.7 $\mu g/m^3$ rather than 5 $\mu g/m^3$. Do staff have monitoring data that show the nonlinearity of this relationship rather than a linearly proportional one?

p. 4-68, l. 26 The reduction in additional estimated risk is miniscule for the scenario of the $98^{th}$ percentile standards of 30 or 25 $\mu g/m^3$. Rather than just making the statement that there are reductions, staff should quantify the percent reduction and the number of cases reduction, particularly since the figures are too small to accurately extract these numbers.

p. 4-74, l. 1 While this sentence is correct, staff should point out the confidence limits for this scenario are substantially greater than those for other $98^{th}$ or $99^{th}$ percentile daily standards.

p. 4-75, l. 25 The material provided on sensitivity analysis for alternative standards, specifically for hypothetical thresholds, gives short shrift to this topic that staff have identified as the single most important driver of considerations for the setting of PM indicator variable standards. No graphs are provided, and the reader is required to search through several appendices to extract the detailed information that should be provided either in tabular or more likely in figure form within the main body of the PM staff paper.

p. 4-82, l. 6 — This summarization is inaccurate. Staff need to qualify the conclusion here by indicating that among the choices examined, 25 provides the largest estimated risk reduction. The largest risk reduction will always occur the smaller the set point in $\mu g/m^3$ of the percentile daily standard.

**Chapter 5.**

General Comments

These will be made at the time of the meeting.

Specific Comments

p. 5-8, l. 7 — Staff acknowledge in this paragraph that the data available do not either support or refute the existence of thresholds for the effects of PM on mortality across the concentration ranges of the studies that are available. While this is appropriate, later on in this paragraph staff take the position that it would be difficult to detect thresholds and that studies that have tests for this haven't been able to distinguish between linear and various nonlinear models. The wording in this paragraph implies that staff will *a priori* consider analyses with the linear model. However, to biologists and toxicologists, the more appropriate *a prior* model would be nonlinear. The position taken on this issue will drive the basis of the PM risk assessments, the choice of the level of the standards, and the reductions in expected adverse health outcomes.

p. 5-10, l. 8 — In this paragraph staff discuss some evidence for long term mortality associated with $PM_{2.5}$. What do CASAC panel members think is a long enough time period for past exposure to influence long term mortality? The answer to this question has significant impact on what might be the appropriate level for a standard that is to protect against long term mortality since, for example, PM levels decreased 22% over the time period from 1988 to 1995.

p. 5-12, l. 27 — For the ACS study describing an annual level of 7.5 $\mu g/m^3$, over what time period did this annual level hold? Different weight should be given to this number depending upon the length of time over which it represented the average exposure to $PM_{2.5}$.

p. 5-16, l. 6 — The staff have indicated that they believe it is more appropriate to give the most weight to the base case risk estimates in the absence of evidence of clear thresholds. Since the risks associated if one invoked a 10 $\mu g/m^3$

exposure threshold range from 36–51% of those for the base case, is the same statement about thousands of premature deaths per year holding still correct? The concern here is that staff may be understating the extent of uncertainty associated with very low levels of exposure to PM.

More importantly though, as this reviewer has previously noted, the reliance on base case rather than threshold levels reflects a bias for linearity over nonlinear models for low level exposures — a position contrary to most biological mechanisms resulting in toxicity or adverse outcome. In fact, a range of plausible thresholds is more informative of the extent of uncertainty about risks and the consequences of various potential standards.

p. 5-25, l. 25 Staff note that they give consideration to the point where confidence levels become notably wider for the basis of the long term average concentration to use. Are staff recommending two standard deviations from the mean, one standard deviation, or what as a general strategy? This observation further strengthens the likelihood that there is an effective biological threshold for PM.

p. 5-27, l. 14 What risk level is predicted from the concentration response functions for a $PM_{2.5}$ value of 7.5 $\mu g/m^3$ in the ACS study compared to the mean value of this study, which was 14 $\mu g/m^3$?

p. 5-28 In Table 5-2, staff have used the base case as 7.5 $\mu g/m^3$, which represents a value 2.5 standard deviations from the mean of the representative study. This is equivalent to a 99.5% confidence interval on the lower tail of the $PM_{2.5}$ annual mean distribution. Given all of the uncertainties that have been discussed, selection of this as the base case is not warranted. For current standards, the 98$^{th}$ percentile of means would have to drop to 30 $\mu g/m^3$ or less, and, for the considered alternative ranges of 12–14 $\mu g/m^3$, decreased incidence would begin to occur at 40 $\mu g/m^3$ for the 98$^{th}$ percentile.

This table should stand on its own and include footnotes that would clarify, for example, that these calculations are based on a percent rollback using the highest monitor in an area. Table 5-2 will be a highly used reference table, and the reader should not have to go looking elsewhere for all of the information to interpret what exact scenario has been looked at. Some standardized health metric should be included in all tables so the reader can make useful comparisons.

p. 5-36, l. 22 Why repeat almost verbatim the same information that is contained on page 5-27, lines 1–9?

| | |
|---|---|
| p. 5-51, l. 25 | If a concentration based form is more reflective of the health risks posed by elevated PM concentrations because it gives proportionally greater weight to days when concentrations are well above the standard, then how do the staff reconcile this with the analyses that show that relatively little contribution to the overall risk is associated with the elevated short term exposure days for mortality? Moreover, as noted by Public comment on the CD, a short term standard invokes the assumption that nonlinearity in the C-R exists. |
| p. 5-58, l. 22 | "thoracic region of the lungs" should be changed to "thoracic region of the respiratory tract" since the thoracic region comprises the tracheobronchial and alveolar regions and is by definition the lungs. |
| p. 5-69, l. 12 | This paragraph is extremely speculative concerning the potential range from which a $PM_{10\text{-}2.5}$ standard could be selected. Specifically, the less than half values for the 98$^{th}$ and 99$^{th}$ percentile of the current daily $PM_{10}$ standard "could be interpreted as providing support for consideration of $PM_{10\text{-}2.5}$ standards that are less than half of the values" is highly speculative and not warranted in the opinion of this reviewer. |
| p. 5-74, l. 25 | The discussion given here for the range of a potential 24-hr $PM_{10\text{-}2.5}$ standard is inconsistent with the discussion on page 5-73, lines 17 and 18, whereby staff qualify their recommendations to indicate that they believe the uncertainties present in the database would lead to standard levels being towards the upper end of the ranges discussed. |

# Mr. Rich Poirot

## Review Comments on Chapter 7 of Janurary 2005 2nd Draft EPA PM Staff Paper: EPA STAFF CONCLUSIONS AND RECOMMENDATIONS ON SECONDARY PM NAAQS

R. Poirot, VT DEC, 3/29/05

Because of the strong, linear and causal relationship between ambient concentrations of fine particles and visual air quality, the possibility of a secondary PM-2.5 standard to protect visibility from adverse effects has been considered in several past EPA PM standards reviews, and was also discussed in detail in EPA's 1979 Report to Congress on "Protecting Visibility". In all of these past reviews, up through and including the August, 2003 first draft EPA PM staff paper, various obstacles have been identified such that staff recommendations for a secondary PM2.5 standard have been postponed for future consideration – typically pending "analysis of new data…" The current January 2005 2nd draft staff paper breaks from this tradition and includes extensive analyses of recently available urban PM2.5 speciation data and continuous PM2.5 mass data (presented primarily in Chapter 6 and the associated technical appendix by Schmidt et al., 2005), and specifically recommends consideration of a range of secondary PM2.5 standards at the present time.

Staff can be highly commended for their responsiveness to previous comments from this committee, as well as for the detailed and insightful exploration of options and clear illustrations with measurement data. That such illuminating measurements are available also reflects well on past EPA decisions for development and implementation of important new PM2.5 monitoring networks. The detailed Schmidt et al. report is especially remarkable considering the short time since the last round of review comments.

The proposed secondary PM2.5 standard is intended to protect visibility primarily in "urban" or non-remote areas (as the 1999 Regional haze rule pertains specifically to rural "Class 1" Federal National Parks and Wilderness areas). This is a very important consideration, as visibility impairment by regional haze in remote Class 1 Federal parks and wilderness areas and visibility impairment in urban and suburban areas are often very different effects. During winter for example, thermal inversions that may assure exceptionally clear visibility from higher-elevation Class 1 areas may also contribute to exceptionally poor visibility in the valleys below. The impracticality of having a single standard to protect both Class 1 areas and urban areas has been cited as an obstacle in past PM standards review cycles, and has been addressed in the 1999 EPA Regional Haze Rule by essentially normalizing each Class 1 area to its own current "baseline" conditions, based on the 2000-2004 period. Secondary PM2.5 standards are not needed to further protect these Class 1 areas, and can now be directed to provide protection elsewhere.

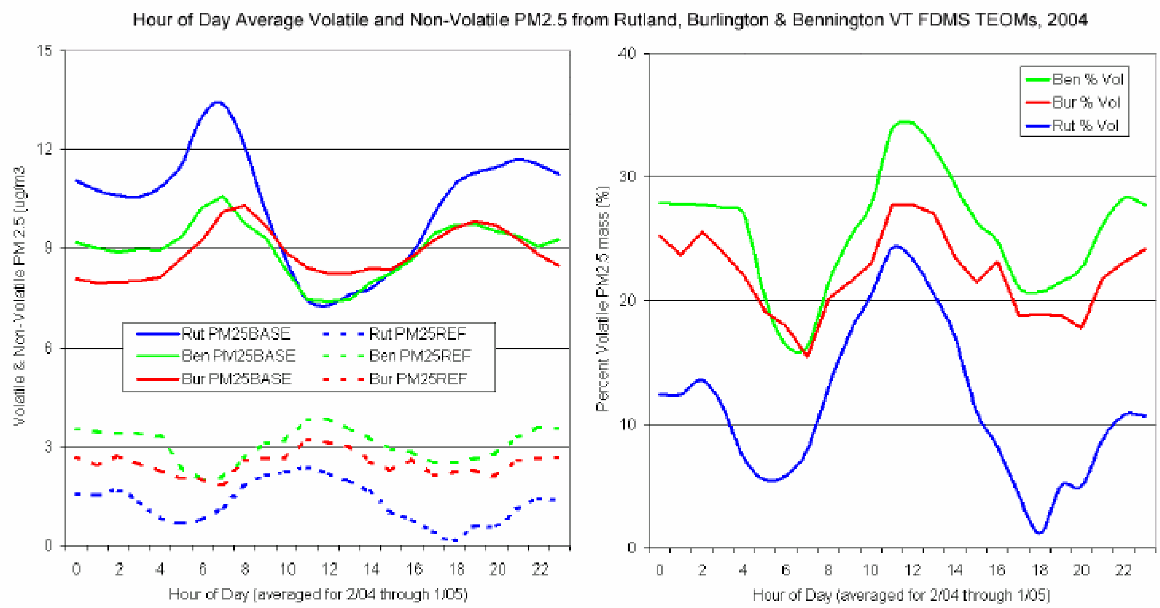
The proposed range(s) of secondary standards includes an indicator (PM2.5 mass), averaging time (4 to 8 daylight hours), level (30 to 20 ug/m3) and form (90th percentile "or slightly higher"). While each of these components can be considered separately, they also need to be viewed collectively, as each modifies the others. For example, the sub-daily 4 to 8-hour averaging time strengthens the (already tight) relationship between the PM2.5 indicator and the associated visual effect by sliding the "window of view" into the driest part of the day, when fine particles in the ambient air are most similar to fine particles as we operationally define them through artificially dried filters and continuous instruments. Compared to the current 24-hour (inadequate) secondary standard, the sub-daily averaging time also modifies the relative stringency of the proposed level of the standard, which in turn is directly modified by the percentile form of the standard. There are essentially an unlimited number of level/percentile combinations that would represent a standard of approximately equal protection and stringency.

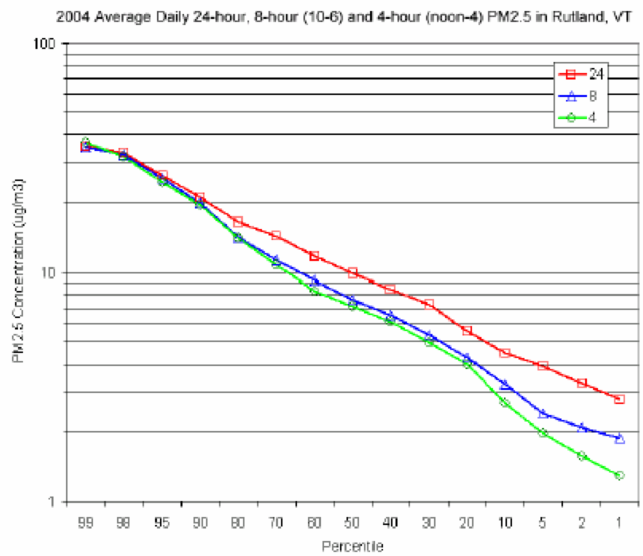
The principal justification cited for the sub-daily averaging time is the (drying effect) improved fit between the fine mass indicator and visibility effect, which (along with an urban focus) also substantially diminishes the differences between the Eastern and Western US. A remarkable set of statistics on p. 7-5 and 7-6 illustrate this point. Based on Class 1 IMPROVE measurements (and EPA regional haze guidance, which considers nighttime visibility in Class 1 areas of equal importance to daytime), average

(IMPROVE site) visibility on the haziest 20% of days in the East (20 km) is 5 times lower than in the West (100 km). It should also be noted that IMPROVE sites tend to be intentionally located at relatively high elevation – to be as regionally-representative as possible – and that these sites generally tend to be at higher elevations in the West than in the East. However, when similar calculations are made using the predominantly urban (and much lower elevation) Speciation Trends sites, the 400 % East/West difference is reduced to 33 %. And when similar calculations are made based on the (dryer) 4-hour daylight mean, the difference between the haziest 20 % days in the East (26 km) and the West (30 km) is only 15%. This relatively small difference does not seem sufficient to preclude setting of a single nationally applicable standard, which – for an equal concentration of PM2.5 – would tend to provide a slightly higher level of protection in the more arid west. Other regional or local differences can be accommodated by the (no required attainment date) flexibility associated with a secondary standard, and/or the option to set more stringent local standards in areas where visual air quality is considered more important.

In addition to this (drying) benefit of the proposed daylight-only averaging time, another benefit that might be considered & discussed is the relatively greater importance of daytime visibility (in urban/suburban areas). While nighttime visibility is important – especially for those seeking "the wilderness experience" in remote areas, it is arguably much less important than daytime in urban/suburban areas, when most of us are awake and there's something to see other than lights, which unavoidably impair urban/suburban visibility of anything else at night. For this reason alone, and also because "average" visibility impairment is not something people perceive, the shorter, (4-hour) daylight-only averaging time is well justified. The staff analysis also shows better regional consistency and fit with (dry) PM2.5 mass for the 4-hour mean, as an 8-hour mean is likely to include more humid (or dark) hours at northern latitudes in winter.

An additional indirect benefit of the sub-daily averaging time, which would presumably be based on the extensive and growing network of continuous PM2.5 instruments, is that it would both make use of, and also force closer scrutiny on, the quality of that data over less than 24-hour averaging times. While there has been a substantial national investment in continuous PM2.5 monitoring (with more proposed under NCORE), and while these data are routinely reported to the public via AIRNOW, and will soon be used as a basis for short-term health assessments, there has been very little attention paid to the quality (precision & accuracy) of the hourly data. "Adjustment equations", which vary considerably by state and method, are used to fit the aggregated 24 hours of hourly data to the (aggregated) 24-hour FRM filter data, but none of these adjustments (that I know of) considers the quality of or potential biases in the hourly data. The degree to which the various continuous instruments – which typically all have heated inlets to remove water – modify the ambient aerosol by reducing other non-water volatile PM species in not well characterized. It is likely, however that such volatile losses will tend to vary by region, season, hour of day, species and source type. Adjustments based on 24-hour aggregations will tend to over-adjust the non-volatile species and under-adjust the volatile species, and health studies based on this data will tend to mischaracterize diurnal exposure patterns and source influences. This diurnal variability in volatile vs. non-volatile species is illustrated in Figure 1 below, based on FDMS (Filter Dynamic Monitoring System) TEOM data from 3 VT sites. Note that the volatile component of the FDMS-defined mass ("REF") tends to peak in an absolute (left) and relative (right) sense at mid-day. During daylight hours this volatile fraction can, on average, represent 10 to 30% of the total mass. This hourly volatile fraction is also the answer to a different question – which is: what is the difference between hourly fine mass as it is defined by a 30 degree TEOM and an FDMS TEOM? Both are commonly deployed, along with a number of other instruments, in ambient monitoring networks. EPA guidance on specific methods and any data adjustment methods will need to be provided to implement any sub-daily standards. This is not to suggest that sub-daily filter sampling should be required or that data quality uncertainties should be held up as an obstacle for standards promulgation. The hourly data are in fact being used for other reasons, and there's nothing like making them somewhat "important", as in use for a secondary standard, to force a much more careful look at, and evaluation and understanding of their quality.




It can also be noted that a sub-daily averaging time will affect the relative stringency of a standard, but will not necessarily result in a more stringent (or protective) standard than a 24-hour averaging time for any given level (concentration) and form (percentile). Figure 2 shows cumulative frequency distributions of daily 24-hour, 8-hour and 4-hour daylight averages for the past year of FDMS TEOM data from Rutland, VT. It may be noted that concentrations are similar for all averaging times at the higher percentiles ($98^{th}$, etc.), but tend to be lower for the shorter averaging times at lower percentiles. These relationships are likely to vary by site and season, but in general, it seems likely that the 4 or 8-hour levels will not be hugely different from 24-hour averages for higher percentile forms of a standard. It can also be noted in Figure 2 that the level and form of (any) standard can't be viewed in isolation from each other. For any given averaging time, all plotted combinations of level and form would, for this site, represent an equally stringent standard.

The proposed level of the secondary standard is one of the most challenging aspects of the (public policy) decision process. While the relationship between (daytime) PM2.5 and visibility impairment is extremely tight, well-characterized, linear and essentially has no lower threshold, the selection of any specific level of desired protection (an "adversity threshold") is more complex, variable, and necessarily dependent on subjective value judgments. See attached 3/28/05 memo and summary table from Robert Rowe for additional detail on proposed the level and form of the secondary standard.

A somewhat parallel issue exists in the selection of annual and short-term primary health standards in the absence of a clear-cut lower effects threshold. For secondary visibility protection, no "margin of safety" is required and economic costs and benefits would need to be considered in the implementation phase. As Dr. Rowe's memo and summary table indicate, economic valuation studies – especially in the eastern US have not been conducted or updated in recent years, and it would be highly desirable for EPA to support some of this kind of analysis in the near future.

EPA's general approach in proposing a level for a secondary standard includes several reasonable "bounding conditions". The current short-term secondary standard can be clearly observed to allow visibility impairment which is clearly "adverse" – especially considering that shorter-term 1-4 hour PM concentrations can and often do exceed 100 ug/m3 on days when the 24-hour concentration does not exceed 65. EPA also sites several examples of (primarily Western) local visibility standards (all from relatively "scenic" areas like Denver, Phoenix and Lake Tahoe, where its reasonable to assume a higher level of protection is desired than would be the case nationally). EPA's proposed levels in the range of 20 to 30 ug/m3 seem to fall reasonably in between these two bounding extremes. EPA also sites the potential for (natural) background conditions to impose a lower limit on the level of the standard, and further includes a series if illustrative photographs from several urban areas that show compelling changes in visual air quality above and below the range of standard levels they have proposed. In my view, the proposed levels are reasonable and reasonably derived.

The proposed level is substantially modified (softened) by the proposed form of "$90^{th}$ percentile or slightly above". While a somewhat "softened" standard may be desirable to balance costs, benefits and the current uncertainties over the "adversity" of effects (how much new secondary non-attainment can we handle?), it should also be noted that any secondary standard is inherently softened by the Clean Air Act. Primary standards are to be attained in 5 years (or up to 10 years at the Administrator's discretion), but secondary standards have no required attainment dates. Thus there is no logic at all for the convenient, recent approach of setting a secondary standard equal to the primary standard. The pace of the progress toward secondary attainment can vary by location, severity of exceedance, and associated local and regional costs and benefits of control. In my view, the use of the (very soft) $90^{th}$ percentile as the form of a visibility standard is not especially well-justified if visibility protection is the desired goal.

In considering alternative forms for the secondary standard (section 7.3.5, page 7-13), staff determines first that a concentration-based form is desirable, in part "because it gives proportionally greater weight to days when concentrations are well above the level of the standard than to days when the concentrations are just above the standard." Staff also argues, "a concentration-based form better compensates for missing data and less-than-every-day monitoring". However, the proposed 4-hour averaging time would represent the first effort to rely specifically on the continuous data, which aren't characterized by less-than-every-day monitoring, and the proposed 90th percentile form would essentially disregard the worst 108 days in a 3-year period, regardless of how high they were above the standard.

Having thus concluded that a concentration-based form, averaged over 3 years, is needed, (which I think is probably reasonable to provide for relatively stable compliance metrics), staff further argues (p. 7-14) that "constraints on the number of days in which a standard can be exceeded should be appreciably tighter for a standard intended to protect against serious health effects than would be appropriate for a standard intended to protect against visibility impairment". I think this may sound more logical than it really is. Of course health effects are more important than aesthetics, but the unlimited time frame for remedying secondary non-attainment clearly accommodates this, and I see no logic to singling out this specific element – the form – of a secondary standard as the component that should be more lenient than a for a primary standard.

On the contrary, sensitivity analyses for many of the health studies cited in the CD and in earlier sections of the staff paper indicate that adverse health effects result from concentration changes across the entire range of PM2.5 concentrations, and are not uniquely driven by the most extreme events. But for visibility, the adversity of the effect increases with PM2.5 concentration. Impaired visibility, due predominantly from light scattering from fine particles, and most notable to most people during daylight hours, IS the primary way in which the public perceives air pollution. In Vermont's small state government, I end up taking most of the e-mails, letters to the Commissioner, and phone calls, etc. from concerned citizens if they pertain to poor visibility conditions. People call and say things like "I can't see across the lake today; what's going on?" The 90th percentile isn't a humanly perceptible concept, and would, regardless of "stability", specifically discount the most adverse and perceptible conditions.

Some justification for the 90th percentile is provided (p. 7-14) by invoking intent for consistency with the regional haze program, which includes a focus on the haziest 20% of the days. The 90th percentile would be the middle (the median actually) of the worst 20% days. However, regional haze in class 1 areas and visibility impairment in urban valleys are often decoupled from, and at times opposite to, each other. The haze rule includes additional focus on protecting the cleanest 20% days, and the metric for assessing and tracking (the cleanest and) haziest 20% days is the mean, not the median. The 92nd percentile or slightly higher would be a better starting point if consistency with the regional haze program was the goal. The haze rule also requires that active steps be taken over time to reduce the (average of) the current "baseline" worst 20% days at all sites, down to the level of "natural background" – an ambitious goal, but over a relatively long 65-year time frame, with periodic checkpoints to assure that progress is being achieved. The baseline period (2000-2004) has recently ended, and we are now in the first period of required improvement, with SIPS due between now and 2008, to assure the first progress checkpoint by 2018. At that time, progress of approximately 20% of the difference between recently past (2000-2004) conditions and 2064 "natural background" conditions is required by the haze rule. Thus, whatever PM2.5 concentrations are associated with current (approximately) 92nd percentile conditions are required to improve significantly such that that associated PM2.5 concentration will need to represent a substantially higher percentile in the fairly near future. If consistency with the regional haze program is desired, a secondary PM2.5 standard should be pointed toward where we are headed, and not backwards at a point in time that is already behind us.

For the above reasons, I would encourage staff to consider a form of the secondary standard that is significantly above the 90th percentile. If compromise is necessary, a level toward the upper end of the proposed range, combined with a higher percentile – perhaps 98th to maintain consistency with the primary standard (if it remains at this percentile). This would help make the form of the standard more consistent with human perception of adverse effects, and would also assure that the level of the standard exceeds a level judged to be adverse over the widest possible range of locations.

I would also encourage EPA to move forward on the proposed secondary standard in the current PM standards review cycle, while at the same time supporting the updated evaluations of economic costs, benefits and perceptions of adversity that will be needed to determine the appropriate pace at which progress toward attaining these standards is reasonable in different regions and urban areas.

**Minor Comments on Chapters 6 & 7:**

p. 6-16, line 1&2: "Aesthetic benefits of better visibility also include improved road and air safety". Change "Aesthetic" to "welfare".

p. 6-46, Aquatic effects: consider adding a reference to and discussion of :
VT DEC (2003) Total Maximum Daily Loads for 30 Acid-Impaired Vermont Lakes, Vermont Department of Environmental Conservation, Water Quality Division.
http://www.anr.state.vt.us/dec/waterq/planning/docs/pl_acidtmdl.pdf
This was required by and approved by EPA in 2003, and includes calculations of critical loads and exceedances for S & N deposition to 30 acid-impaired VT lakes.

p. 7-1, line 8: I suggest adding "intended by statute" between "standards" and "to address", as staff later concludes that current standards do not in fact adequately address visibility impairment associated with fine particles.

Section 7.3.1, Section 6.2.5.2, or elsewhere: The existing local programs for protecting visibility in Denver, Phoenix, Lake Tahoe, VT and elsewhere are nicely summarized here. But it might be useful to include somewhere some discussion on the implementation of, and consequences that may have resulted from, implementation of these standards. In VT for example, new source applications are evaluated for potentially significant contributions to exceedances of our state visibility standard, and/or permit applications are likely modified in advance to avoid significant impacts. In Denver, I understand (personal conversation with Dan Ely) that the standards lead to visibility forecasts and associated "no burn days" for residential wood burning, carpool incentives, etc. and that there were resulting improvements in visibility and reductions in PM concentrations from these standards well in advance of the establishment on national PM2.5 standards.

## Dr. Frank Speizer

**Critique of Staff Paper on PM—Jan 2005—Comments of Frank Speizer on Chapter 5 (Recommendations for primary NAASQS.)**

General comment: The chapter is substantially strengthened from the previous draft, particularly in the way it takes seriously the request of CASAC to deal with the 3 specific issues involved in the setting of a standard: eg. Indicator, Averaging times, and Statistical form. Staff does a good job of summarizing the CD finding supporting the potential causal association of PM and significant health effects and certainly extends the finding from the 1997 status to provide less uncertainty of the significance of those findings. This is particularly the case for PM2.5. With regard to the course fraction, the document tries to deal with the remaining uncertainties but leaves open the possibility that the standard would be more arbitrary and would have more remaining uncertainty.

There are at least two important constraints under which Staff has had to operate in preparing this document. Both seem to have been imposed by law. One is the fact that PM10 could no longer be considered a criteria pollutant and other is that cost is not supposed to be a factor in suggesting a standard. Dr. Vedal has offered a detailed assessment of the role of cost in constraining the recommendation.

With regard to the issue of not being able to use the PM10 findings, this reviewer believes that they still provide important clues as to potential health effects that cannot be ignored. Unless one believes that all of the action of PM10 reported in the past was due to PM2.5, we must have a standard that includes the course fraction. PM10-2.5, although not optimal, is not only justifiable but also necessary to protect the public health with an “adequate margin of safety”.

By providing a range of levels for both the short and long term PM 2.5 along with an assessment of the statistical form of the standard, the administrator will be in a position to choose the trade offs.

The situation is less clear for PM10-2.5. We must deal with the uncertainty in choosing a standard, “with a margin of safety”. I think the Staff paper has justified this position and therefore the matter becomes one of choosing the “right numbers”. One suggestion that is partially covered in Chapter 3 where direct comparison of finding of PM10, 2.5, and 10-2.5 are summarized is that the apparent effect of PM10 is not fully accounted for by PM2.5. The degree that these findings show heterogeneity is an indicator of the uncertainty and must be taken into account in setting the PM10-2.5 level and statistical form. Just how far to judge the degree of non-compliance that will result in whatever decision is taken should take a secondary roll in the decision.

Specific Comments:

Page 5.14, Table 5.1:

Insert in title of table that Mortality is an incidence figure by adding "per 100,000" (if that is the right denominator).

Page 5.19, lines 21-28. This is an important discussion point that results in the judgment that the form of the standard must remain at a 2.5 um cut point. It clearly represents a compromise if one is setting a national standard and is the best judgment call as long as regional standards are not possible.

The lay out of Sections, 5.3.2 (Indicator); 5.3.3 (Averaging Times); and 5.3.4 (Short and Long term considerations) provides as excellent logical basis by which the staff decisions parallel the CD.

Page 5.28, 5.29, Table 5.2. Table layout. The shading of the first column in each section by the 12-15 annual average figures, then separated into the 98 and 99 percentile is confusing. Suggest space the 98 and 99 percentiles so it is clearer that this is a repeat of the same standard level.

Page 5.36, lines 16-28. Staff is concluding that short term standard be between 13 and 12ug, bringing in both demonstrated harm and margin of safety for mortality.

Page 38-41 Tables 3a and 3b. I do not find these tables compelling. There simply indicate that there are regional differences in being able to meet any proposed standard. I think we knew this before and this simply raises the issue to a level that provides grounds for people to complain. I do not see how it should or does affect the judgment on making the call as to form and timing of standard.

Pages 5.43, table 5.4. In contrast to above this puts a more intuitive reality of what the different levels of standard do to mortality risk rather than compliance risk. As discussed on page 5.45 these estimates of risk reduction are real and substantial across the board.

Page 5.65, line 9-10. The lack of evidence is really a lack of studies of long term effects of specific PM 10-2.5 measurements. Therefore to say no need for an annual standard seems inappropriate unless one can assume that all of the data from findings of annual PM10 was from the unmeasured PM2.5. I am not comfortable with that assumption.

Page 5.69, line 1-5. Not clear how approximate PM 10-2.5 level of 65-85 became equivalent to a PM 10 of 150. That would mean that PM2.5 was 75. Not likely.
Suggest we have EPA Staff produce Table 5.6 for PM10 for comparison purposes.

Page 73, para 8-18. In discussing the uncertainties it seems to me that a comparison with PM10 is needed (as is contained in Chapter 3 and shown in the appendices to Chapter 3), as this would mitigate against some of the lack of data rather than lack of evidence.

Page 5.78 List of key uncertainties and research questions. Need to add paragraph on susceptibility and interaction with personal behavior and or genetics.

# Dr. Barbara Zielinska

**Comments on the EPA-OAQPS Staff Paper for Particulate Matter (2nd Draft).**
**Chapter 2: Characterization of Ambient PM**
**Barbara Zielinska**

In general, Chapter 2 is very well written and presents an accurate and concise summary of Chapters 2, 3 and 5 of the PM Criteria Document. My answers to specific charge questions to CASAC PM Panel from Dr. Martin's (OAQPS) memo of January 31, 2005 are as follows:

> <u>Question 1</u>. *To what extend are the air quality characterization and analyses clearly communicated, appropriately characterized and relevant to the review of the primary and secondary PM NAAQS?*

In my opinion, the summary of ambient PM properties (Section 2.2), measurement methods (Section 2.3), concentrations, trends, spatial and temporal patterns (Section 2.4 and 2.5) and background levels provide an appropriate background for the review of the primary and secondary PM NAAQS contained in Chapters 5 and 7. I have a few specific comments regarding these Sections:

a) A brief discussion concerning chemical composition of $PM_{2.5}$ and $PM_{10\text{-}2.5}$ in Sections 2.2.3 (Chemical Composition) and 2.4.5 (Components of PM) do not adequately portray the compositional differences of PM depending on their sources. Although organic carbon (OC) is listed as one of the main components of ambient PM, the lack of information concerning the composition of OC is not mentioned. OC fraction is especially abundant in fine and ultrafine PM (as it could be seen from Figure 2-16) and it could play an important role in PM toxicity. Clearly, the investigation of OC composition should be identified as one of the important future research objectives. Also, there is nearly no discussion concerning the chemical composition of $PM_{10\text{-}2.5}$ in different areas of the country and the heterogeneity of this size fraction of PM.
b) PM mass measurement errors could be better characterized in Section 2.3.5 (Measurement Issues), especially in relation to PM $_{10\text{-}2.5}$ measurements. Since the majority of $PM_{10\text{-}2.5}$ is determined by subtraction, it would be important to state how uncertain these concentrations are.
c) Since the proposed averaging time for the secondary PM standard is a subset of 24-hr period (i.e. 4 to 8 hrs), it would presumably rely on continuous PM monitoring methods. Thus, it would be useful to discuss the current limitations of these methods in Section 2.3 and emphasize the need for their improvement and further development.

Section 2. 7 discuss to what extend the centrally monitored PM concentrations represent surrogate of human exposure to ambient $PM_{2.5}$. My concerns are:

a) The spatial gradient of fine PM concentration is not adequately characterized in this section. Recently, there have been numerous reports (including data from the Los Angeles PM Supersite, Zhu et al., 2002) that the ultrafine and fine PM concentrations at the source-dominated ambient locations (e.g., along freeways and surface arterials) are many times higher that those measured at "central sites". For example, the peak $PM_{2.5}$ concentration of 90 $\mu g/m^3$ was measured in the Terminal Island area (diesel-dominated port area in Long Beach, CA) and this is about an order of magnitude higher than the local background PM concentration of about 9 $\mu g/m^3$ (Fujita et al, 2005). Similar differences were observed on freeways, along the major truck routes in the LA basin on weekdays, when the proportion of truck traffic was high. Clearly, the commuters and the residents of areas adjacent to the major arterials are exposed to much higher PM concentrations than those measured by "central sites".

b) My other concern is the lack of acknowledgement in this document of the importance of indoor sources to potential health effects. Indoor environment typically contains many potentially toxic pollutants (such as cigarette smoke, dust mites, etc.) that may complicate greatly the evaluation of human exposure to ambient PM. Although it is true that the PM of indoor origin does not influence exposure to PM of ambient origin, how one can separate these effects as far as human health is concerned?

c) The staff states on page 2-68: "Although the spatial variability of PM2.5 varies for different urban areas, overall, some degree of uniformity results from the widespread formation and long lifetime of the high regional background of secondary $PM_{2.5}$". I suppose staff means secondary sulfates and nitrates, since there is really no information concerning the composition of secondary organic aerosol in different areas. However, are nitrates and sulfates really important as far as the health effect is concerned?

_Question 2._ *To what extend have appropriate distinction been made between fine and coarse fraction particles with regard to properties of ambient PM, spatial and temporal patterns of ambient PM, and relationships between ambient PM and human exposure?*

To the extend the information is available, the staff paper presents appropriate characterization of coarse particles and emphasize the differences between $PM_{2.5}$ and $PM_{10-2.5}$. However, the problem is that there is really not enough data concerning the spatial variability of the coarse particle composition and the effect of this variability on human health. Coarse PM composition is very different in different settings and may vary from the dominance of crustal material to the road dust with adsorbed motor vehicle emissions and biological material.

_Question 3._ *Does the information in Chapter 2 provide a sufficient air quality-related basis for the human health and visibility assessment presented in later chapters?*

Section 2.8 explains briefly the relation between ambient PM mass and visibility impairment. I found this section relevant to the proposed secondary standard deliberation in Section 7. My minor comment is that there has been some discussion recently concerning the IMPROVE algorithm for calculating light extinction, particularly in relation to the 1.4 factor for organic carbon, that is not mentioned in this section. Also, the problem of the gas-particle partitioning of semi-volatile compounds and its relation to true ambient PM concentrations is not discussed here.

Minor comments:

1. Page 2-16, Section 2.3.2. The abbreviation for coefficient of haze is defined in the Abbreviation and Acronyms section as COH. In the text, both COH and CoH are used.
2. Page 2-29, Figure 2-9. The Y-axis needs a label.
3. Section 2.4.5, page 2-41-43, Figure 2-15. I wonder if the differences between measurement protocols for IMPROVE and EPA STN network influence the comparison of rural and urban $PM_{2.5}$ composition?
4. Section 2.5.1, Figures 2-19 and 2-20. The text on page 2-46 and 2-50 refers to 24-hr fine and coarse particle concentrations, but the figure captures refer to hourly observations.
5. Page 2-55, lines 29-30, Figure 2-28. There is something wrong with this figure.
6. Page 2-68. The abbreviation MSA in not defined in the Abbreviation and Acronyms section.

**References:**

Fujita, E., D. E. Campbell, W. P. Arnott and B. Zielinska (2005). Evaluations of Source Apportionment Methods for Determining Contributions of Diesel Exhaust to Ambient Carbonaceous Aerosols. Paper in preparation.

Zhu et al. (2002). Study of Ultrafine Particles near a Major Highway with Heavy-Duty Diesel Traffic. Atmospheric Environment, **36**, 4323-4335

---

**Revised – May 23, 2005**

**Comments on Chapters 5 and 7 of the PM Staff Paper (2nd Draft).**

**Barbara Zielinska**

1. I believe that the summary submitted by Drs. Sverre Vedal and Frank Speizer accurately summarize the discussion held during the CASAC meeting on April 6-7, 2005, in Durham, NC, regarding Chapter 5. I support their conclusions regarding the recommended level and statistical form of $PM_{2.5}$ standard. Specifically I favor the option of setting a 24-hour $PM_{2.5}$ standard at concentrations in the range of 35 to 30 µg/m$^3$, together with an annual standard in the range of 15 to 13 µg/m$^3$. I also support 98th percentile form, since it provides greater stability.

2. As far as $PM_{10-2.5}$ standard is concerned, I'm in favor of a more specific indicator for the toxic portion of coarse PM, if there is any. Obviously, more health data that focus specifically on this size fraction of ambient PM is needed. The indicator could be, for example, paved road dust, some specific industrial activities, or urban coarse PM. In my view such an indicator of coarse particle standard, if set at an appropriate level, would be effective in providing health protection for urban population in case such a protection

was needed, and also allow for conducting more health studies and collecting more ambient data. In my opinion, the option that favors specific $PM_{10-2.5}$ sampling sites is less optimal, since it seems more subjective.

3. I support Staff recommendations regarding secondary $PM_{2.5}$ standard to address urban visibility impairment. Specifically, I think that the 4 hr averaging time within daylight time period and 30 to 20 μg/m$^3$ range are the most appropriate. Slightly higher than 90$^{th}$ percentile form seems to be adequate as well, although "slightly higher" should be defined more specifically with a better justification.

# Dr. Jane Q. Koenig

Review of Staff paper
Feb/march 2005
Jane Koenig

I believe this document is generally well written (although longer than needed) and I am pleased to see that EPA staff agrees with many in the air pollution field that the current PM standard needs to be tightened.

Ch 3
I agree with the staff that health effects information in the CD are sufficient to use for quantitative purposes, that such information does not suggest a numerical threshold, and that it is appropriate to assume a linear concentration/effect relationship.

In the discussion of Evidence Based considerations on pp 5-25-5-27, I applaud the introduction of the Precautionary criterion. I have been surprised that EPA hasn't appeared to follow this criterion in the past, even though my interpretation of the CAA is that it is mandated!

I think Ch 5 would be improved by the addition of a succinct statement spelling out the recommendations for a new standard. For example, the statement in the Clean Air Report. Regarding a recommendation "the first option is to keep the existing annual standard of 15ug/m3 while lowering the 24 hour standard from 65 to between 25 and 35 ug/m3. The second alternative is to lower the annual standard to between 12 and 15 ug/m3, while keeping the daily standard at 35-40." The chapter could begin with this statement and then give the justification, and repeat the statement in the conclusion. This would allow individuals to scan the document--few will read it due to its length--and determine the staff recommendation.

I find Ch7 a concise statement of the need for a secondary standard and have no suggestions or additions.

# Dr. Petros Koutrakis

## CHAPTER 2: CHARACTERIZATION OF AMBIENT PM

Overall Chapter 2 reads well and presents many new interesting data. The authors should be commented for the thorough job they have done.

I have a few minor comments which are discussed below.

**2.2.1, page 2-2, Lines 5 - 20**

If I remember well, during our previous CASAC PM meeting there were some concerns about the presentation of the Whitby et al data. There was some discussion about including some new data on size distributions which I do not see here. This is not very critical but it would be nice to use some real and new data on particle distributions.

**Table 2-2, page 2-12**

This is mostly correct but it gives the impression that all ultrafine and coarse particles are removed after they travel short distances. Some qualifier should be used for this statement such as "a large fraction". Also big particles can travel long distances under some certain circumstances. For example the Sahara particles that reach Europe and South Eastern US are not necessarily the small size tail.

**2.3.1., Page 2-15, Line 2**

This is not 100% true. The coarse particles sample contains a large fraction of fine particles. [10% of the original air sample which can be up to 30-40% of the coarse particle mass].

**2.3.1., Page 2-15, Line 23**

I think it is a little early to offer this as a great alternative to the existing TEOM. There are not that many field tests out there to demonstrate its supremacy.

**2.3.1., Page 2-16, Line 10**

Unfortunately, the company that has licensed the CAMM sampler did a very poor job in constructing and marketing this sampler. I am afraid to say that at this point that CAMM does not seem to be a viable candidate, in spite its great promise. However, the beta-gauge approach seems to have earned some ground and may be an alternative if it is co-located with a filter-based method for calibration purposes.

**2.4.1., Page 2-27, Line 27**

There was a lot of concern about the potential effect of decreasing sulfates in the Northeastern US. Some papers had claimed that nitrate concentrations would increase because of the removal of sulfates. The results presented here do not support this hypothesis. This is a welcome finding.

**2.4.1., Page 2-27, Line 28**

I am not sure that the comparison between sites is the best way to examine heterogeneity among sites within a city. This approach overestimates heterogeneity. Do not forget that these sites were selected to represent different locations for which one would expect some differences.

From the analysis presented here one can suggest that compliance will depend to a great degree on the site selection and this is not completely true.

From the exposure assessment point of view comparing city mean concentrations to site specific ones is a better way to examine heterogeneity.

**2.4.4., Page 2-41, Line 8**

This also means that the particle health effects found using PM mass concentrations are not related to ultrafine particles!

**Figure 2-15, Page 2-42**

Fig 2-15 suggests that PM levels are higher in urban areas as the authors report. It also important to state that particle composition is similar in the urban and rural areas within the same geographical area.

**Figure 2-23, Page 2-53**

The authors present Figure 2-23 to say that for the same annual mean, there is a great variation among the 98th percentile daily values. I think there is a more important message here which is: for annual concentrations below the air quality standard all 98th percentile values are below the daily standard. Therefore, the annual standard controls the daily standard. This is a very important point that has to be kept in mind when deciding about the new daily standard.

Of course one could analyze the data in a different way: for instance present distributions of annual means for the same 98th percentile daily concentrations, but I think it is less meaningful.

**Table 2-5, Page 2-61**

Some concentration values reported here, e.g. 0-4 micrograms/m3, are below the IMPROVE detection limit. This is a serious issue, although it is difficult to go around it.

**2.7.2., Page 2-68, Line 18**

Not sure that I understand the previous sentence.

**2.7.2., Page 2-68, Line 23**

The real question we should address is the following: to what extent a given site within an urban area represents the average population exposure. As I mentioned above comparing pairs of the different monitoring sites is misleading because it exaggerates the differences. A more realistic, but not perfect, approach is to compare specific sites with the mean concentrations. The way the issue of heterogeneity is presented in the Staff Paper is not relevant to the exposure assessment question we want to address.

Also another important issue which is completely ignored is the measurement error. At best mass measurements are plus or minus 5% for fine particles and considerably higher for coarse particles. So even two co-located measurements can easily be different by several micrograms/m3. Therefore, a great deal of the differences in daily concentrations is associated with measurement error.

**2.7.2., Page 72, Line 6**

Actually exposure error may bias downwards the estimated risk factors.

**2.7.2., Page 72, Line 11**

Also outdoor measurements may not be good surrogates for some particle components especially the semi-volatile ones such as nitrates and organic carbon.

## CHAPTER 3: POLICY-RELEVENT ASSESSMENT OF HEALTH EFFECTS EVIDENCE

Chapter 3 is well written. Although it is difficult to summarize the health effects information presented by the Criteria Document, this chapter appropriately focused on the most important issues.

I only have a few general comments which are the following:

1) The authors should have given some more emphasis on toxicology;

2) Some more emphasis should have been given to source related health effects. There is limited but important information that suggests that combustion particles are more toxic than crustal particles and;

3) The chapter gives the impression that there is plenty of epidemiological evidence for coarse particles which is not the case.

**3.3.2., Page 3-23, Line 24**

The Staff should mention here that the number of coarse particle epidemiologic studies is extremely limited. The authors give the impression that there are plenty of coarse particle studies which is not true.

**3.4.2., Page 3-34, Line 23**

Considering the very low infiltration efficiency of coarse particles from the outdoor to the indoor environment, I am amazed that the limited coarse particle epi studies have found effects.

**3.4.3., Page 3-35, Line 18**

Yes it is not clear but I could assume that particle composition and home ventilation rates which may differ by city can be a reason for this heterogeneity.

**3.5.1., Page 3-39, Line 29**

The diabetes evidence is quite strong, so one could drop the "possibly".

**3.5.1., Page 3-40, Line 18**

Note that exposure studies have examined relationships between outdoor concentrations and personal exposures for children and other susceptible populations. These studies found no exposure differences between the susceptible and healthy groups.

**3.6.1., Page 3-46, Line 7**

As I have commented in Chapter 2, site-to-site comparisons do not make that much sense from the exposure assessment point of view. The issue here is how population exposures compare with concentrations observed at a given monitoring site. Therefore, concentrations at a given site should be compared to the city mean concentration which is better characterized by the mean of all sites.

**3.6.5.3, Page 3-60, Line 25**

I am not sure what greater associations mean. Lower p values or larger slopes?

**3.6.6, Page 3-66, Line 20**

I am not sure if I saw a discussion on harvesting. If this is correct, should harvesting be included in this section?

## Chapter 5

**STAFF CONCLUSIONS AND RECOMMENDATIONS ON PRIMARY PM NAAQS**

I have some minor comments discussed below. I will keep my major comments regarding the adequacy of the proposed standards until the CASAC meeting.

**5.3.1 - Page 8., Line 2**

This sentence needs some editing.

***5.3.2 - Page 17., Line 9***

I do not think this is a question of being appropriate as stated by this sentence. In my opinion both the criteria document and staff paper did not deal exhaustively with the source/composition issue. Although there is limited information on source specific effects, there is some indication that people living closer to traffic (e.g. the Netherlands and Southern California, and Six Cities study) are at higher risk. Also other combustion particles such as oil combustion particles have been found to be toxic.

Of course the evidence may not be sufficient and one may not be able to justify a source-specific standard at this time. However, ignoring the entire issue and conveniently focusing on mass is not the most rigorous approach.

**5.3.2 - Page 19., Line 7**

This is a hypothesis with limited evidence, which has been blown out of proportion in the Criteria Document.

**5.3.2. - Page 19., Line 17**

See my comment above regarding the health effects of specific sources. This statement is very strong and does not represent reality.

## 5.3.3. - Page 22., Line 13

I applaud this decision because it is both practical and realistic; meanwhile it recognizes the importance of short term exposures as an emerging issue that may have to be dealt with by the future PM standard reviews.

This approach should have been adapted in the case of source/components as well.

**5.3.3. - Page 23., Line 20**

This point is correct and should form the basis for making decisions about the new PM standard. In chapter 2, the staff paper shows the relationship between annual and 98th percentile 24-hour measurements. From this data it was clear that cities meeting the annual standard were also compliant with the daily one. Interestingly, the authors of chapter 2 did not make this very important point.

**5.3.5.1. - Page 37., Line 19**

Is it possible to give a few examples of areas with high peak-to-mean ratios?

**5.3.6.1. - Page 51., Line 20**

It is not very clear to me what the Staff paper is trying to do here. This is like opening the Pandora's box. It seems to me that if more than one site is used within a metropolitan area, then the average concentration should be used for calculating risks.

Furthermore, the sites should be selected to represent a large spectrum of exposures within the area. I do not see the purpose to select sites which are extremely correlated and they are within 10%.

Post-CASAC meeting comments by Petros Koutrakis

Fine Particles:

I was very pleased to see that there was a consensus among the panel members regarding the fine particle standard. I think this is a reflection of the research efforts made since 1997 to address key scientific questions concerning the health effects of fine particles. My recommendations include reducing the daily standard from 65 to 30-35 micrograms per cubic meter, while maintaining the 98th percentile rule for the previous 3 year average. For the annual standard I would suggest a slight reduction from 15 to 14 micrograms per cubic meter. This would force cities with relatively infrequent episodes but with frequent single daily concentrations ranging from 20 to 30-40 micrograms per cubic meter to reduce concentrations.

Coarse Particles:

I feel we did not make much progress with the coarse particle standard. This was mostly due to the lack of data. I have always, in principle, supported having separate fine and coarse particle standards, rather than standards for fine (PM2.5) and PM10. This is because there is a contradiction when the contents of the measurement for one standard (PM2.5) is entirely contained within the amount for another standard (PM10).

Furthermore, setting up a coarse particle standard poses a great challenge for two reasons: First, one would not expect serious adverse health effects associated with exposures to natural soil dust and; second, many components of coarse particles such as natural dust, sea salt and pollen, although they regulated they can be controlled. Thus it would not make sense to regulate soil dust, seas spray, or pollen if we can do nothing about them.

In contrast road dust (from paved roads) and industrial coarse particles can pose a serious threat to public health. Although limited information exists about the toxicity of road dust, it is likely that these particles are extremely potent as they contain many components which we know are toxic (Batalha et al 2002). These include latex particles from the tires, combustion particles from exhaust emissions, metals from tires and brakes, and lubricant from the engine. In addition road dust is rich in secondary particles such as nitrates and sulfates which can serve as nutrients to many micro-organisms which are associated with road dust. Thus the observed associations between coarse particles and health effects found so far by limited epidemiological studies is best explained as due to the road dust component of these particles. This is because individuals living in large cities are more likely to be exposed to road dust particles than coarse particles originating from unpaved areas inside or outside the cities. We should not forget that under normal atmospheric conditions coarse particles stay airborne for short periods (minutes to hours). Long range transport of coarse particles happens infrequently.

I propose that the we set up a coarse particle standard which includes only road dust and particles from specific industrial activities (specific indicators). Today we have the tools to distinguish road dust from natural dust. For instance, electron microscopic analysis can be easily done for individual coarse particles (while not as easy and precise for fine particles) and these results can be used for source apportionment studies.

If road dust and other industrial coarse particles be used as the standard indicator, then we can develop a strict standard which can be used to protect public health. Conversely, if the indicator encompasses soil, sea salt. and pollen then it will be hard to set a suitably low standard because it will be hard to achieve compliance. Finally, and most importantly, we can control exposures to road dust in a cost effective way. We know that street sweeping is effective and has helped cities such as Santiago, Chile to reduce PM10 levels.

I hope my comments will help to make a scientifically sound and practical decision about the coarse particle standard. I will be happy to provide more information if needed.

Reference:

Batalha, J. R. F., Saldiva, P. H. N., Clarke, R. W., Coull, B. A., Stearns, R. C., Lawrence, J., Krishna Murthy, G. G., Koutrakis, P., Godleski, J. Concentrated Ambient Air Particles Induce Vasoconstriction of Small Pulmonary Arteries in Rats, Environmental Health Perspectives, 110(12): 1191-1197 (2002).

# Dr. Allan Legge

Revised Comments: April 13, 2005

Review of the EPA - OAQPS ‘Second Draft’ Staff Paper on Particulate Matter entitled “Review of the National Ambient Air Quality Standards for Particulate Matter”, January, 2005. Comments by Allan H. Legge

I. Chapter 6: Policy - Relevant Assessment of PM - Related Welfare Effects
Note: Emphasis on Section 6.3 Effects on Vegetation and Ecosystems

II. Chapter 7: Staff Conclusions and Recommendations on Secondary PM NAAQS

I. Comments on Chapter 6
Overall, this Chapter is well done. Staff is to be commended for a well written and concise reflection of the key science as presented in the final PM CD as it pertains to effects on vegetation and ecosystems. The ecological risk assessment is reasonable given the required ‘criteria pollutant’ approach. That being said, the ‘criteria pollutant’ approach in this case (i.e., PM) has serious short comings when it comes to ensuring environmental protection of vegetation and ecosystems in the US. This is illustrated in the following discussion.

There is scientific evidence presented in the SP and the PM CD that indicates that forest ecosystems at a number of locations in the US “are now showing severe symptoms of nitrogen saturation” ( SP page 6-37, lines 13-18). The SP makes the point that this is the result of chronic long-term additions of reactive nitrogen (Nr) species that have been accumulating over time. The SP also makes the point that the forest ecosystem deterioration issue is broader and more complex than just Nr accumulation. The SP notes that “The most significant PM-related ecosystem-level effects result from long-term cumulative deposition of a given chemical species (e.g., nitrate) or mix (e.g., acidic deposition) that exceeds the natural buffering or storage capacity of the ecosystem and/or affects the nutrient status of the ecosystem” ( SP pages 6-31, line 31 and 6-32, lines 1-3 ). A key point implied here and elaborated later in the SP text is that PM deposition is only partially responsible for the observed ecosystem-level effects and that the extent of the role of PM deposition in these ecosystem-level effects needs to be determined. While this has scientific merit, the question must be asked as to whether knowing the role of PM alone will improve the protection of vegetation and ecosystems in the US? The answer to this question is critical because forest ecosystems are responding to the cumulative total load which has resulted from the chronic long-term deposition of both PM as well as gases and not to PM alone.

While Staff has done a commendable job within the context of the ‘criteria pollutant’ approach, it is strongly recommended that in the future that the ‘Agency’ give serious consideration to a philosophical shift from the ‘criteria pollutant’ approach to the European approach of ‘critical loads’ when it comes to ensuring protection of vegetation and ecosystems in the US.
I. Specific Comments: Chapter 6.
1. Page 61-1, lines 15-18.
There is an omission in this paragraph with respect to addressing organic compounds which were covered in the PM-CD. While it recognized that these are not criteria pollutants controlled by the NAAQSs under section 109 of the Clean Air Act (CAA), it would be beneficial

at least to indicate that there is some control of these substances under Section 112, Hazardous Air Pollutants (U.S. Code, 191) as indicated in the PM-CD. It is important to remember that some of these organic compounds occur in the particle phase.

2. Page 6-28, line 28-29.

How does reference 'SAB, 2002' differ from EPA (2002)? In the reference section these two references appear to be the same ( see page 6-77, lines 41-43 and page 6-80, lines 30-33).

3. Page 6-41, lines 15-16.

Reference is made to 'cloud deposition'. Shouldn't this read 'occult deposition'?

4. Page 6-43, line 26.

Reference is made to '(Smith, 1990a)'. The reference list on page 6-80, line 41 refers only to 'Smith, W. H. (1990)'.

5. Page 6-46, line 1.

Spelling. Should read '- - the nutrient cycling model, NuCM, to'

6. Page 6-50, lines 3-4.

The title refers to 'Indirect Vegetation and Ecosystem Effects' which is quite broad. The text ,however, focuses on 'radiation and climate conditions'. The title needs to be changed.

7. Page 6-54, lines 24-26 and page 6-55, lines 3-4.

There is an inconsistency between these two statements in the text. It is first noted that "Data from these deposition networks demonstrate that N and S compounds are being deposited onto soils and aquatic ecosystems in sufficient amounts to impact ecosystems at local, regional and national scales." It is them noted that "Unfortunately, at this time there is only limited long-term ecosystem response monitoring taking place at the national level."

8. Page 6-78, lines 35-38.

Citation needs to be corrected. Should read as follows:

Hornung, M.; Langan, S.J. (1999) Nitrogen deposition: sources - - -. In: Langan, S.J., ed. - - - - . Dordrecht, The Netherlands: - - -; pp. 1-13. [Environmental Pollution, Volume 3].

II. Comments on Chapter 7.

Staff notes "that further reductions in ambient PM would likely contribute to long-term recovery and to the prevention of further degradation of sensitive ecosystems and vegetation" ( SP page 7-20, lines 10-11) and that "national standards alone may not be an appropriate means to protect against adverse impacts of ambient PM on ecosystems and vegetation in all parts of the country" (SP page 7-20, lines 16-18). This is true. The problem is that the current 'criteria pollutant' approach does not allow the 'Agency' to adequately address the matter of total cumulative load of all of the pollutant stressors to which the ecosystems are responding. Reducing PM will help but will not be effective in a timely manner to help increase the protection of ecosystems currently showing adverse responses from the results of cumulative deposition. As noted under the comments for Chapter 6, a philosophical shift in the 'Agency' from the 'criteria pollutant' approach to a 'critical loads' approach is recommended. This more realistic holistic approach is far more likely to improve the environmental protection of vegetation and ecosystems than emphasizing PM alone. While Staff has recognized the 'critical loads' concept one has the sense that it is not enthusiastically embraced.

II. Specific Comments: Chapter 7.

1. Section 7.5 Summary of Key Uncertainties and Research Recommendations Related to Standard Setting.

i) Page 7-23, lines 22-27 and page 7-24, lines 1-2, Recommendation (2).

It is a positive step that 'Staff' has recognized that PM plays a role in cumulative long-term environmental impacts and that PM's contribution to long-term environmental impacts is not known at this time. It is questionable, however, that research simply focused on determining the percentage of the total deposition contributed by PM would be useful in and of itself. Ecosystems respond to the cumulative deposition of both PM as well as gases in both wet and dry form. There is no question that "better tools and monitoring methods should be developed." That being said, one needs to know what air pollutants of whatever form that vegetation and ecosystems are exposed to and how they respond over the short as well as long-term (i.e. key biological and chemical indicators need to be adequately characterized and monitored). Further, it is also extremely important that the monitoring and research takes place where the ecosystems are located. While it is recognized that urban environments are important, it is equally important for vegetation and ecosystems in rural and more remote environments to be recognized as important.

ii) Page 7-24, lines 3-11, Recommendation (3).

This recommendation by 'Staff' needs to be rethought. The thinking is too short sighted. While it may be true that there is likely "immense variability in sensitivity to PM deposition across U. S. ecosystems", understanding that variability only in the context of PM will not be that helpful towards meeting the goal of improved environmental protection. It would be more important and more helpful to characterize and quantify the range of the responses of ecosystems and ecosystem functions to cumulative loading from all forms of air pollutants not just PM. This more holistic approach would be more likely to yield data capable of forming a sound scientific basis for predictive models. The 'critical loads' concept should be considered sooner rather than later in the U.S.

# Dr. Paul J. Lioy

**Review of the EPA-OAQPS Staff Paper for Particulate Matter**
**By Dr. Paul J. Lioy,**
**Professor, and Deputy Director For Government Relations**
**The Environmental and Occupational Health Sciences Institute**
**Robert Wood Johnson Medical School –UMDNJ**

## Chapter 2

Sections 2.1 through 2.3, provide an accurate summary of the information found in the PM Criteria Document. A very good job. My only concern is the lack of information that describes potential qualitative and quantitative differences between the annual and daily composition for each of the four seasons.

### Section 2.4

I find the information accurate and useful; however, there is one significant issue. The document does not adequately define the quantitative range of error (as opposed to the variability) for the FRM, either for daily samples or as carried over to construct an annual average. The lack of error information for both the daily and annual average leads to a problem. The $PM_{2.5}$ analysis summaries take as a given that the concentrations reported are accurate to three decimal places. Thus, when constructing the regional trends for $PM_{2.5}$ (Figure 2-8) you are assuming that an annual average mass difference of 0.2 $ug/m^3$ (or 0.6 $ug/m^3$) is real. Please justify the scientific validity the staff's interpretation of the results, especially for the Northeast. Within the limits of analytical error, the difference may be zero.

Table 2-3 needs information for Northeastern cities of New York, and Boston. I do not necessarily believe that Philadelphia is representative of the areas north along I-95. Figures 2-6, and 2-7 suggest that such information is available.

Figure 2-16 is excellent, and should provides a basis for discussions about annual source contributions.

Figure 2-19, and 2-20 do the counts mean "hourly"? (are these continuous monitoring data?), or does it mean "# of 24 hour samples"?

Page 2-71. For the current forms of the standard ambient monitors are a useful surrogate for exposure. However, as stated in my review of the criteria document, the detection of cardiac health effects in populations at risk, may require consideration of a shorter term standard, e.g. 1 hr or 8 hr, in future reviews. If these information continue to be coherent in future field and laboratory studies, a PM central monitor may no longer be adequate to address "hot spots" of emissions and human contact with high short term exposures to PM. Thus, I suggest providing some room for considering exposure based forms of the standard in the future evaluations of the PM standards.

**Chapter 3**

Generally a good summation of the results and current evidence. Further, there is a clear discussion of the GAM. The issue of thresholds difficult and not resolvable given the current state of knowledge. The Staff has provided a reasonable analysis of the situation, and how to approach the problem.

**Chapter 4**

The staff provided a reasonably clear foundation for the risk assessment. There are a few areas of concern that need to be re-examined with respect to the assumptions and issues associated with the short term standard relative to the annual standard.

**1**. Figures 4-8a and 4-8.b are misleading. They project an image that as the pollution goes up to around 25 to 30 ug/m3 that the deaths go up. On the surface this is true, and, based upon this observation, the greatest concern for risk reduction would be focused on reducing the *peak number* of non-accidental deaths. However, a more accurate representation of the data for Detroit would be to divide the non-accidental deaths by the # of days in which the ambient air concentration is at a specified value. The result would be in deaths/day, which is probably a more representative value of pollution impact. A graph of the excess daily death rate vs. concentration would show that the excess daily death rate in Detroit slowing increases from approximately 0.15/day at 10 ug/m$^3$ to a peak of approximately 2.3/day between 45 and 50 ug/m$^3$. (note, either of these values could be a bit higher or lower as I only could interpolate from the graphs) There is little data above this concentration range; therefore, further projections would very uncertain. The values for excess death rate suggest that at higher concentrations particles with either higher toxicity or higher population exposures (more time outside or near the source) could be increase the daily death rate in Detroit. The result for Detroit may be anomalous, or representative of the other areas, but such information was not provided in the staff paper. At a minimum this observation needs to be more thoroughly explored by the staff. If this type of relationship is realistic, the results suggest that the strategies needed to bring down the annual average would not be the same as those needed to deal with the peak concentrations, and the effects caused by the higher levels.

As many have stated throughout this process, the science in the CD indicates that PM is a complex mixture, (see Table 2.2 in the staff paper). Further, it as been shown in many studies that peak concentrations can be driven by daily and seasonal specific events, and chemistry. As a result, the concentration and source patterns on a day of high levels of $PM_{2.5}$ may not be he same as what drives the $PM_{2.5}$ levels on most days. A result reported by multiple investigators over the past 30 years.

**2.** A role back model may be appropriate for designing strategies to reduce risk and achieve the long term (annual) standard, since the idea would be to role back the emissions from all primary sources, However, I am not sure that it is appropriate for the $PM_{2.5}$ peaks. The results associated with figures 4-10 and beyond suggest such a conclusion, since the annual average is not affected by decreases in the peak until the values are decreased to around 25 ug/$m^3$, and not at concentrations above 35 to 40 ug/$m^3$.

If we were dealing with the primary pollutant CO, then a role back model would be valuable for both a long term (if we needed one) and a short term standard. In that case, the reductions made to mean the mean would in fact be correlated with the reduction of the levels at the 98%tile+. However, the nature of the extremes for $PM_{2.5}$ are not just the sources of primary $PM_{2.5}$, but also secondary formation processes which are governed by homogeneous and heterogeneous chemistry, and meteorology: A point clearly discussed in the Criteria document and mentioned throughout Chapter 2 of the Staff paper. Thus, the accumulation rates could be different for the 90%tile+ of measured concentrations and the pollutants accumulated could be different. In the end, the exposures to the populations at risk may also be much different both in terms of quality and quantity of material that people contact on a given days. Without analyses similar to Figures 4-8a and b and my suggested 4-8c for other cities, I think the risk issues related to the short term standard have not been adequately addressed in the staff paper.

The staff should start with Figures 4.8a and 4.8b, modify to a figure that examines the mortality /day on the Y axis, and complete he same evaluation for other urban areas. Detroit could be anomalous or similar to other urban areas within the US.

**3.** *After the Staff reviews any new results, it may find that Detroit is truly an anomaly. If that is the case, then the role back approach would be adequate for reducing risk. If, however, Detroit is not an anomaly, the role back method could still be used to set both the annual and the 24 hour standard, a point suggested in sentences 3-5 on page 4-59. The Staff, however, would have to provide qualifiers since the 24 hr standard may or may not be protective on the highest days.* This would be due to focus being on the average composition of the particulate mass, and not secondary or other particles that will contribute to the mass on the higher exposure days.

---

*Paul J. Lioy, Ph. D.* *Date: 4/11/05*
*Deputy Director and Professor*
*The Environmental and Occupational Health Sciences Institute*
*UMDNJ-RWJMS*
*170 Frelinghuysen Road*
*Piscataway, NJ 08854*

**Comments on Chapter 5 of the PM Staff Paper and the Discussions at the CASAC Meeting held on April 6-7, 2005, Durham, NC.**

1. The results derived from the $PM_{2.5}$ risk assessment were presented in Table 5.2. Although interpretable, they were not easy to understand, especially when I attempted to make simple comparisons among the different sets of long term and short term health risk assessments. The Staff should build upon the discussion at the meeting and construct three dimensional plots consistent with those presented for Detroit by Dr. Fred Miller.

The plots should be completed for each city used for the Staff paper risk assessment, and the figures should be discussed with respect to the influence of the various design values on decreasing long term and short term mortality. From the work presented in this chapter, and chapters 3 and 4, it appears that a primary goal should be toward reducing the level of the primary 24 hour $PM_{2.5\ NAAQS}$. Based upon the results in Table 5.2 the value I believe the value should be reduced from 65ug/$m^3$ to somewhere in the range between 30 and 35 ug/$m^3$; however, the "3 dimensional" plots could refine my current opinion.

2. Based upon my review of the risk assessment for long term health effects, the annual $PM_{2.5}$ NAAQS could remain at, 15 ug/$m^3$, or the value could be reduced to 13 or 14 ug/$m^3$. I will withhold final thoughts on the level for annual standard until after I see the "3 dimensional" figures for each of the other cities. However, the current tables do provide the needed information, and at the present time I would be comfortable with an annual value that does not exceed 15 ug/$m^3$. I am interested in seeing whether or not the influence of the daily average is washed out in cities that have lower peak concentrations or fewer episodes (a point made by Dr. P. Koutrakis during the meeting).

3. The 24 hr standard should use the $98^{th}$ %tile form as a robust bench mark, and the value should be derived from the highest monitor in the area of concern.

4. Based upon the results summarized in Chapter 3 and 4 of the staff paper, the coarse (thoracic) particle risk assessment had much weaker information to draw upon for risk characterizations. I commend the Staff for their attempt at completing such an assessment at this time. I agree with the Staff findings on page 5-65 that: 1. support a 24 hour $PM_{10-2.5}$ NAAQS, and 2. cast doubt about the evidence for a $PM_{10-2.5}$ long term NAAQS at this time.

5. A 24 hr $PM_{10-2.5}$ NAAQS should take the $98^{th}$ % tile form, and the value could fall within the range between 50 and 75 ug/$m^3$.

6. My major concern is that the current proposals for a $PM_{10-2.5}$ 24 hr NAAQS do not provide any compensation or relief for the differences in composition associated with urban thoracic particles and rural thoracic particles. Crustal material associated with rural dust or dirt has been well characterized for well over thirty years. There are volumes of information on the elemental and ionic composition of dirt or dust. Further, there have been few, if any, studies indicating the potential for severe health outcomes after exposure to dust particles. What is only beginning to be understood is the nature of the composition of urban street dust. These materials have thoracic and extra-thoracic size fractions and include materials typical of rural dust. However, urban dusts or re-suspendable dirt will also carry adsorbed materials obtained from general urban, and human activities. Examples would include oils and greases, rubber, wastes from home activities, and condensates from motor vehicle emissions. Therefore, my recommendation is to establish a $PM_{10-2.5}$ standard that is defined as either a national urban thoracic particle, or resuspendable road dust thoracic particle NAAQS, a $UPM_{10-2.5}$ or $RRDPM_{10-2.5}$, respectively. Either can address the concerns raised by the epidemiological studies in

the CD, and the risk analyses described in the Staff Paper. In addition, either will also provide an NAAQS that would focus on the potential for population exposure to $PM_{10-2.5}$ rather than the general air quality. Therefore, a $UPM_{10-2.5}$ or a $RRDPM_{10-2.5}$ would be effective in protecting public health.

7. Although either a $UPM_{10-2.5}$ or $RRDPM_{10-2.5}$ would be acceptable as an NAAQS, an alternative approach that could satisfy many of my concerns would be based upon the method for implementation of an $PM_{10-2.5}$ NAAQS. This could be accomplished by defining the air quality measurement siting criteria in term of the locations where one would anticipate significant population exposure. The network could be defined in such a way that it would only include sites within a defined distance and height from urban roadways that are near population centers or hubs. This approach would exclude farms, and other rural locations as monitoring sites. Industrial areas that are not within a specified distance of a town or development, based upon dispersion modeling simulations, should also be excluded from having PM10-2.5 sampling sites. An implementation approach for a 24 hr $PM_{10-2.5}$ NAAQS is more complicated and would require many rules and exceptions. Therefore, I would rather have the EPA consider establishing a 24 hr national $UPM_{10-2.5}$ or $RRDPM_{10-2.5}$ standard at this time. As more health and exposure data becomes available the form of the $PM_{10-2.5}$ NAAQS could change.

*Concluding Comment:*

Chapter 5 is going to require some major revisions as well as minor revisions. Because of the former there will need to be more discussion among the members of the committee. Thus, in my view, Chapter 5 is a work in progress requiring another review by CASAC to achieve consensus; especially as related to the 24 hr Thoracic Particle Standard. I truly want the CASAC recommendations to the administrator to be based upon the most recent information since we have raised a number of important scientific issues for $PM_{10-2.5.}$

# Dr. Morton Lippmann

Chapter 4

General Comment

Chapter 4 provides a straightforward description of the rationale, procedures used, and the results obtained in the Staff's characterization of the human health risks resulting from population based exposures to PM of outdoor origin. The methods have evolved with significant input from this and prior CASAC and Council panels, and represent the state-of-the-art work by a highly professional team of contractor and EPA professionals.

Chapter 5

General Comments

In my view, this chapter provides a fair and balanced presentation of the issues, as well as of the science as it was presented in the PM CD and the Abt Associates risk assessment.

The major issues that I believe need discussion by the CASAC PM panel and further elaboration in the text of Chapter 5 are:

1) Coarse thoracic concentrations for the epidemiological study in the Detroit metro area were indexed by measurement data in Windsor, Ontario (which is close to downtown Detroit). On page 5-68, line 5-8, it is argued that the population exposures should be indexed upward because more recent data indicate that downtown Detroit concentrations are about twice those of Windsor (with Windsor having concentrations similar to those in Detroit suburbs). I agree that such differentials in concentration in subsequent years should be considered, and are appropriate to the discussion of "margin of safety". It should also be noted that the morbidity effects in that study were those for the Detroit metro area containing 2.1 million residents (p. 4-21), and that more of them reside in suburbs and intermediate density areas than in the center city. Thus, if an adjustment in the coefficient of response is warranted, it should be something between none and a reduction by half.

2) The issue of retaining an annual NAAQS for coarse thoracic particles is raised on page 5-74, lines 22 - 24, but there is no discussion of what specific levels of an annual average NAAQS would provide what additional degree of public health protection, as was done for the PM2.5 NAAQS.

3) Item (6) on page 5-78 should also include a reiteration of the recommendation, on page 5-64, lines 5 & 6, for continuous monitoring.

SPECIFIC COMMENTS:

| Page | Line(s) | Comments |
|---|---|---|
| 5-59 | 20 | Change "these" to "the". |
| 5-73 | 1 | Change "a" to "little". |

[Note: An expanded listing of comments on Chapter 3 from Dr. Lippmann is imbedded in Dr. Fred Miller's comments found on pp. B-11 through B-16 above.]

# Dr. Joe Mauderly

Draft Comments on Second Draft of OAQPS Staff Paper on Particulate Matter

Joe L. Mauderly 3/26/05

**General comments**

Much weight is given to multi-city studies, and that's reasonable given the current state of our knowledge. However, I don't think that sufficient caveats are given concerning the heterogeneity among cities in the magnitude of PM concentration-response relationships. Perhaps that's the best we can do at this point and indeed, it is likely that heterogeneity of exposure within cities and the effects of short-term spikes (which is a related issue) will eventually be found to be more important variables than differences among cities. I'd like to see that stated more clearly.

**Specific Comments**

Chapter 3:

P 3-6, L 21: Recent studies also demonstrate that PM can move into brain by pathways other than systemic circulation.

P 3-8, L 12-16: It should be acknowledged that most of the data referred to here resulted from non-physiological doses.

P 3-51, L 15-17: It is not clear why "colinearity" could not also occur if the concentrations of different pollutants were also raised or lowered simultaneously due to meteorological conditions.

P 3-52, L 16-19: Here we have authors' speculations portrayed as "information". The example given is a reference to a quote in the CD by Pope et al. that the association between $PM_{2.5}$ and mortality was less plausible than an association between $SO_2$ and mortality. The staff paper offers it as a "conclusion". The quote in the CD is a bit different: "the absence of a plausible toxicological mechanism --- further suggests –". The fact is, we didn't have much of a plausible mechanism for $PM_{2.5}$ either a few years ago. One might suspect that if we threw an equivalent amount of time, money, and incentives at $SO_2$, we'd generate some "plausibility" for that pollutant as well. Regardless, Staff should be careful not to portray speculation as "information".

P 3-56, L 4: "Suggested" is misspelled.

P 3-61: Here is an example of glossing over likely inter-city heterogeneity. It is stated that when data from many cities are combined, there is no apparent pattern with season. Indeed, why should there be? Pollution composition and levels are affected differently by season in different locations. Seasonal variation is not a variable that one would expect to attack best by combining data from cities in different regions of the country. The final statement may be true – that present evidence does not support a quantitative assessment of effects of season. It is not clear that our best effort has been directed at this issue.

Chapter 4

P 4-2, L 6: Is the Agency developing elicitation approaches for PM, or as a general tool? As I recall the Agency felt that it had the ability to do this in the last Ozone review.

P 4-3, L 7-9: This statement seems to conflict with the quantitative estimates of differences in health outcome among the different forms of the standard. I guess is depends on what one means by "precise".

P 4-28, 19-22: This is confusing. If the studies included were not required to have reported a statistically significant linkage to concentration, then how do we have the confidence to use them to estimate a concentration-response function?

Chapter 5

P 5-9 &10: The logic of the connections being drawn between short-term exposures and health effects and long-term average concentrations is not clear. Of course one would expect that short-term peak exposures could cause effects in areas meeting the annual standard. That doesn't mean that exposures at the level of the annual standard caused significant effects. 24 hr (or potentially shorter) standards are intended to take care of this – how could one set an annual standard to ensure that no harmful short-term exposures would occur? If that were the case, one wouldn't need a 24 hr standard at all. There must be a rationale here, but it's not clear.

P 5-13, 14, & 15 (including Table 5-1): The discussion of numbers of deaths and "incidence" is confusing. This is an important point to clarify, because it impacts heavily on subsequent statements. The table lists numbers of deaths, which is a useful metric. The text following the table discusses numbers of deaths per 100,000 population – a different, but also useful, metric. The text also talks about percentage incidences – still another useful way to look at the data. What is the definition of "incidence"? The table should list both numbers of deaths and some population-normalized mortality parameter, to allow easier comparisons among cities.

P 5-76, L 15-26: This paragraph appropriately notes "spatial variability" as a research need. I don't think the text sufficiently identifies the need to know more about both spatial and temporal variability. The issue of inter-city variability should be clearly emphasized here, as well as variability within a city. Inter-city (regional) differences in seasonal variability is also a research need.

Chapter 7

P 7-17, L 4: Staff recommends a secondary standard with a less than 24 hr averaging time – which, of course, would require such monitoring. Although Staff is not recommending a primary standard averaged less than 24 hrs, the impact of such a standard on controlling and assessing human health effects should be discussed. The health sections should also note this proposal and its potential ramifications.

# Dr. Roger O. McClellan

**Review Comments**

**on**

**"Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information (Second Draft, January 2005)"**

by

Roger O. McClellan, DVM, DABT, DABVT, FATS
Advisor, Toxicology and Human Health Risk Analysis
13701 Quaking Aspen Place NE
Albuquerque, NM 87111

Revised: May 20, 2005

## Summary

These revised personal comments (May 20, 2005) are based on my review of "Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information (Second Draft, January 2005)" and associated supporting documents and review of the comments provided by 28 interested parties. I did not participate in the CASAC meeting on April 6-7, 2005 because I was undergoing major surgery on April 6, 2005. As of this date a transcript of the meeting is not yet available for review.

I found the written comments of "interested parties," especially those that focused on the scientific basis for setting the NAAQS for PM, to be well written and useful in evaluating the adequacy of the EPA Staff Paper and supporting documentation. I was disappointed to learn that the "interested parties" were only given 3 minutes each to present comments at the CASAC meeting. The value of the comments from "interested parties" in informing the CASAC deliberations certainly warrants giving those parties more time to summarize their views. My views on the scientific adequacy of the Staff Paper have been influenced by the points made by interested parties, points that were either ignored or down-played in the EPA Staff Paper.

In my professional judgment, the Staff Paper, in its present form, does not represent a balanced and scientifically adequate synthesis and interpretation of the scientific evidence relevant to setting/revising the NAAQS for PM. I urge the CASAC to request the Agency to revise the Staff Paper and return it to the CASAC for re-evaluation.

The present document does not provide a balanced summary of the information available on $PM_{2.5}$ exposure-response relationships. The document gives excessive weight to highly uncertain complex calculations based on multiple extrapolations to estimate excess morbidity and mortality for exposure to $PM_{2.5}$ across the United States. The manner of presentation is such

that uncertainties are masked and with each step in the presentation a more certain and precise estimate is provided of the consequences of exposure at relatively recent ambient exposure concentrations. The approach taken leads to calculations of excess risk even for cities previously demonstrated to not have statistically significant excess risks. The failure of the document to deal with such obvious contradictions will undoubtedly lead to an erosion of public confidence in even the current NAAQS for $PM_{2.5}$, let alone any revised standard.

Based on the information in the Staff Paper and consideration of the comments of interested particles, I believe it would be appropriate for the Staff Paper to recommend: (a) the continued use of a $PM_{2.5}$ indicator; (b) with an annual averaging time standard as high as 15 $\mu g/m^3$; and (c) with a 24-hour averaging time standard as high as 50 $\mu g/m^3$ with a $98^{th}$ percentile statistical form. The choice of lower numerical values by the Administrator would represent a decision to increase the margin of safety.

The Staff Paper substantially overstates the scientific information available for evaluating $PM_{10-2.5}$ exposure-response relationships across the United States. In the absence of adequate data the Staff Paper attempts to cloak consideration of a $PM_{10-2.5}$ indicator by noting the evidence is less than is available for $PM_{2.5}$. The evidence for a $PM_{10-2.5}$ indicator must stand on its own. The Paper should acknowledge that this database is extremely weak. The Staff Paper approach pleading for a $PM_{10-2.5}$ indicator can be viewed as an attempt to use $PM_{10-2.5}$ as a "place-holder" indicator for a coarse PM standard. I initially thought that approach might be appropriate. However, on reflection I have concluded that in the absence of a scientific basis specifically for a $PM_{10-2.5}$ indicator, the choice of such an indicator would be arbitrary and capricious. The arbitrary selection of a $PM_{10-2.5}$ indicator as a National Ambient Air Quality Standard would be especially inappropriate for areas where $PM_{10-2.5}$ consists primarily of crustal material. I am supportive of considering a $PM_{10-2.5}$ urban indicator.

In closing, I express concern as to the current efforts by EPA staff to change the modus operandi of the CASAC. In my professional opinion, based on past service as Chairman of the CASAC and as a member of numerous CASAC panels, the past modus operandi of CASAC has been consistent with the language of the Clean Air Act and the Federal Advisory Committee Act. Moreover, the activities of the CASAC have clearly had a positive impact on the setting of science-based NAAQSs for criteria pollutants and, thus, the implementation of the Clean Air Act. The use of "closure letters" to document when the CASAC has concluded that the Criteria Documents and Staff Papers are scientifically sufficient for regulatory decisions has been an important part of the CASAC modus operandi. In the absence of compelling arguments to the contrary I see no basis for a change in the CASAC modus operandi. Why tamper with a successful approach!

## INTRODUCTION

This document serves as a record of my comments on the EPA Staff Paper entitled "Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information (Second Draft, January 2005)." I start with a brief discussion of the CASAC modus operandi. I then proceed to provide both general and specific comments on Chapters 3, 4 and 5 of the draft document and then offer specific comments on the four elements (indicator, averaging time, statistical form and numerical level) for potential primary NAAQS for $PM_{2.5}$ and $PM_{10-2.5}$.

## **Past is Prologue**

This section of my report is an expression of concern with changes being made, out of public view, in the modus operandi of the Clean Air Scientific Advisory Committee (CASAC). My concerns are based to a large extent on my historical involvement, involvement that began soon after EPA was created, with the review of the scientific basis for new and revised National Ambient Air Quality Standards for Criteria Pollutants. The past modus operandi of the CASAC has been highly successful in ensuring independent, critical review of the scientific basis for setting and revising the NAAQS for criteria pollutants. The basis for the changes that are being made in the modus operandi of CASAC have not been publicly articulated. However, the changes in the modus operandi of the CASAC would appear to relegate CASAC to an advisory status that is not consistent with its independent status under the Clean Air Act. The CASAC is not merely another advisory committee, it is a very special independent scientific committee with extraordinary responsibilities as part of a national program to ensure air quality.

Prior to the creation of the EPA in 1970, responsibility for administering the Clean Air Act and earlier air quality statutes was vested with the National Air Pollution Control Administration which had an independent Clean Air Advisory Committee. With creation of the EPA, a number of “inherited” advisory committees, including the Clean Air Committee, were abandoned. In their place the EPA created a Science Advisory Board which had a number of discipline- oriented committees; Health, Engineering, Ecology, etc. I served as a member of the original EPA Science Advisory Board Executive Committee by virtue of my chairing the Board’s only original issue-oriented standing committee, the Environmental Radiation Exposure Advisory Committee.

In the early 1970s, air quality issues were handled by the SAB on an *ad* hoc basis. An example was the handling of a review of lead as an air pollutant. Lead had not been included as one of the original criteria air pollutants. The National Resources Defense Council (NRDC) took legal action to have lead listed as a criteria pollutant and ultimately prevailed in the Appeals Court (NRDC vs Train). Thus, EPA was required to prepare a criteria document on airborne lead and the decision was made to have the document subjected to external peer review. In the absence of a formal clean air scientific review committee, I was asked to chair an *ad hoc* committee to review the lead criteria document. The *ad hoc* committee met in public sessions, reviewed the report, received input from the EPA staff and heard public comments. Our initial conclusion was that the original criteria document on airborne lead was inadequate and needed to be substantially revised. EPA was operating under a court-ordered deadline to issue a NAAQS for lead, a deadline that did not allow time for revision of the lead criteria documents. However, the Agency and interested parties persuaded the Court to extend the deadline to allow preparation of a scientifically adequate document rather than merely meeting an arbitrary “date certain” deadline. The *ad hoc* committee reviewed subsequent revisions of the document. Ultimately, a document was created that the *ad hoc* committee approved as being a scientifically adequate basis for setting the National Ambient Air Quality Standards for lead and issued a “closure letter” to the EPA Administrator. The key points being made are that the scientific basis for the NAAQS for lead was reviewed and a decision was made by the *ad hoc* committee as to when the documentation was scientifically adequate for regulatory decision making.

In my opinion, the approach taken by the *ad hoc* committee dealing with lead as a criteria pollutant influenced the decision of the Congress in amending the Clean Air Act in 1977 to explicitly call for the creation of an independent Clean Air Scientific Advisory Committee (CASAC). The CASAC, in accordance with the Clean Air Act (1977), has periodically reviewed the scientific basis for setting and revising the NAAQS for all the criteria pollutants. I have participated in most of those reviews and served as Chair of CASAC (1988-1992). CASAC has reviewed all of the Criteria Documents for criteria air pollutants prepared by EPA's Office of Research and Development and in some cases, health assessment documents for specific pollutants, such as diesel exhaust. In every instance, the CASAC modus operandi has included rigorous review of the document, receipt of input from the EPA staff and receipt of extensive written and oral comments from interested parties. All of these activities have been carried out in public sessions. On many occasions, the CASAC has offered comments to the Agency on multiple draft documents and, when it deemed the documentation scientifically adequate for regulatory decision making, provided a "closure letter" to the EPA Administrator. Without question, the CASAC has played a critical role in ensuring that the "final" criteria documents were of high scientific quality.

As the Criteria Documents grew in size the CASAC recognized the value of having documentation that could bridge from the science of the criteria document to the regulatory decision-making process. This was the genesis of the "Staff Papers" prepared by EPA's Office of Air Quality Planning and Standards. The CASAC, as with the Criteria Documents, reviewed the Staff Papers, received EPA input, received public comments and deliberated in public sessions on the scientific adequacy of the documentation. Frequently, the CASAC advised the Agency that the current version of the Staff Paper was not scientifically adequate and needed to be revised. In a manner similar to that followed with the Criteria Documents the CASAC provided a "closure letter" on the Staff Paper to the EPA Administrator when it deemed the Staff Paper scientifically adequate for regulatory decision making.

The discussion here is not intended to be an exhaustive review of all of CASAC's activities; rather the review has focused on the modus operandi of CASAC as a standing independent scientific committee. The activities of the CASAC, in my opinion, have been in accord with the language and intent of the Clean Air Act (1977) and consistent over time with the evolution of CASAC practices that have received substantial public and legal scrutiny. The modus operandi has proved successful in helping to ensure that the NAAQSs are science-based.

It now appears that parties within the EPA, but unknown to the public, are proposing to change the modus operandi of the CASAC. The arguments for change have been made in "administrative sessions" of the CASAC and, thus, have not been made public. As best I can discern the arguments are intended to relegate the CASAC to a status as an ordinary Scientific Committee operating under the Federal Advisory Committee Act (FACA) rules. The motivation for the change has not been articulated and it should be publicly articulated. Does the Agency believe that its ability to carry out the mandates of the Clean Air Act have been impaired by rigorous CASAC review and the use of a "closure letter" process? If so, this should be publicly documented. I would argue that to the contrary, even the delays resulting from CASAC's call for more rigorous documentation of the science have contributed to more defensible NAAQSs. Is the argument one that the CASAC is operating in a manner that is different from some other EPA FACA committees? If so, then the differences need to be publicly documented. Even if differences do exist in how CASAC operates versus other FACA committees does not make the CASAC past modus operandi inappropriate. The critical issue is whether the CASAC has and is operating in a manner consistent with the Clean Air Act language calling for an independent CASAC. Over the past 25 years, the Chairperson and members of CASAC have appeared before Congressional Committees on numerous occasions. My impression is that the Congress has consistently held a favorable view of the CASAC's modus operandi and its role in implementing the Clean Air Act. I am not aware that either the Congress or Executive Branch have advocated changes in how CASAC carries out its responsibilities.

In the absence of a publicly articulated basis for change that is also consistent with the Clean Air Act language establishing CASAC, why should the CASAC modus operandi be changed? In short, why tamper with a successful modus operandi!

**Chapter 3. Policy Relevant Assessment of Health Effects Evidence**

**A. General Comments – Chapter 3**

1. This chapter is not always balanced in its presentation of evidence and gives excessive weight to information that will support more stringent standards. It could be improved as noted below and in specific comments.

2. The chapter should be word-searched for all uses of the term – PM. In most cases it would be appropriate to be specific and use either $PM_{2.5}$ or $PM_{10\text{-}2.5}$. At this stage in the NAAQS review process, it is appropriate to be specific and avoid generalities such as PM.

3. The chapter should be word-searched for all uses of terms like "thoracic" particles or "accumulation" mode particles. In most cases it will be appropriate to be specific about the nature of the evidence for a specific PM indicator, i.e., $PM_{2.5}$ or $PM_{10\text{-}2.5}$ or $PM_{10}$.

4. The chapter underplays the substantial heterogenicity in excess risk for cities across the USA and Canada by giving only summary estimates from the NMMAPS (90-city) (Dominici *et al* 2003a) and Canadian (8-city) (Burnett and Goldberg, 2003) studies. The individual city estimates from both studies should be given in figures to complement Figure 3-1.

5. Most of the text appropriately summarizes the weak and inconsistent evidence for an association between urban $PM_{10\text{-}2.5}$ exposure and excess health outcomes. Unfortunately, the summary portions of the chapter overstates this evidence. The nature of the evidence is at best suggestive of a weak association, it certainly does not raise to a level that can be considered suggestive of causality.

6. The document should avoid the ambiguous characterization of the strength of the evidence for a $PM_{10\text{-}2.5}$ exposure-response association as being less than that for $PM_{2.5}$. The evidence for a $PM_{10\text{-}2.5}$ exposure-response relationship must be evaluated on its own merits.

**B. Specific Comments – Chapter 3**

Pg 3-2, line 7-9: Sentence does not make sense – "$PM_{10\text{-}2.5}$ are less well correlated" with what?

Pg 3-5, line 26-28: I suggest revising to read – "Removal of particles from the air stream by the extra-thoracic region is less efficient for accumulation-made fine particles, and thus, penetration of particles to the tracheobronchial and alveolar regions is increased (CD, 6-105)."

Pg 3-8, line 20: This statement is misleading, ROFA is an example of ROFA and nothing else despite the attempts of EPA ORD to portray it as a proto-typical example of combustion PM. Reword – "Administration of residual oil fly ash (ROFA) has been shown ---"

Pg 3-10, line 17: I have previously provided the authors of the CD a paper by Rothenberg *et al* (Surface area, adsorption and desorption studies in indoor dust samples. Am. Ind. Hyg. Assn. J. 50: 15-23, 1989) that clearly shows that only small quantities of formaldehyde are associated with PM and the dose from gas-phase formaldehyde to the upper airways is substantially greater than the delivered dose for formaldehyde associated with PM. Remove this inappropriate example – it is "folk lore" not supported by science.

Pg 3-11, line 12: Reword to avoid the implication that ROFA is a proto-typical combustion particle. Reword – "For example, using data from residual oil fly ash exposures, --"

Pg 3-17, line 4-5 and Figure 3-1. This figure should be complemented with figures showing the individual city results from the NMMAPS and Canadian Multi-City reports. It is misleading to present only the single composite value from the NMMAPS report (Dominici *et al*, 2003a) and the Canadian Multi-city Study (Burnett and Goldberg, 2003) and not present the individual city excess risk estimates for comparison with other individual city values shown in Figure 3-1. Many of the individual city values in Dominici *et al* (2003a) and Burnett and Goldberg (2003) have statistical "power" equal to or greater than the individual city values shown in Figure 3-1 and many of these cities fail to show statistically significant $PM_{10}$ effects.

Pg 3-22, line 7-10: For this key value give the confidence interval.

Pg 3-39, line 12-14: In my professional opinion, it is a stretch to indicate that limited suggestive evidence of an association between short-term (but not long-term) exposures and various mortality and morbidity effects raises to a level that can be considered "suggestive of causality." It would be more appropriate to simply say – "the very limited body of evidence is suggestive of an association between ---."

Pg 3-41, line 10-11: This statement is misleading in view of the limited, weak evidence for $PM_{10-2.5}$. The staff should not try to use the "$PM_{2.5}$" cloak to make the case for $PM_{10-2.5}$ effects. $PM_{2.5}$ and $PM_{10-2.5}$ are fundamentally different with regard to origin and chemistry as argued in many places in the CD and SP and in fact the differences can be used to argue for a distinction between the two particle sizes modes with regard to their potential for producing an increase in adverse health effects.

Pg 3-42. Several examples of sloppy writing by referencing PM rather than being explicit as to whether it is $PM_{2.5}$ or $PM_{10-2.5}$.

Pg 3-45, line 5-7: Reword – "determined by the difference between $PM_{10}$ and $PM_{2.5}$ measurements"

Pg 3-47, line 27: Why not be specific – $PM_{2.5}$ rather than accumulation-mode particles.

Pg 3-67, line 1: In my professional opinion, the evidence does not support the statement – "but suggestive evidence of causality for short-term exposures to $PM_{10-2.5}$." There is at best weak evidence of an association but it falls far short of evidence of causality.

Pg 3-67, line 4: I agree that the quantitative assessments for $PM_{2.5}$ can inform decisions on the NAAQS for $PM_{2.5}$. However, I am concerned at the excessive reliance being placed on complex statistical models and calculations to derive highly uncertain estimates of excess morbidity and mortality related to $PM_{2.5}$ exposures. I do not think the evidence of a very weak association between $PM_{10-2.5}$ exposure and health outcomes is sufficient to warrant the use of the quantitative assessment for $PM_{10-2.5}$ in making decisions about the setting of a NAAQS for $PM_{10-2.5}$ that would be applicable to the entire U.S. for which $PM_{10-2.5}$ is remarkably varied in origin and chemical composition.

## **Chapter 4: Characterization of Health Risks**

### **A. General Comments – Chapter 4**

1. As best I can discern it accurately summarizes the findings presented in the Technical Support Document (Abt, 2005). Unfortunately, the Support Document tends to selectively use information that supports arguments for more stringent standards. The support document fails to adequately relate the high degree of uncertainty that exists in understanding $PM_{2.5}$ exposure-response relationships across the United States at the present time. Moreover,

quantitative data frequently tends to convey a level of certainty that does not exist in the underlying data.

**B. <u>Specific Comments – Chapter 4</u>**

Pg 4-1, line 3: Reword – "assessment that <u>was</u> conducted"

Pg 4-6, footnote: The chapter should minimize use of the term PM in favor of being specific as to the indicator being discussed, $PM_{2.5}$, $PM_{10}$ and $PM_{10-2.5}$, especially in discussing results.

Pg 4-56, line 26: The appropriate term is "similar" rather than "comparable." Many things can be compared, only some things are similar.

## <u>Chapter 5: Staff Conclusions and Recommendations on Primary PM NAAQS</u>

**A. <u>General Comments – Chapter 5</u>**

1. This chapter conveys an excessive degree of certainty with regard to scientific knowledge of $PM_{2.5}$ exposure-response associations and $PM_{10-2.5}$ exposure-response association. The authors have got it right on page 5-75, line 12-14 –"Staff believes" it is important to continue to highlight the <u>unusually large uncertainties</u> associated with establishing standards for PM relative to other single component pollutants for which NAAQS have been set." This accurate statement should also be placed at the beginning of the chapter to serve as guidance both for the authors and readers. The authors need to be especially mindful that although numbers can be presented very precisely they may still have substantial underlying uncertainty. In short, the precise calculations of the Abt, 2005 Support Document, should be used with a high degree of caution. I personally give limited weight to the calculated estimates of health impacts for $PM_{2.5}$ concentrations in the range of and below the current standards.

2. It is critical that the terminology used throughout this chapter be a specific as possible, especially with regard to the use of phrases like thoracic coarse particles and fine particles and specific indicators such as $PM_{10}$, $PM_{2.5}$ and $PM_{10-2.5}$. It is important to recall that the $PM_{10}$ standard was never set to be solely protective of coarse particles; it was originally set to be protective of both coarse and fine particles.

It would be appropriate to run a search to identify in the Staff Paper all uses of the term PM and then verify if it is used appropriately. In some cases it may be appropriate to use PM as a general descriptor. However, in most cases a more specific term such as $PM_{10}$, $PM_{2.5}$ or $PM_{10-2.5}$ may be more appropriate.

3. The chapter should be critically reviewed and revised as necessary to create a more neutral tone. The present chapter in many places conveys the view that the EPA was right in 1997 in setting the $PM_{2.5}$ NAAQS and the intervening events have proved the EPA right in spades. I do not hold the same view. The level of uncertainty was very substantial in 1997 when the $PM_{2.5}$ standard was set largely based on inferences from $PM_{10}$ exposure-health associations. It is arguable as to the extent uncertainties have been reduced since the $PM_{2.5}$ was set. Moreover, only modest data are yet available on $PM_{2.5}$ exposure-health associations.

It is noteworthy that the extent to which $PM_{1.0}$ would be a more appropriate indicator than $PM_{2.5}$ has not been rigorously evaluated largely because of a lack of $PM_{1.0}$ monitoring data. As a key EPA staffer said in the 1990's review, "you are never going to get $PM_{2.5}$ monitoring data unless you set a $PM_{2.5}$ standard." He was right – we now have $PM_{2.5}$

monitoring data, but no $PM_{1.0}$ data. I wish he had advocated the need for monitoring data on $PM_{1.0}$, $PM_{2.5}$, $PM_{10\text{-}2.5}$ and $PM_{15\text{-}2.5}$.

4. The biased tone of the chapter is especially apparent in the discussion of information on $PM_{10\text{-}2.5}$ exposure-response associations. The extent and strength of the evidence is regularly over-stated and inappropriate conclusions drawn with regard to the setting of a $PM_{10\text{-}2.5}$ NAAQS.

5. The section on "Indicators" represents "revisionist history" at its best. This section needs to be rewritten to reflect reality. The history of developing scientific knowledge on PM exposure-health associations and the setting of the NAAQS for successive PM indicators is very inter-twined. When the first PM NAAQS was set in 1971 using a Total Suspended Particulates (TSP) indicator most of the available epidemiological data was based on Black Smoke, coefficients of haze or even "Stinking Smog Days." With the setting of the TSP standard, monitoring data began to be collected on TSP. Subsequently, epidemiological studies were conducted using TSP as the exposure metric.

In the 1970s and early 1980s, a large amount of human data on the fractional deposition of radioactive particles of different sizes in various regions of the respiratory tract began to be reported. This gave impetus in 1987 to the setting of a size-based PM indicator, the $PM_{10}$. As an aside, it might well have been set at $PM_{15}$ which is yet another story. However, most of the epidemiological data available to set the $PM_{10}$ standard were based on the TSP metric. Subsequently, during the late 1980s and 1990s, substantial monitoring data became available on $PM_{10}$. Interestingly, some limited monitoring of $PM_{2.5}$ was discontinued during this time period in deference to the selection of a $PM_{10}$ indicator.

In 1997, a NAAQS was set using $PM_{2.5}$ as an indicator. This standard was largely based on inferences made from epidemiological studies using $PM_{10}$ as the exposure metric. As an aside, the standard might well have been set using a $PM_{1.0}$ indicator which is yet another story. The setting of the $PM_{2.5}$ NAAQS resulted in the deployment of the $PM_{2.5}$ monitoring network for "regulatory compliance purposes." Now in 2005, the $PM_{2.5}$ NAAQS is being reviewed largely using $PM_{10}$ epidemiological data and somewhat more $PM_{2.5}$ data than existed in the mid-1990s. In addition, as a result of the Supreme Court decision, consideration is being given to a $PM_{10\text{-}2.5}$ standard based on very limited data on urban $PM_{10\text{-}2.5}$ exposure-health associations and consideration of the $PM_{10}$ exposure-health database. Using a 0-10 scale, I would argue that if the scientific knowledge for $PM_{10}$-related effects is 8, then the $PM_{2.5}$ evidence might warrant a 5 and the $PM_{10\text{-}2.5}$ evidence is something less than 1.

It is appropriate to ask if these are the best indicators for PM? The answer is we do not know! It is quite possible that $PM_{1.0}$ might be a better indicator for fine particles and their effects. It is also quite possible that $PM_{15\text{-}2.5}$ might be a better indicator for coarse particles and their health effects. In addition, several PM components might warrant consideration as indicators on a size-specific chemical mass basis. The point I want to emphasize is that EPA should not attempt to rewrite history as though it arrived at today's position as a result of some science-based decisions in which careful consideration was given to a range of options. The options have always been limited by EPA-dictated orientation to "regulatory compliance based monitoring" rather than using a science-based strategy.

Unfortunately, I think we will be in the same deplorable "science deficient situation" in 2010 or 2011 attempting to defend a $PM_{2.5}$ "house of cards" unless a conscious decision is made to create a different kind of scientific basis for decision-making on the NAAQS for PM. A starting point is to abandon the current excessive focus on "regulatory compliance monitoring" in favor of a more balanced approach that gives equal weight to "acquisition of new scientific information." This would certainly include $PM_{1.0}$ monitoring and related epidemiological studies.

## B. <u>Specific Comments – Chapter 5</u>

Pg 5-7, line 6: Reword – "by which $PM_{2.5}$ exposure" ---" Avoid using the ambiguous term – PM – when a more specific term, $PM_{2.5}$ or $PM_{10-2.5}$ is appropriate.

Pg 5-8: Numerous examples of inappropriate use of PM.

Pg 5-17 to 5-21: The section on indicators needs to be revised to reflect a more accurate picture of what occurred historically over the past 25 years as discussed above. The present version presents a "revisionist historical" version that is disconnected from what really happened. A key background point is that epidemiological investigations can only evaluate "PM indicators" that have been measured. Initially, the bulk of the PM measurements were made using Black Smoke and coefficients of haze as indicators and, increasingly, during the 1960s and 1970s using Total Suspended Particulates (TSP). With the NAAQS set in 1971 using TSP as an indicator "the law of the land" dictated TSP monitoring. Hence, more TSP data was available for the conduct of epidemiological studies. In the early 1980s, there was discussion of a range of potential indicators including $PM_{15}$ and $PM_{10}$. In 1987, the PM NAAQS was set using a $PM_{10}$ indicator. Again, the "law of the land" required $PM_{10}$ monitoring. Hence, as $PM_{10}$ monitoring data became available more epidemiological studies were conducted using it (recall the NMMAPS research was conducted using the $PM_{10}$ indicator because it was available). Following the 1997 PM NAAQS promulgation with $PM_{2.5}$ as an indicator a new national monitoring network using $PM_{2.5}$ monitors was deployed. The $PM_{2.5}$ monitoring data from this network is just becoming available and will increasingly be used in epidemiological studies. Unfortunately, a lack of PM data will preclude direct comparisons between $PM_{2.5}$ and $PM_1$ as indicators.

In short, over the last three decades it has been the regulatory compliance monitoring data that has served as the primary input for the conduct of epidemiological investigations, i.e. the light under the "regulatory lamp post." Indeed, there has been considerable "back-filling" of support data for the successive PM standards (TSP, $PM_{10}$ and $PM_{2.5}$) after they have been promulgated. The TSP standard was set primarily on Black Smoke data, the $PM_{10}$ standard was initially set primarily on TSP data and the $PM_{2.5}$ standard was initially set based primarily on $PM_{10}$ data. Indeed, today the most substantial epidemiological database being used to re-evaluate the $PM_{2.5}$ indicator is that based on epidemiological studies using $PM_{10}$ as the indicator.

It is absolutely ridiculous for the SP to imply that the EPA Staff considered both a $PM_{2.5}$ and $PM_{1.0}$ cut-point (see pg 5-20, line 1-17). This was simply not possible. There is essentially no epidemiological data on a $PM_{1.0}$ indicator. Why not? Because the "regulatory lamp post" was set up at 2.5 µm. Unfortunately, neither the EPA staff nor CASAC were forward looking with regard to alternative size-based indicators. If they had been, at least a limited network of 1.0 monitors would have been deployed. The same can be said with regard to the failure to deploy appropriate monitors to collect $PM_{10-2.5}$ data that could be used in epidemiological investigations that would provide a scientific basis for considering a $PM_{10-2.5}$ standard rather than the "hand-waving" approach of the present document.

The same case can be made for obtaining epidemiological evidence on specific PM chemical components. There will continue to be an absence of evidence on PM components until such time as specific chemical components are widely monitored for extended periods of time and the resulting data used in epidemiological investigations.

Pg 5-20 to 5-21: While the emphasis on improving the Federal Reference Method for $PM_{2.5}$ measurements is appropriate, I would argue it is insufficient. Why not equal consideration given to the $PM_{10-2.5}$ indicator?

Pg 5-22, line 6: It is a stretch to say there is a "growing body of studies that provide additional evidence of effects associated with exposure periods shorter than 24 hours (e.g. one to several hours)." It would be more appropriate to characterize this as – "a limited body of data suggests." The jump from this inaccurate characterization to "consideration of a short-term standard in the future" is totally inappropriate. I urge the agency to not attempt to create a "self-fulfilling" prophecy. The issue of selecting "averaging times" is much more complex than simply finding some studies showing changes with short-term exposures to some regulated compounds or mixture. The same arguments made in the next paragraph against "multiple-day averaging times" apply to averaging times less than 24 hours.

Pg 5-22, line 33: Reference is made to both a "significant harm level program" and the Air Quality Index. Both of these should be more fully described since I suspect some CASAC members and many readers will be unaware of some of the subtle characteristics of these activities as regards different PM indicators.

Pg 5-55, line 12-21: Why not stick with the facts, i.e. "Because epidemiological evidence for PM effects is dependent upon information from the "regulatory compliance monitoring" network that has successively focused on TSP, $PM_{10}$, and $PM_{2.5}$ metrics, these are the only potential indicators that can be considered based on scientific information." The rest of this paragraph, which is future oriented, should become a separate paragraph. In a new paragraph, reference should be made to the need for regular monitoring of a few selected PM chemical components as a basis for future epidemiological investigations.

Pg 5-58, line 1-4: It is inappropriate to characterize – "a growing but still limited body of evidence on health effects – that directly use an indicator of $PM_{10-2.5}$." Why not say it the way it is – "There has been only limited evaluation of the association between urban $PM_{10-2.5}$ exposure and adverse health effects."

Pg 5-59, line 23-26: This is masterful double-talk.

Pg 5-60, line 8-16: Why not be straightforward – "The present very limited data evaluating the association between urban $PM_{10-2.5}$ exposure and adverse health effects does not allow rigorous evaluation of the presence or absence of a threshold for this possible association."

Pg 5-60, line 17 and beyond: The authors should explicitly call attention to the challenge of evaluating $PM_{10-2.5}$ exposure-health association against a background of $PM_{2.5}$ exposure-health outcomes.

Pg 5-63, line 10-16: Why not be direct – "Since EPA has in the past elected to use $PM_{10}$ and $PM_{2.5}$ as indicators and deploy associated monitoring networks for these indicators, the EPA has no choice today other than consider a $PM_{10-2.5}$ indicator for thoracic coarse particles." There is no need to try to recast the sow's ear as a silk purse.

As a westerner, I appreciate line 14-16. However, I should note that it is more difficult to implement and comply with EPA's "natural events policies" than it is to offer this "do not worry, the EPA is really here to help you" statement.

Pg 5-64, line 23-28: In my opinion, this is an over-statement of the extent and nature of the evidence. The phrase "causal associations" should be removed. On page 5-65, line 8, is a more appropriate statement – "evidence suggestive of associations between short-term exposures and morbidity effects."

Pg 5-67, line 15-17: Earlier in the SP, the staff appropriately calls attention to the difficulty in evaluating the $PM_{10-2.5}$ data for the cities where effects are dominated by $PM_{2.5}$ exposures. Now the staff appears to have forgotten those limitations. In my opinion, it is totally inappropriate to use analyses based on Seattle, Toronto and Detroit to argue for 24 hour $PM_{10-2.5}$ standards that are going to have their greatest impact in the Southwest U.S. (see Table 5-6).

I suggest the staff re-write this section with less "hand-waving" and say – "The results of a very limited number of studies of the association between $PM_{10-2.5}$ exposure and morbidity and mortality are not sufficient for use in setting a 24 hour $PM_{10-2.5}$ standard." In my opinion, it is inappropriate to argue for setting a $PM_{10-2.5}$ standard based on limited studies of Seattle, Toronto and Detroit and two very limited studies of Phoenix and Coachella Valley, CA. The proposed range of 65 to 75 µg/m$^3$ with a 98$^{th}$ percentile form would place 40 to 45% of the counties in the Southwest in the "not likely to meet standards" "category." I make this point to help my colleagues on the CASAC PM Panel and the EPA staff appreciate the serious ramifications of the process we are involved in. I will also recall for you that none of the cities in this region on an individual city basis had statistically significant positive association between $PM_{10}$ and adverse health outcomes in the NMMAPS evaluation.

Pg 5-74, line 25: It is inappropriate to use a circular argument based on the $PM_{10}$ standard as a basis for setting the $PM_{10-2.5}$ standard.

Pg 5-75, line 12-14: This sentence is the most important sentence in the entire document and should be set in bold type with "unusually large uncertainties" set in red type.

Pg 5-76 to 5-78: I find it disappointing that the EPA staff could be so myopic as to write two pages on uncertainties and research needs and fail to address the issue of the appropriateness of a 2.5 µm cut point versus 1.0 µm cut point. This is an extremely important issue that needs to be addressed based on scientific evidence from epidemiological investigations. It is important to recall that the 2.5 µm cut point was not selected over a 1.0 µm cut point based on rigorous scientific and technical debate. The 2.5 µm cut point, as I recall, was selected based on one EPA supervisor's views of the technical feasibility of designing and operating a $PM_{2.5}$ monitoring device versus a device with a 1.0 µm cut point. The outcome of this argument, now known to be without merit, has shaped regulations for PM that have billions of dollars of impact annually. I ask my fellow CASAC PM panel members how confident they are that a $PM_{2.5}$ standard is more appropriate than $PM_{1.0}$?

I urge the Clean Air <u>Scientific</u> Advisory Committee PM Panel to challenge the EPA staff on the wisdom of continuing down the $PM_{2.5}$ path as though it is the only option. It is time to shift from a "regulatory compliance dominated" monitoring network to one that balances "regulatory compliance" with the potential for "acquisition of new scientific knowledge" and the opportunity for improved science-based standards.

## <u>Comments on Staff Recommendations on NAAQS</u>

### A. <u>General Comments</u>

In my opinion, the EPA staff are placing excess emphasis on the results of quantitative risk assessments conducted using highly uncertain and highly selected input data. I seriously question the validity of the calculations of excess morbidity and mortality for $PM_{2.5}$ concentrations approaching and below the current annual $PM_{2.5}$ NAAQS – 15 µg/m$^3$ or the 24-hour average concentrations below about 50 µg/m$^3$. The calculated excess health effects for

$PM_{10-2.5}$ are even more uncertain, and in my opinion, are of limited use in making a decision on the selection of a $PM_{10-2.5}$ indicator to be used nationwide.

## B. Primary $PM_{2.5}$ NAAQS

1. Indicator: As noted elsewhere there is no choice but to continue with a $PM_{2.5}$ indicator for fine particles. I personally wish that data were in hand to consider a $PM_{1.0}$ option. I strongly suspect with monitoring data available on both $PM_{1.0}$ and $PM_{2.5}$ and associated epidemiological results, that $PM_{1.0}$ would prove to be the better indicator and result in more efficient and effective control strategies.
2. Averaging Times: There is only limited data available to consider any options other than 24 hour and annual averaging times.
3. Statistical Form: I favor retaining the 98$^{th}$ percentile form for the 24 hour averaging time and averaging over three years for the Annual Standard.
4. Numerical Level: In view of the substantial uncertainties in the database available on $PM_{2.5}$ I favor retaining the current annual standard of 15 µg/m$^3$ and revision of the 24 hour averaging time standard with a value as high as 50 µg/m$^3$ with a 98 percentile form.

## C. Primary $PM_{10-2.5}$ NAAQS

1. Indicator: The selection of a $PM_{10-2.5}$ indicator is without scientific merit and would represent an arbitrary and capricious choice based solely on the perceived need to have a "place holder" coarse PM indicator. Alternatively, I would find any urban $PM_{10-2.5}$ PM indicator acceptable.
2. Averaging Time: No scientific basis. In my opinion, there is not an adequate database for selecting an annual averaging time or a 24-hour averaging time standard in the absence of restricting the $PM_{10-2.5}$ indicator to urban aerosols.
3. Statistical Form: No scientific basis in the absence of restricting the $PM_{10-2.5}$ indicator to urban aerosols.
4. Numerical Level: No scientific basis in the absence of restricting the $PM_{10-2.5}$ indicator to urban aerosols.

## Dr. Günter Oberdörster

### Comments on OAQPS Staff Paper, 2nd draft (G. Oberdörster)

Overall, the staff paper provides a very good, although rather lengthy, summary of the CD, and recommendations for the primary PM standards are well developed and justified. The document now clearly distinguishes between fine and thoracic coarse particles throughout the staff paper. Air quality characterizations and analyses are clearly communicated, and it provides in my view a very sufficient background for the human health and visibility assessments. I have only a few comments and suggestions for changes:

In Chapter 2, the word "Aitken" is consistently misspelled which should be corrected. Also, on page 2-3, the term "microns" should be replaced by "micrometer" or "µm".

Add the word "mass" before "ratios" in the figure legend on page 2-45.

Page 3-5, lines 28-30: Although it is correct that fractional deposition expressed by respiratory tract region is greatest in the alveolar region for these particles, on an epithelial cell surface area basis the deposited dose is greatest in the tracheobronchial region, which in many cases is a more important determinant for effects than fractional regional deposition.

Page 3-6, line 22: Delete the word "rapidly".

Page 3-8, lines 12-16: The statement that studies using intratracheal instillation of ambient particles from different locations can cause lung inflammation and injury is not very enlightening, any particle given at high enough doses can do this. What is of importance is whether these effects are greater compared to a reference particle of low toxicity given at the same doses.

Section 5.5, Summary of key uncertainties and research recommendations:

Although this section may not be a major part of the staff paper, my suggestion is to emphasize more an important concept for future research and express the need to strengthen the database to be used for regulatory purposes. Thus, extending the thoughts expressed in point #2 on page 5-77, lines 6-9, regarding the need to help identify PM components and characteristics, or sources of PM, that may be linked with various effects, the staff paper could more clearly outline the concept that future research – and probably future NAAQS – should not only focus on PM *per se*, but on specific chemical constituents. This relates to all sizes of PM ($PM_{10-2.5}$, $PM_{2.5}$ and $PM_{0.1}$). With a greater emphasis on PM chemistry, and increasing respective database, there might perhaps be a future PM standard combining size and specific chemical components. (*The chemical composition of the PM mixture – regardless of the size – varies by location, and it is no surprise that results from different groups using PM from different locations report widely varying effects: For example, at this year's SOT meeting, results of a presentation entitled "Exposure to concentrated ambient particles does not affect endothelial vasomotor function in patients with ischemic heart disease" contrast with other reports where such effects have been observed. The major contribution of sodium chloride in the urban aerosol of the SOT presentation likely explains these differences*). Thus, PM mass standards alone may have to be changed in the future by considering chemistry, which would also allow [researchers] to identify and go after sources. Such focus on the chemistry would also be in line with the staff paper's recommendation with respect to a secondary PM standard where chemistry is emphasized (pg. 7-22, lines 18-20). On another issue of PM components, the CD and also the staff paper devote significant sections to ultrafine particles, yet there is no mention of this potentially very important component of $PM_{2.5}$ in these recommendations. One would be that monitoring of UFP in several major cities should begin so that time series epidemiological studies can be completed in time for the next NAAQS review. With increasing activities in toxicology on potential effects of nanoparticles (<100 nm), it is becoming more and more obvious that nano-sized particles – which include ambient ultrafine particles – do, indeed, have very specific toxicological properties, specifically with respect to translocating to many different tissues of the organism following inhalation.

**Dr. Robert D. Rowe**

# Memorandum

| | |
|---|---|
| **To:** | Fred Butterfield |
| **From:** | Robert Rowe |
| **Date:** | 4/13/2005 |
| **Subject:** | January, 2005 PM Staff Paper |

My comments on the January, 2005 Draft Staff Paper for PM focus on materials damage and visibility aesthetic welfare effects.

### *Materials Damage*

As identified by Warren White, high levels of $PM_{10-2.5}$ can be expected to be associated with adverse materials soiling impacts, and should be identified as another consideration for a primary/secondary standard for $PM_{10-2.5}$.

### *Visibility*

The Draft Staff Paper visibility section is well conceived and well written.
The Staff Paper has clearly identified that the public places importance (e.g., value) on visibility aesthetics, both inside and outside of Class I areas.
Currently, the secondary 24 hour standard of 65ug/m$^3$ is too lenient to protect visibility values. This standard allows miles of visual ranges in the single digits, which has been found undesirable in virtually every study conducted over the past several decades. See, for example, the citations provided by EPA on how citizens and government agencies rate such low visibility levels, as well as Table 1 in Chestnut and Dennis (cited in the Staff Paper) reviewing eastern and western urban visibility economic studies where citizens were willing to pay significant amounts to improve visibility over current conditions, based on annual average levels, and willingness-to-pay (WTP) to improve poor visibility levels of a limited set of poor days (McClelland et al. and Carson et al./Mitchell et al.**)**. I am not aware of any evidence (nor has the public raised any) that suggests the current standard is protective of visibility value.
Reflecting the public importance of visibility, EPA has made constructive steps to design a well-reasoned visibility specific secondary standard in terms of the locations of interest, indicator measure and time period. Some concerns remain with the level and form of the proposal.

- *Location.* Focusing on non-Class I locations appears reasonable given the regional haze program specifically targets Class I areas. However, one cannot simply assume the regional haze program will also protect urban areas across the country due to differences in locations of cities and Class I areas and worsening haze in the western U.S..

- *Indicator*. The use of a $PM_{2.5}$ indicator is a reasonable means to focus on key light scattering particles and to avoid issues with weather impairment of visibility.

- *Indicator period*. A 4 to 6 hour middle of the day period focuses on a consistent year-round daylight period of when visibility matters and when the impacts of relative humidity are the least confounding across locations. This is a significant improvement over using a 24 hour period indicator. A rolling 4-hour period may ultimately be desirable to consider, but presently adds a complication that seems likely to provide limited additional benefit given other uncertainties. In support of this approach, increased continuous monitoring is warranted.

- *Level*. The proposed level of 30 to 20 ug/m$^3$ (resulting in visual range of 25 to 35 km, or 15.5 to 21.7 miles) is selected as being conservative primarily vis-a-vis levels selected as being adverse in Denver (50 km), Phoenix (36 to 48 km) and Lake Tahoe (48 km); levels suggested by other cited evidence that is less directly relevant; and staff visual observation of simulated conditions with these and alternative levels.

  - Based on the cited evidence, the proposed level clearly is conservative for the Mountain West, and likely is conservative for other locations with scenic vistas in excess of the 15 to 22 mile range in the proposed standard..

  - Whether the proposed level is conservative for other Midwest, East, and Pacific Coast locations (where many of the violations may arise) is not certain from the available evidence. One might presume that a level lower than the Mountain West is reasonable due to differences in humidity and the nature of viewing distances and objects (and thus is why a level is proposed that is much lower than appropriate for cities in the Mountain west). Older economic studies (reviewed in Chestnut and Dennis, cited in the Staff paper) clearly reveal public WTP to improve average daytime visual range in eastern cities from poor levels (typically 15 miles or less) to improved levels (typically 18 to 38 miles). Tolley et al (cited in Chestnut and Dennis) also found that at existing visibility levels, Chicago area residents reduced behaviors such as viewing, participating in outdoor activities, and increased TV viewing. Nothstein[3] and others (see Chestnut and Dennis) report public values for improving visibility on the West Coast. Thus, these studies support the importance of improving visibility in these locations from conditions allowed under the current standard, but (as so far interpreted) are less informative about the appropriate level and form. This reemphasizes the importance of EPA moving forward with public attitude and value studies, as proposed in 2000, to refine the basis for a level and form of this standard.

- *Form*. Given the uncertainty in setting a level, at least for some locations, EPA proposes a lenient statistical form of the standard (90% of days target). This allows 36 visibility impairment days per year, in addition to bad weather days. In some locations, this would exempt a large share of days in the visibility impairment season, and thus may provide limited protection. On that basis, a 90% form seems too lenient. The Denver and Phoenix standards identified are more restrictive. McClelland et al. (1991) find a significant WTP for improving visibility on the worst 25% of days in two eastern cities, and a pilot economic study in Cincinnati (Carson et al., Mitchell et al.) finds overwhelming support (about 90% with WTP) for programs that would improve poor visibility on 13 to 29 days a year (both studies attempted to separately account for visibility versus health benefits).

While a tightening of the secondary standard to protect visibility is warranted, the specific level and form of the proposed standard is a policy decision as to whether it is worse (more social harm) to continue with something similar to the current secondary standard (which may allow

significant visibility impairment) or to move forward to select a more protective secondary standard based on limited information recognizing that the level selected may not be optimal.
**Recommendations**

Based on my review of the literature, experience, and professional judgment, I recommend:

- Moving to the structure of a visibility secondary standard as outlined in the Staff paper – use of a 4 hour mid-day standard for non-Class-I areas, but with a statistical form greater than 90%, as EPA may recommend based on further evaluation).

- EPA augment its analysis and presentation by providing data on how alternative proposed 24 hour primary standard would translate in terms of the comparable level under the proposed visibility specific standard structure (with a > 90% statistical form), and provide information about the locations where violations would occur (such as the map provided in EPA handouts at the meeting).

- At a minimum, a secondary standard would be set with a level "comparable" with the primary standard, but expressed in a visibility specific standard structure – rather than just adopting the primary 24 hour standard as in the past.

- EPA should move forward with new public attitude and valuation studies to enhance the inputs to the determination of the specifics of the visibility standard for the next cycle.

### *Minor comments*

1. Page 6-12. I would like to get a copy of Schmidt et al. 2005.

2. Page 6-16. Line 10. Remove "significant". Lines 11-14 unnecessarily repeat lines 5-6 and 7-8, with some of this repetition removed.

3. Page 6-23. Lines 3-6 needs a citation.

4. Page 6-25. Lines 6-9 have repetition.

**Additional References**


Carson, R.T., R.C. Mitchell, and P.A. Ruud. 1990. *Valuing Air Quality Improvements: Simulating a Hedonic Equation in the Context of a Contingent Valuation Scenario.* In Visibility and Fine Particles, edited by C.V. Mathai. AWMA Pittsburgh, PA.

Mitchell, R.C., R.T. Carson, and P.A. Ruud. 1989. *Cincinnati Visibility Valuation Study: Pilot Study Findings*, Report prepared for the Electric Power Research Institute, Palo Alto, CA.

Nothstein, G. 1998. *An Evaluation of Public Willingness to Pay for Improvements in Visibility and Air Quality*. Masters paper for the Department of Environmental Health, University of Washington, Seattle.

McCcclelland, G., W. Schulze, and D. Waldman et al. 1991. *Valuing Eastern Visibility: A Field Test of Contingent Valuation*. University of Colorado report prepared for the U.S. EPA under Cooperative Agreement #CR-815183-01-3. Washington, D.C.

# Dr. Jonathan M. Samet

COMMENTS ON CHAPTER 4, OAQPS STAFF PAPER ON PM NAAQS

Jonathan M. Samet
James D. Crapo

March 30, 2005

## INTRODUCTORY COMMENTS

This chapter provides the general methodology, findings, and sensitivity analyses for EPA's risk assessment of $PM_{2.5}$ and $PM_{10\text{-}2.5}$. It is supported by the full Technical Support Document and associated appendices. The methods used in these documents have undergone review by CASAC as well as public comments. The chapter considers the morbidity and mortality burden associated with PM and the benefits of attaining the current standards, as well as several scenarios of more stringent standards. The findings of the risk assessment figure centrally in the recommendations of the Staff Paper.

## GENERAL COMMENTS

- One major concern with the current version of the chapter is the clarity of presentation. Readers need to struggle through dense prose and jargon-ridden prose to identify key aspects of the methods and findings. Concern about the document's style is more than cosmetic, as the risk assessment needs to be clearly presented so that there is no ambiguity as to its findings. In this regard, key terms are sometimes used incorrectly or inconsistently across the chapter. The chapter could be substantially shortened.
- Figure 4-1 provides an overall framework for the risk assessment that could be used to shape the chapter. It shows where sensitivity analyses are carried out and even numbers them by subscript. This potentially valuable framework is not subsequently utilized, however. We suggest that the chapter refer to it repetitively as the risk assessment methods and findings are described. The various sensitivity analyses might be listed in expansions of the "diamonds" on the figures.
- Subheadings might be more effectively used to guide the reader through the individual sections of the chapters. For example, clearly listing "assumptions" and "sensitivity analyses" so that the distinctions are clear and uniformly worded across sections.

## METHODOLOGICAL CONCERNS

- The selection of C-R relationships is premised in the concept that locally-derived coefficients are likely to be most appropriate. The Staff Paper mentions the possibility that the suite of potential confounding and modifying factors may vary from location to location. Is there a basis for assuming substantial variation? Is effect modification anticipated on the relative risk scale on which the risk assessment is carried out? There is evidence that coefficients from single-city time-series analyses tend to be biased upwards, in comparison to those from multi-city analyses (Dominici et al, in press).

Additional variability is introduced by variations in methods from analyst to analyst. These issues need discussion.

- In calculating the burden of associated incidence, the risk assessment uses either the predicted background or the lowest measured level in the utilized epidemiological analysis for the counterfactual.  We suggest that the background level be used throughout to eliminate a needless difference in approach across locations.  While there may be some further uncertainty in extending the C-R relationship beyond the lowest measured level, the larger uncertainty comes with the reliance on a linear, non-threshold model.
- The analyses in this chapter highlight the impact of the assumption of a linear nonthreshold model in overestimating actual risk. The absence of data near the threshold does not imply the absence of a threshold. Threshold models should be emphasized in this risk assessment. A major research need is for more work to be done to determine the correct threshold.
- Uncertainty receives comment throughout the chapter. Its inherent asymmetry needs acknowledgment; i.e., uncertainty is greater for scenarios set at lower and lower concentrations.

**SPECIFIC COMMENTS**

Page 4-2, first paragraph:  There are methods for characterizing uncertainty beyond probabilistic judgments of "health scientists."

Page 4-2, line 8:  Confused sentence conceptually; Is the reference to statistical variability or to population variation—quite distinct concepts?

Page 4-2, second paragraph:  See comments above.  Ideally, a multi-location analysis would be done, if the data were available.  Reliance on single-city analyses by individual analysts suffers from both variation in methods and limited precision.

Page 4-3, line 8:  "precise measures" should be "certain measures", one of many examples of careless wording.

Pages 4-6 and 4-7:  The discussion of causality remains muddled.  As a first question, EPA should determine whether $PM_{10}$ or $PM_{10-2.5}$ is causally associated with injury and adverse health effects and then select epidemiological or population indicators of the injury to health for use in the risk assessment.  The sentence concluding the first paragraph on page 4-7 is not clear. There is also inconsistency in the chapter's discussion of the level of causation inferred for $PM_{10-2.5}$ which is given as "causally related" here but "suggestive" elsewhere (see page 3-67, line 1; page 4-40, line 23).

Page 4-8, line 14:  should read:  ""intended to provide protection from health effects of ambient PM."

Page 4-27, line 22:  would not use the phrase "mortality incidences" here or elsewhere in the document.  Consider "mortality events".
Page 4-53, full paragraph:  The discussion of the basis for selecting the "thresholds" should be expanded.

**RESPONSES TO EPA QUESTIONS**

- Question 3, PM-related health effects, risk assessment, and health-based standards (Chapters 3, 4, and 5).

Chapter 3 offers a general review of the epidemiological literature on thresholds (Section 3.6.6). The focus on this topic is applauded and the consideration of a threshold represents the largest factor in subsequent quantitative risk assessment in Chapter 4. This discussion reviews some of the relevant epidemiological literature but has no grounding in relevant toxicologic or mechanistic considerations. It does not lend direct support to the thresholds picked for sensitivity analyses in Chapter 4. A figure should be used to explain the slope adjustment in the "hockey stick" models.

- Question 4a.

In general, the set of health endpoints selected is appropriate and supported by relevant studies. We are concerned by the reliance on single-city analyses as a precedent and urge that multi-city analyses, once available, be used in future risk assessments. In this instance, there is not great variability across the C-R relationships selected.

- Question 4b

With regard to inclusion of mortality associated with $PM_{10-2.5}$ in the risk assessment, we are in agreement with not including such estimates. The epidemiological literature is mixed and there are inherent limitations to their findings, including the problem of measurement error for this derivative PM indicator and the difficulty of estimating a possibly separate effect from that of $PM_{2.5}$.

- Question 4c

- Question 4d

With regard to the handling of uncertainty in the risk assessment, an overview of the model is supplied in Figure 4-1, and key sensitivity analyses are indicated. Pages 4-37 through 4-41 offer a descriptive summary of the findings of these analyses. This section might be strengthened by adding the quantitative findings of these analyses, rather than including very limited verbal descriptions. It is unfortunate that a more comprehensive, quantitative characterization of uncertainty has not been undertaken, even if it only took into account several sources of uncertainty simultaneously. The chapter acknowledges this limitation of the risk assessment. There is also likely to be directionality to the degree of uncertainty, with greater uncertainty around effects at lower, compared with higher PM levels. Overall, the chapter tends to understate uncertainty, both through style, (e.g., inclusion of numerically specific estimates,. e.g., "403" deaths rather than "400" or "about 400", and by not bringing together the individual sensitivity analyses.

- Question 5. We agree with the general views and approach taken by staff in Chapter 5. We agree with the emphasis on the quantitative risk results for $PM_{2.5}$ and with a general approach on the use of $PM_{10-2.5}$ risk assessment.

- Question 6.  We agree generally with the proposed alternatives for primary standards for fine particles.  The range of proposed standards are consistent with the available scientific information.

- Question 7.  We agree with the proposed alternative standards for thoracic coarse particles.  The proposals are generally consistent with the available scientific information.

## Dr. Sverre Vedal

**April 2005**
**Final comments on PM Staff Paper draft**
**Sverre Vedal**

**Chapter 5 (Recommendations for primary NAAQS)**

Overriding issues:

1. Motivation for revised fine PM NAAQS.

The vast majority of recent epidemiological findings on the association between concentrations of fine PM and an array of health outcome measures have been based on studies in setting where fine PM concentrations were below the current NAAQS. The uncertainties in the epidemiological findings, as reasonably well summarized in the PM Staff Paper (Ch.3 and 5), in my opinion still preclude making solid conclusions as to the causal role of fine PM based on these findings in isolation, although certainly causation is likely. When these epidemiological findings, however, are viewed together with currently available findings from human experimental and toxicological studies, there is enough to allow a judgment that lowering of current standards is sensible and prudent.

2. Recommendations for revised fine PM NAAQS.

a. Importance of the constraints under which the recommendations are made.

The choices and alternatives recommended for consideration as new fine PM standards, as well as the process for arriving at them, reflect the constraints under which these choices were made. Arguably the most significant constraint is the inability to consider costs. Another constraint is the largely informal nature of the risk assessment that, in addition, does not fully incorporate the uncertainties in the concentration-response functions.

The framework for currently recommending lower standards is based on the observation that effects are demonstrable at concentrations substantially below the current standards and on estimated concentration-response functions. Given these, lower standards are recommended by considering both the estimated impacts on health and the number of counties that will be affected by different choices of standards. Such an approach operates at the fuzzy interface between health protection (estimated improvements in health outcomes) and feasibility (the number of counties that would be out of compliance with defined scenarios). In the absence of much guidance as to the threshold concentrations below which no effects occur, it is not clear how the process focuses interest on a range of alternative standards. It would seem naively that continued lowering of the standards would result in further estimated improvements in health, and in more counties not meeting these standards. As long as feasibility is conceivable, such an approach unavoidably results in recommending ever lower standards until background concentrations are reached.

The sensible alternative, but one which EPA staff are constrained from using, is one that attempts to recommend standards in light of both estimated benefits and estimated costs. Although all agree that improvements in health are desirable, these are not ends to be achieved absolutely, *no matter what*. That is, without providing examples, it is not difficult to envision scenarios where the cost of achieving these improvements in health is too high. While the challenges in carrying out a meaningful, and acceptable, cost-benefit analysis of NAAQS are considerable, it should be realized that approaches which do not incorporate such analyses will not make full use of all of the information needed to judge what risk is unacceptable.

The other constraint was the lack of a formal, probabilistic risk assessment that incorporates all of the uncertainties in the concentration-response function. Not incorporating all of these uncertainties in the concentration-response functions results in the ranges around the estimated impacts being too narrow and, in turn, the ranges around the estimated impacts of changes in the standards as well.

b. The *ad hoc* nature of the recommendations.

While the recommended changes to the NAAQS appear to be based on a thorough quantitative assessment, the choices reflect an *ad hoc* approach largely necessitated by the constraints noted above (see above). This is reflected by inexact statements such as in considering an annual standard "*somewhat below* ("my italics) the averages of the long-term concentrations" (5-25, line 20) and "… providing an adequate margin of safety to prevent pollution levels that may pose an unacceptable risk of harm ..." (p. 5-36, line28 and p. 5-55, line 3). How does one determine when risk is unacceptable? Unacceptable risk could be determined either in the political arena (unlikely), with or without the use of cost-benefit analysis, through use of either cost-benefit analysis or cost-effectiveness analysis alone, or through other means, but none of these is at play here. The term therefore has little meaning here and its use only serves to strengthen the impression that the recommendations are, within defined limits, somewhat arbitrary.

However, as noted earlier, given the constraints under which staff recommendations were made, the *ad hoc* approach taken by staff may be as good as can be reasonably done. This is not to say the recommended changes to the NAAQS should not be based on a more firm quantitative foundation, but given the inability to consider costs, and the nature of the risk assessment, this may be about what is realistically possible.

3. Recommendations for coarse PM NAAQS.

It is true that there are some epidemiological, and even some experimental, studies that indicate adverse effects of the coarse fraction. However, I continue to have reservations about the advisability of proposing a coarse PM standard at this time. These reservations are based on two issues. First, the nonspecificity of the coarse PM mass metric, while somewhat of an issue for fine PM, is a much more significant issue for coarse PM. Coarse PM includes crustal PM that evidence indicates has little toxicity in the concentrations of interest here, as well as road dust and biological material that are more worrisome. Coarse PM composition varies from setting to setting. Note on p. 5-62, line 3, it is noted that a short-term coarse standard would afford protection "*in some urban areas*." This is an indication of the type of qualifier that might need to be considered should such a standard be proposed. This nonspecificity of a mass metric for coarse PM with respect to composition would seem to preclude serious consideration of a coarse PM mass-based standard, if used in isolation, that is, without any qualifiers.

Secondly, given the more marked spatial variability of coarse PM concentrations across an urban area, monitor placement becomes a more critical issue than it does for fine PM. What strategy is in place for the optimal siting of coarse PM monitors?

Without specific plans to deal with these two issues, in my opinion they are serious impediments to recommending a standard at this time.

Specific points:

p. 5-14 (Table 5-1). I would find the addition of incidence rates, rather than just absolute incidence, to be helpful. Also, it should be pointed out that the reason the 95% confidence ranges for estimated impacts of long-term exposure are never negative is that, as opposed to the case for short-term exposure, these are all based on only the ACS relative risk estimate whose 95% CI did not cross 1.0. Short-term exposure effects are based on individual city effect estimates where this was not always the case. This is also relevant to the statements beginning on p. 5-45, line 20. Also, as noted in comments on chapter 4, the influence here of hypothetical thresholds on impact estimates is dramatic.

p. 5-16 (line 6). It is stated that in the absence of evidence for thresholds, most weight is placed on estimates that assume none. I don't believe it is more likely that risk extends down to base concentrations than to some higher concentrations, even though these cannot, and possibly never will, be identified. I would therefore favor placing more weight on the impacts of a range of thresholds than on the base case.

p. 5-19 (line 20). It is not clear how a mass indicator controls ultrafines. Further, I would argue against the indication that reduction in sulfates and acids would likely result in most risk reductions, given that it is likely they are not the more toxic components of the fine PM mix.

p. 5-25 (line 14). "Consistency" and "robustness" are overworked terms. The cohort studies are not necessarily consistent if one considers the Veterans study, and the ACS findings are not robust to the effects of $SO_2$. Therefore, the findings are consistent and robust only in a very defined sense. See also p. 5-33, line 25.

p. 5-26. The process described in determining what level of standard should be considered for an annual fine PM standard is indicative of its arbitrariness. The average fine PM concentrations of the initial (21 $\mu g/m^3$) and the recent (14 $\mu g/m^3$) periods used in the ACS studies was used to focus on a range based on the standard deviations of this average. At this time we do not know which time period is most relevant for the effects observed, so I suppose an average is as good as anything, but this is only crude reckoning. Further, the link between average concentrations in any study and what should be recommended as a standard is fuzzy. What is the relationship?

p. 5-28 (Table 5-2). It would also be helpful if the estimates for the 65 $\mu g/m^3$ 24-hr 99$^{th}$ %ile case were also presented for annual standards of 14 through 12 $\mu g/m^3$, and not just for the 15 $\mu g/m^3$ case, unless this concentration is never reached under the lower annual standard scenarios, and therefore adds nothing to the presentation.

p. 5-49, line 24. It seems counterintuitive, although it may be correct, that use of average vs. highest monitor results in larger estimates of incidence.

p. 5-52, line 8. Population-oriented monitors would not typically be viewed as being the most appropriate for providing information that could be used to protect people residing in localized areas of elevated concentrations.

p. 5-63, line 7. One could consider, with justification, that the crustal component of coarse PM could be "eliminated from consideration."

**Chapter 3 (Health effects)**

There are an inordinate number of instances where the summary of findings and their interpretation are overstated, misleading or inaccurate. I detail these here in the order in which they occur:

1. The APHEA revised estimates were in fact reduced by 30%, and therefore not identical (p.3-16, line 21).
2. Bringing in the Hong Kong study on effects of limiting sulfur in fuels as a buttress of the PM time-series studies (p.3-19, line 28) is inappropriate, given that only $SO_2$ concentrations, and not PM concentrations, were reduced by this intervention.
3. While the modifying role of level of education of the PM effect on mortality in the ACS could be characterized, as was done, by stating that those with the lowest level of education showed "larger and more statistically significant" effects (p. 3-21, line 14), it is also true that in those with greater than a high school education the effect was nil.
4. Downplaying the apparently confounding effect of $SO_2$ on the PM-mortality association in ACS by arguing that $SO_2$ is merely part of the causal chain in the formation of sulfate is not credible.
5. The discussion of strength of association (p.3-32) sidesteps the issue by focusing on the strength of the evidence, which is a different matter altogether. The estimates of effect range, in fact, from very weak to relatively weak.
6. The discussion of robustness (p. 3-33) is overly optimistic in its characterization of the impact of model specification, and on co-pollutant confounding. For example, it ignores the $SO_2$ effect in ACS, the study on which the risk assessment is based.
7. There is surprisingly still confusion about confounding vs. effect modification. Much of what is include in the discussion of effect modification (p.3-53) is actually a discussion of confounding:

i. Line 7 describes an approach used by Schwartz to assess the likelihood of confounding, and the terminology of the last sentence of the paragraph confirms this by referring to the lack of dependence of the PM associations on the correlation of PM and the gases.

ii. In the next paragraph, effect modification is addressed by noting that there was no pattern between PM effect estimates and average concentration of gaseous pollutants, which is appropriate (although I doubt this point requires two pages of figures to support it), but confusion is again instilled when the effect of PM is described as being independent (p. 3-53, line 21), a term that is typically applied in discussions of confounding, not effect modification. The above approach, by the way, is only one way of assessing effect modification by gases. Another would be the assessment within a single study of whether the effect of time-varying PM is modified by time-varying concentrations of the gases in the time-series models, through addition of interaction terms, for example. This is typically not done, and perhaps it is not advisable. However, the conclusions regarding gases not modifying the effects of PM may the too strong, since this is little studied.

iii. Bringing in the "transfer of effects" argument from the context of measurement error of two pollutant measures (p. 3-44, line 9) into a discussion of effect modification and confounding (p. 3-53, line 28) is a misappropriation.

iv. Finally, the discussion of pollutants acting together (p. 3-56, line 1 and onwards) is indeed a discussion of effect modification, as is appropriate in this section intended to address effect modification, but the section then ends with a discussion of confounding.

8.  The discussion of temporality (beginning p. 3-56) largely misinterprets what is meant by the term in the context of criteria used to assess the likelihood of causation.  Hill used it to address whether the “horse came before the cart,” or vice versa, and indicates that the notion is most relevant to studies of diseases of “slow development.”  In the air pollution context, this would refer to the setting of cohort studies if, for example, diseased persons tended not to move away from more polluted regions, but healthier persons did – an example of the horse (exposure) coming after the cart (disease).  In the time-series setting, it might refer to increases in pollution concentrations associated with a particular outcome that actually occurred following the outcome in time, which is not plausible, and if observed, would call into question the observed associations.  In the time series or cohort context, investigating this is not straightforward, since concentrations are correlated over time; this might explain why this is not done.

9.  The discussion of lag structure, averaging times, etc. should be included under a different heading than “temporality.”  Although these pertain, in a general sense, to time, they are not relevant to temporality in the context of arguing for causation, which is the context here. Further, the discussion of lags (p. 3-57) is muddled and does not link up well with the risk assessment.  True, use of single day lags typically underestimates effects of distributed lags. However, use of a best single day lag is biased upwards when trying to estimate the effect of a single day.  Since the risk assessment, and the form of the standard, both pertain to single 24-hour periods, the concentration-response function should reflect that.  If these were to make use, somehow, of distributed lags, then it would appropriate to use estimates for the concentration-response functions that are in fact based on them.

10.  The claim that “investigators have reported quantitative results only for the strongest associations, after testing associations over a range of lags and finding a reasonably consistent pattern across lags” (p. 3-57, line 25) is wishful thinking, as is the claim that lags are often chosen *a priori* (p. 3-58, line 7), based on much of my review of the literature.  Best lags are often chosen regardless of consistency of effects, and typically after reviewing the findings.

11.  Regarding seasonal differences (p.3-60), in those uncommon studies in which there actually has been exploration of differences across season, these are generally found to be present. However, I agree with the SP that the lack of consistency precludes using this observation to modify risk assessments.

12.  Based on Figure 3-1 (p. 3-18), I disagree that coarse PM short-term mortality effects should not be considered in the risk assessment (p. 3-67, line 6).  These effects, particularly those on cardiovascular deaths, are at least as impressive as those for fine PM, if not more so.

<u>Minor issues</u>.

p. 3-9, line 5.  Is it really being suggested here that PM can result in such severe impairment in lung function as to cause heart injury, that is, cor pulmonale?

p. 3-15, line 23.  Another explanation for the larger effect estimates compared to the total number of NMMAPS cities is that these are quite a select subset of cities.

p. 3-25, line 12.  Regarding the NMMAPS findings on gaseous confounding of the hospitalization effects, I would be more interested in what the authors of the SP conclude than the authors of the report.  The soundness of this approach to assessing confounding was questioned in CASAC comments on the CD.

p. 3-43.  The discussion in the section, “Air Quality Data in Epidemiological Studies,” includes a great deal that should be instead included in the next section on Exposure Error.

p. 3-51, line 15. Collinearity is most likely due to meteorology influencing the suite of pollutants together, rather than pollutants originating from the same source or being part of the same causal pathway.

Appendix 3A. It is unclear why some of the study descriptions are italicized and what the rationale is for ordering of studies.

**Chapter 4 (Health risks)**

Overriding issues:

Although this chapter is generally well reasoned, and the multiple sensitivity analyses are very helpful, the risk assessment is still not quite adequate. First, the risk analysis is relatively informal in nature, by which I mean that no attempt has been made to take a more probabilistic approach to incorporating the full range of uncertainties in the risk estimates. As noted (p. 4-1, line 22), the probabilistic aspect is manifested only in the statistical uncertainty in the concentration-response functions. The approach to addressing this full range of uncertainties is largely through sensitivity analyses, which is one acceptable, albeit informal, alternative approach. Further, it is not clear why a more comprehensive integrated uncertainty assessment would necessarily require expert probability judgments (p. 4-2, line 4), although that is one approach to take. Perhaps expert judgments are required in order to obtain priors for a Bayesian uncertainty analysis, but my inadequate background in this area doesn't allow me to go further here. If such judgments are required, then there seems to have been adequate time to have them obtained.

Secondly, not all sources of uncertainty are addressed. Two that pertain to the concentration-response functions are model selection and publication bias. I would recommend adding these to the discussion of the "empirically estimated C-R relationships" (p.4-38). These omissions are particularly acute for short-term exposure effect estimates.

I am not convinced that short-term coarse PM effects on mortality should not be considered in the risk assessment. See Figure 3-1 (p. 3-18) where total mortality effects are not much different than for fine PM, and cardiovascular mortality effects are arguably more convincingly present than for fine PM.

A small point that may require elaboration is whether the risk assessment needs to take account of different scenarios of PM concentrations over time. Specifically, I presume that settings where, to take two extreme examples, PM concentrations remain unchanged at a fixed level above baseline are treated the same as those where there is marked fluctuation in daily PM concentrations, but where the cumulative increases above baseline are the same as in the first case. Strictly, the interpretation of the model coefficients based on time-series studies is that of an estimated change in outcome for a given *change* in PM concentration. This is relevant to the discussion of estimation of incidence (pp. 4-24 & 25).

The discussion of how lags were chosen for use in the C-R functions is not clear (pp. 4-35 to 37). Did the SP authors select lags from the original papers, or were the "best" lags used when these were identified as such by the authors of the papers? It sounds as if 0 and 1 day lags were used in all instances except when the original authors stipulated otherwise. If so, I find this approach to be unsound and one that introduces yet another element of uncertainty into the C-R functions.

Some of the findings of the risk assessment sensitivity analysis should motivate intensive future research. First is the dramatic impact of hypothesized thresholds (p. 4-57, etc.). The second is the effect of model specification (p. 4-56, first paragraph on the Moolgavkar analyses).

Minor issues:

p. 4-35, line 3 and further. See my discussion of confusing confounding and effect modification, which is equally relevant here.

p. 4-52 and 4-59. What do the superscripts 18 and 19, respectively, refer to, or are these vestiges from an earlier version?

p. 4-61, Figure 4-9. What are the two separate estimates for St. Louis?

p. 4-62, Figure 4-10. This also shows impacts of alternative standards (not just current standards), in contrast to what the legend indicates. Also, a range of 24-hr concentration standards is presented, again in contrast to the legend, and this range does not include a concentration of 65 $\mu g/m^3$. Note that Figure 8-3 in the Abt report has the correct legend.

p. 4-67, line 23 and table 4-13 and table 4-15. The "design value" terminology is not intuitive (even after a review of the Schmidt document) and needs a very brief explanation.

**Chapter 7 (Recommendations on secondary PM NAAQS)**

I have no specific comments to make about this chapter apart from echoing a point made by some of my colleagues on the CASAC encouraging a future change from the criteria pollutant approach to a critical load approach.

# Mr. Ronald H. White

Revised Comments of Ronald White on Chapter 3
EPA Particulate Matter Staff Paper – Second Draft
April 12, 2005

## General Comments

Overall, the discussion of the health effects literature in Chapter 3 accurately interprets the information presented in the October 2004 PM Criteria Document. The discussion and interpretation of the results from the health studies is balanced, if anything somewhat overly cautious and conservative in tone, and generally clearly stated. The discussion of potential factors that can affect the interpretation of quantitative results, such as measurement error, model specification and lags, is also generally well done. The discussion of the effect of co-pollutants in interpreting the results of PM health studies would benefit from a clearer discussion of EPA's approach to interpreting quantitative results from multi-pollutant studies.

## Specific Comments

Pg. 3-6, line 7: Breathing patterns are appropriately raised here as affecting particle deposition, but then are not discussed. Some mention of the effect of activity state on particle deposition would be appropriate in this paragraph.

Pg. 3-40, line 26: The correct reference to the conclusions in the Criteria Document of the discussion of fetal/infant health effects is pg. 8-222, not 8-335 which discusses the implication of infant mortality on life-shortening estimates. Also, a more accurate paraphrasing of the CD's review and conclusions on this topic is that results from this emerging limited body of research, though mixed, are suggestive of a possible PM effect and more research is needed to further elucidate the potential risks from PM exposure for these health outcomes.

Pg. 3-41, lines 8–10: The summary statement on susceptible and vulnerable populations omits any reference to the discussion in the previous paragraph of low SES and more highly exposed vulnerable populations. At a minimum, the summary conclusions for this section should acknowledge the emerging though more limited evidence for increased vulnerability for these populations.

Pg. 3-42, line 10: The conclusion in the Criteria Document regarding the hypothesis of a harvesting effect from the time series studies analyses is on pg. 8-334, not pg. 8-329.

Pg. 3-59, lines 25-27: It would be appropriate to note here that the Second Draft Health Risk Assessment does include quantitative estimates of short-term $PM_{2.5}$ health effects from the "as is" scenario using distributed lag models as part of the sensitivity analyses.

Pg. 3-61, lines 27-28: change "PM" to $PM_{2.5}$

Revised Comments of Ronald White on Chapter 4
EPA Particulate Matter Staff Paper – Second Draft
April 12, 2005

General Comments

The selection of health endpoints, epidemiologic studies and concentration-response functions for $PM_{2.5}$ and $PM_{10-2.5}$ is appropriate and well supported. However, the discussion in Section 4.5.2, which describes results from analyses of alternative $PM_{10-2.5}$ standard levels, is somewhat confusing in sections. For example, the discussion of the health protectiveness of alternative $PM_{10-2.5}$ standard levels (pgs. 4-74 to 4-75) comparing design values to 98$^{th}$ and 99$^{th}$ percentile concentrations lacks clarity as currently written.

I concur with the SP and CD assessment that the weight of evidence from the $PM_{10-2.5}$ short-term mortality health literature currently is significantly less compelling than for $PM_{2.5}$. However, the CD discusses several studies that found statistically significant associations between short-term $PM_{10-2.5}$ exposure and cardiovascular mortality (pg. 8-303) and notes that while the associations of short-term $PM_{10-2.5}$ exposure and respiratory mortality generally do not reach statistical significance and have broader confidence intervals than for cardiovascular mortality (which may well reflect the issue of low study power due to the small number of respiratory mortality cases in these single city studies), "…the findings may well reflect actual associations between mortality and $PM_{10-2.5}$, at least in some locations" (CD, pg. 8-304). Given this information, inclusion in the Risk Assessment of short-term mortality associated with "as is" $PM_{10-2.5}$ levels and potential mortality reductions due to attainment of the proposed $PM_{10-2.5}$ NAAQS alternatives would provide the Administrator with a more complete picture of the potential health impacts related to PM exposure. These mortality estimates would need to be accompanied by appropriate language indicating the increased uncertainty associated with the $PM_{10-2.5}$ mortality estimates in comparison to the $PM_{2.5}$ estimates.

The risk assessment and analysis of the potential health impacts of the "as is", attainment of current $PM_{2.5}$ and alternative standards scenarios has been well done. Both the RA and SP adequately identify and describe the uncertainties associated with the risk assessment, and in general the sensitivity analyses are well done and informative regarding the impact (or lack thereof) on the RA results from the uncertainties that were examined. However, given the importance of the threshold uncertainty issue to the risk assessment results, a more detailed probabilistic analysis using air quality and health data of the quantitative implications of various threshold values using air quality and health data would have been particularly helpful to more definitively assess this uncertainty. Given that such an analysis if started now would not be completed in time to inform this current PM NAAQS review, EPA should consider undertaking such an analysis in a timeframe that allows for the results to inform the next PM NAAQS review. It might be also be helpful to provide some indication of the impact on the results of the quantitative risk assessment from integration of the sensitivity analyses that address what EPA considers as the key uncertainties.

I share the concern expressed in Dr. Lioy's comments on Chapter 4 of the SP regarding interpretation of Figures 4-8a and 4-8b, and concur with his suggestion that the more interesting analysis would be to present the daily mortality as a function of concentration-days. Based on the presumption of a linear, no-threshold concentration-response function, it is not surprising that the largest number of deaths occur at the concentrations with the largest number of days at those monitored values. It would appear from visual inspection of these two figures that the number of deaths per concentration-day increases substantially in comparing values between approximately 10 to 50 ug/m$^3$, which as Dr. Lioy suggests might reflect differences in particle composition and therefore toxicity, population exposure, or some combination of the two. As Dr. Lioy also notes, this observation from the Detroit data may or may not be representative of the national picture, but given the implications for development of control strategies further analysis of this approach for other areas of the U.S. might prove informative before concluding that higher concentration levels are relatively unimportant for mortality-related health risks. It should be also be noted that the overall aggregate contribution of higher $PM_{2.5}$ levels to the mortality burden would be more substantial if the C-R model assumed a threshold for short-term mortality, especially at the higher cutpoint values used in the sensitivity analysis.

Specific Comments

Pg.4-16, line 20: Unlike the reference to a threshold for study precision related to mortality studies (≥9 natural log of mortality-days), the reference to the criteria for "greater precision" related to morbidity studies not clear either in the SP of RA TSD. If the same precision threshold was used for the morbidity as well as mortality studies, it should be so stated.

Pg. 4-25, line 11: Since this seems to be the first use of the term "policy relevant background" in this chapter, spell out "PRB" and indicate the acronym. The distinction between PRB and the frequently used term "estimated background level" should be clarified if one exists. If they are synonymous, then a single term should be used consistently to avoid confusion.

Pg. 4-58, Figure 4-8b: The mortality data should be presented using the more standard 5$^{th}$ and 95$^{th}$ percentile values, or some explanation for use of the 2.5$^{th}$ and 97.5$^{th}$ percentiles should be provided.

Pg. 4-69, Table 4-13: It should be noted that the annual design values are based on the maximum monitor values rather than the average of monitor-specific annual averages.

Revised Comments of Ronald White on Chapter 5
EPA Particulate Matter Staff Paper – Second Draft
April 12, 2005

General Comments

I agree with the approach taken in SP of emphasizing the risk assessment results for $PM_{2.5}$ in developing the alternative suite of standards for the annual and 24-hour $PM_{2.5}$ NAAQS, and concur with placing greater emphasis on the results of the health studies of $PM_{10-2.5}$ rather than the risk assessment in developing the alternative NAAQS for $PM_{10-2.5}$. However, as indicated in my comments on Chapter 4, I recommend that appropriately caveated quantitative short-term mortality risk assessment estimates be developed for $PM_{10-2.5}$ to improve the scope of information available to the Administrator in setting a thoracic coarse particle standard, or at a minimum should be included as a sensitivity analysis for the $PM_{10-2.5}$ risk assessment.

I agree that the suite of standards developed for fine and thoracic coarse particles are generally consistent with the available scientific information, and are appropriate for consideration by the Administrator. Based on my review of the scientific information, I support the SP finding that the current annual and 24-hour $PM_{2.5}$ NAAQS are insufficient to protect public health, and support a revised annual average $PM_{2.5}$ standard in the range of 12-14 ug/m$^3$ and a revised 24-hour standard in the range of 25-35 ug/m$^3$. In selecting revised $PM_{2.5}$ NAAQS from these ranges, the Administrator should consider the following factors: 1) the amount of public health protection provided by the combination of the annual and 24-hour $PM_{2.5}$ NAAQS; 2) whether the 98$^{th}$ or 99$^{th}$ percentile form of the 24-hour NAAQS is selected, recognizing that the trade-off between improved stability for compliance designations provided by selection of the 98$^{th}$ percentile form of the NAAQS results in less public health protection in the number of days allowed to exceed the NAAQS. If the 98$^{th}$ percentile form of the standard is selected, it would be appropriate to select a numeric value for the 24-hour NAAQS that is 5 ug/m$^3$ lower than would have been selected for a 99$^{th}$ percentile form of the standard; 3) the need for the 24-hour standard to provide some additional margin of protection against very short-term (1-3 hour) peak exposures that preliminary evidence suggests may be an important health concern, especially for cardiovascular effects.

While the amount of scientific evidence supporting a $PM_{10-2.5}$ NAAQS is substantially smaller than for $PM_{2.5}$, the available evidence for morbidity effects, and to a lessor extent mortality effects, supports the need for a $PM_{10-2.5}$ NAAQS to protect public health against 24-hour coarse particle exposures. Currently, there is not sufficient evidence to support establishment of an annual average $PM_{10-2.5}$ NAAQS. As adverse health effects have been found at levels well below the current $PM_{10}$ NAAQS in studies conducted in locations where a substantial portion of the $PM_{10}$ fraction is in the coarse mode, and a more limited number of $PM_{10-2.5}$ studies conducted by necessity primarily in urban areas have found adverse health effects from coarse particle exposures at levels well below the equivalent $PM_{10}$ values, I recommend that the Administrator consider selection of a 24-hour $PM_{10-2.5}$ NAAQS at a level below the equivalent 24-hour $PM_{10}$ value of approximately 70 – 75 ug/m$^3$. As is the case with the $PM_{2.5}$ NAAQS, the Administrator's decision on selection of a 24-hour NAAQS value should consider the need to trade-off the compliance stability afforded by selection of a 98$^{th}$ percentile form of the standard versus the additional health protection provided by a 99$^{th}$ percentile form. In addition, given that the majority of health evidence related to coarse particles is based on studies conducted in urban areas, EPA's implementation policy for the $PM_{10\text{-}2.5}$ NAAQS should likewise focus on urban areas.

Specific Comments

Pg. 5-22, line 32 to pg. 5-23, line 2: The suggestion that a significant harm program and/or the AQI program would be an effective mechanism for protecting public health from short-term (1-3 hour) peak PM concentrations does not seem realistic given the current limitations of these programs and the inherent delays involved with public information dissemination and public response.

# Dr. Warren H. White

**Comments on Chapter 2: Characterization of Ambient PM**
Warren H. White

**General**

Chapter 2 is excellent. It exhibits a sophisticated and nuanced understanding of the CD and underlying research base. The only gap I would highlight concerns phase partitioning and the relationship of collected sample to ambient particle concentrations and delivered doses. Whereas the distinction between different-sized particles receives considerable attention, the problems of distinguishing gaseous from condensed phase semivolatiles (water, ammonium nitrate, many organic compounds) are acknowledged only in passing. In particular, no foundation is laid for preferring the sampling losses prescribed by the Federal Reference Method over alternatives such as explicit characterization of volatilization (e.g. FDMS) or volatile species (e.g. using denuders for nitrates), or truly *in situ* measurements (e.g. nephelometry).

**Specific comments**

Footnote 2, page 2-4, and Y-axis labels in Figure 2-1: I applaud the effort to explain how particle- size distributions are plotted, but this explanation isn't quite right. It's not the "measured concentration difference" on the Y-axis, but instead the "measured concentration difference per logarithmic increment in particle diameter, $[F(D_p+\Delta D_p)-F(Dp)]/[\log(D_p+\Delta D_p)-\log(D_p)]$, where $F(D_p)$ is the cumulative concentration (in counts, surface area, or volume, per $cm^3$ air) of all particles with diameters less than $D_p$". Correspondingly, the Y labels in the figure should be $dN/d\log D_p$ (or $\Delta N/\Delta \log D_p$) and so on, rather than $N/\log D_p$. Note that the Y label in Figure 2-2 is <u>correct.</u>

Line 26, page 2-6: I don't find any real discussion of measurement methods for $PM_{10\text{-}2.5}$ in section 2.3.

Lines 12-25, page 2-10: This paragraph is a little muddled. Accumulation-mode particles <u>don't</u> "remain suspended longer [than ultrafine particles, the subject of the previous 3 sentences] due to collisions with air molecules". They are not any better supported by collisions with air molecules than ultrafine particles are. Particles in the 0.1 – 2.5 μm range remain suspended – and thus accumulate! – because they are too large to diffuse rapidly to surfaces and other particles, and too small to settle out or impact on stationary objects. Similarly, ultrafine particles don't simply "grow rapidly into the accumulation mode" – they are more likely to reach the accumulation mode by coagulating with a particle that is already in that size range.

Line 1, page 2-11: The concept of "intercontinental dust storms" is an interesting one. Dust storms over the oceans? The cited CD passage is "when mixed high into the atmosphere, as in dust storms"; even a 100m-high dust devil is enough to suspend a $PM_{10\text{-}2.5}$ particle for several hours.

Lines 15-17, page 2-11: It is true that elemental carbon is a relatively small component of PM in most areas, and that absorption therefore contributes less than scattering to extinction. But that is not why scattering dominates visibility impairment, which is caused more by extraneous light scattered into the sight path by the intervening atmosphere ("airlight") than by attenuation of transmitted light from the target. One need only consider the view through sunglasses; these absorb strongly, do not scatter (unless they are scratched), and would not be worn if they impaired visibility as haze does.

Line 31, page 2-11, through line 2, page 2-13: It's hard to see how "the radiative properties of the particles … are dependent on … their vertical and horizontal distribution in the lower atmosphere."

Lines 10-12, page 2-13: Has it been established that "absorption of … outgoing [IR] terrestrial radiation by particles" is "primarily" attributable to elemental carbon rather than mineral dusts?

Lines 13-14, page 2-13: A more accurate topic sentence for this paragraph would be "The mix of scattering and absorption by ambient particles is .." or "The relative proportions of scattering and absorption …". The absolute "extent" (or amount) of scattering and extinction is not all that dependent on composition, as the main accumulation-mode species all have comparable refractive indices and densities.

Line 17, page 2-13: "degree of reflectivity" should be "refractive index" or "single-scattering albedo."

Line 2, Page 2-15**:** This is not 100% true. The coarse particles sample contains a large fraction of fine particles. [10% of the original air sample which can be up to 30-40% of the coarse particle mass].
Lines 1-2, page 2-15: Would be clearer and more accurate as "… two streams so that fine-fraction and coarse-fraction-enriched particles can be collected on separate filters."

Line 11, page 2-15: The claim that "the $PM_{2.5}$ FRM has been a robust indicator of ambient levels by meeting the data quality objectives" is likely to mislead unwary readers. The FRM is a robust measurement in the sense of being repeatable, and the FRM sample depends on ambient concentrations, thereby serving as an indicator of them. But the performance of the FRM *as an indicator* – that is, *its relationship to ambient levels* – is simply not something that the cited QA data address.

Page 2-16: Why is increased pressure drop (CAMM) considered a "mass" measurement when light scattering (nephelometer) is considered an "indirect optical" measurement? Pressure drop has nothing directly to do with gravimetric or inertial mass; like light scattering, it correlates principally with particles' aggregate cross-sectional area.

Lines 23-24, page 2-16: Should read "Nephelometers measure the light scattered by ambient aerosols *as the principal component of* light extinction."

Lines 21-23, page 2-17: The "several methods" for distinguishing OC and EC are not distinct "methods" in the way XRF and IC are distinct methods. They are variations of the same method, just as the XRF and IC performed at different labs for different networks employ distinct variations of a common technique. The difference between carbon analysis and XRF or IC is that the OC/EC distinction is sensitive to, and *defined in terms of*, details of the method. I suggest the following revision, starting with a sentence added to the end of the first paragraph.

> … ($NA^+$), organic cations (such as acetate), and phosphate ($PO_4^{3-}$). Particulate carbon is first oxidized to $CO_2$ and then reduced, for measurement as $CH_4$.
>
> Organic carbon (OC) and elemental carbon (EC) are distinguished by the temperatures at which they evolve, according to several different protocols of heating, oxygen availability, and correction for pyrolysis. Thermal optical reflectance (TOR), …

Line 23, page 2-18: "There is no perfect PM sampler under all conditions" is a tautology, and hardly needs stating. I suggest a more informative statement of the problem: "The partitioning of material between particle and gas phases is sensitive to the micro-environment".

Lines 25-27, page 2-18: Any SRM that could be used "to estimate the accuracy … relative to what is found in air" would itself have to be an aerosol. (That is, particles suspended at known concentration in a gas.) Does staff know of <u>any</u> such SRM?

Line 26, page 2-19: What's "unique" about blowing dust?

Line 21, page 2-22: The EPA's exceptional events guideline cited in footnote 22 specifically <u>excludes</u> meteorological inversions from designation as exceptional events: "Because inversions are expected to occur frequently and are part of weather patterns, they are not considered exceptional events for the purpose of flagging data.

Figure 2-9: The Y-axis needs a label.

Lines 17-18, page 2-46: The interpretation of $PM_{10-2.5}$ variability was properly "tempered" in line 22, page 2-39 by noting the importance of measurement uncertainty. The same caveat should be added here.

Lines 13-15, page 2-50: These are redundant. I suggest "…when the abundant $SO_2$ emissions there are rapidly converted to sulfates by increased photochemical activity."

Figures 2-21, 2-22: The Y-axes need labels.

Lines 29-30, page 2-55; Figure 2-28: There is something terribly wrong with this case study. (a) The $PM_{10-2.5}$ scale in Figure 2-28 is exactly 1/6$^{th}$ the $PM_{2.5}$ scale, not "about 3 times as large". (b) $PM_{2.5}$ is shown approaching 3000 $\mu g/m^3$, more than 6 times $PM_{10-2.5}$. Let us suppose the two plots of Figure 2-28 are interchanged. In that case we have a series of source-oriented measurements that clearly show the ratio of fine to coarse dust to be about 1 to 6 in emissions that must have been quite fresh (to be so concentrated). A coarse concentration of 30 $\mu g/m^3$

would not be unusual in the west, and with a little aging we could expect the fine/coarse ratio to increase as larger particles are lost. Therefore, (c) Figure 2-28 (with labels interchanged) implies that significant $PM_{2.5}$ dust increments, exceeding 5 $\mu g/m^3$, can be expected to accompany unremarkable $PM_{10-2.5}$ dust concentrations (e.g. 30 $\mu g/m^3$). This is not quite the picture painted earlier in the chapter, which tends to emphasize the separation afforded by the 2.5 µm size cut. And finally, (d) this activity starts in the evening, peaks around midnight, and is over by 3 or 4 am; is this properly considered a "dust storm", which the American Heritage dictionary defines as "a severe windstorm that sweeps clouds of dust across an extensive area"?

Lines 3-5, page 2-60: These are rather non-obvious statistics. The "site-level median hourly increases" are presumably "hourly increases in site-level hourly medians", and the "average median increase" must actually be something like the "average site-level maximum hourly increase in hourly medians." For any diurnal ***cycle***, all increases are going to average out to zero! In any case, I'm not clear what useful information is conveyed by these numbers.

Lines 24-25, page 2-73: Extinction can't be "calculated" from available pollutant concentrations, it can only be "estimated".

Line 6, page 2-74: Change to "Malm (2000) developed an algorithm for calculating …" or "EPA guidance for tracking progress under the regional haze rule specifies an algorithm for calculating …". This formula was developed by the National Park Service as a standard format for reporting the optical implications of IMPROVE data, and adopted by EPA in its guidance on the Regional Haze Rule, which rests on IMPROVE data. It is a convention developed for reporting and regulatory purposes, with known biases and oversimplifications, and is neither a product nor the best representation of "The IMPROVE visibility monitoring program".

---

**Comments on SP chapter 6**
Warren H. White

I heartily applaud EPA staff's analysis of the desirability and feasibility of a secondary PM2.5 standard related to visibility. It is thoughtful and responsive to previous CASAC reviews, and is solidly grounded in well-focused explorations of the abundant data newly available from the Agency's speciation monitors. Rich Poirot has written a masterly review, and I have only some editorial comments to add.

My main comment is that section 6.2.1, Overview of Visibility Impairment, seems both pedantic and confused, and not a very good introduction to such an exciting topic. It is a dry taxonomic discussion, and I don't see how it really informs anything that follows it. What, exactly, is the distinction intended between "local" and "regional" in the introductory sentence (lines 30-31 on page 6-2)? Is it

a) geographic scale, which affects whether a problem can be dealt with at the city or state levels, or must be kicked up to RPOs?

b) impacts from identifiable sources ("reasonably attributable") *vs.* accumulations of indistinguishable increments, which may come under different regulatory mechanisms?

c) bounded regions (plumes or layers) external to the observer *vs.* diffuse regions engulfing the observer, which require different measurement and modeling approaches?

These are not three versions of the same distinction: for example, most urban "brown clouds", which line 3 of page 6-3 identifies as "localized haze", are not "reasonably attributable". All three distinctions are worthy of discussion in an orientation to visibility impairment, as substantive and inter-related factors that shape our characterization and regulation of haze. The present section instead treats the taxonomic scheme itself as fundamental, and presents these distinctions more as unexamined attributes that help us decide what goes where.

Line 8, page 6-4: IMPROVE does <u>not</u> measure fine-particle precursors.

Line 25, page 6-5: When does urban visibility impairment <u>not</u> result from the combined effect of stationary, mobile, and area source emissions? The qualifier "often" is superfluous here.

Line 1, 6-16: I think you mean to describe road and air safety as "use" rather than "aesthetic" benefits.

Lines 2-3, 6-22: The introductory sentence needs to make clear that this claim applies only in the context of surveys, if then. It's certainly not true in general that "the principal method for recording and describing visual air quality has been through 35 millimeter photographs"!

Lines 10-13, 6-25: It doesn't make sense to tell us that South Mountain is 40 km away and then give the visibility only in deciviews.

Line 25, page 6-65: Why "especially sulfates"?

Line 28, page 6-65: Is there good reason to emphasize black carbon over mineral dust as an absorber, especially in a context that includes outgoing terrestrial radiation, which is long-wave? I don't see support for this in the CD. The emphasis is repeated in line 23, page 6-66.

Line 4, page 6-66, line 19, page 6-67: The CD references are incomplete: *e.g.*, "p. 216" should be "p. 4-216".

# Dr. George T. Wolff

**Comments on the January, 2005 PM Staff Paper**

by

George T. Wolff
(3/2005)

**Specific Comments**

1. p. 2-8, line 31 – Insert "gaseous" in front of "ammonia."
2. p. 2-9, line 2 - Insert "gaseous" in front of "ammonia."
3. p. 2-9, line 3 – Insert "gaseous" in front of "volatile organic compounds."
4. p. 2-50, lines 12 – 26 and Figures 2-21 and 2-22 – It appears that the graphs are true annual quarters rather than climatological seasons. Because of this, true seasonal patterns are likely to be somewhat obscured. It would be more climatologically meaningful if Dec. – Feb., Mar. – May, Jun. – Aug. and Sept. – Nov. values were plotted.
5. p. 2-55, lines 29-30 – "Note that the $PM_{10-2.5}$ scale is about 3 times as large as the $PM_{2.5}$ scale (in Figure 2-28)." Something is wrong here as the $PM_{2.5}$ scale in Figure 2-28 is 6 times larger than the coarse scale.
6. p. 2-58, Figure 2-27 – April 26 – 27, 2002 must have had a significant influence on the annual diurnal average. I am curious to see what this figure would look like if those two days were removed. The observed diurnal pattern may be different.
7. p. 2-60, Section 2-6 "PM Background Levels" – The definition of PRB as stated in the Staff Paper may have been useful in the scientific discussions in the CD, but it is inappropriate for the more policy oriented discussions in the Staff Paper. For the SP, consideration should be given to the background that the U.S. cannot control. We cannot control what is transported into the U.S. from Canada or Mexico. It is naïve to assume that "international agreements" will eliminate all anthropogenic emissions from these countries. U.S. PRB should include PM that is transported in from Canada and Mexico.
8. p. 2-63, lines 8-10 – Not counting any sulfate as part of the PRB background is wrong and biases the PRB too low. Transported sulfate from Canada, Mexico and elsewhere as well as from natural sources should be part of the PRB. In 1981, I estimated that the background sulfate concentration in the U.S. ranged from 0.5 to 1.9 $\mu g/m^3$ based on my data (Ferman, Wolff and Kelly, *J. Air Pollut. Control Assoc.* vol. 31, pp. 1074-1082, 1981) and data from others cited in the paper. In 1982 and 1983, I measured sulfate concentrations at a remote cite in Bermuda of 1.1 and 1.2 $\mu g/m^3$ when the wind was blowing from the SE and SW, respectively (Wolff et al., *Atmos. Environ.* vol. 20, pp. 1229-1239, 1986). Note these are sulfate measurements as sulfate. To incorporate the mass due to ammonium, an appropriate multiplier is needed.
9. More on Background Levels – Based on the above two discussions, I obviously think EPA has underestimated the annual $PM_{2.5}$ and $PM_{10-2.5}$ PRB levels. For a more rigorous approach to estimating the background, I refer the Staff to the well thought out discussions submitted to EPA in October 2003 by George Hidy (EMA comments) and

David Chock (Ford comments). Based on their work, it is obvious that the upper range of the PRB exceeds 12µg/m$^3$. The best way to estimate the background would be to look at the rural sites on the U.S boarders and on the West Coast and average the concentrations on those days when the trajectories came into the U.S. with one caveat. Days with precipitation should be excluded because scavenging will create periods when the PM concentrations are below background.

10. Final Comments on Background Levels – I suggest that EPA read Chock and Hidy's comments on the 24-hour background comments as well. The distributions of these are likely underestimated as well.
11. p. 2-66, lines 25-29 – Indoor concentrations are also affected by meteorological variables such as wind speed and temperature.
12. p. 2-69, lines 26 & 27 – Also wind speed.
13. p. 2-71, lines 1 – 8 – This explanation is pure speculation and should be stated as such.
14. p. 2-76, lines 23-25 – I believe these natural visual range estimates are too high in the east, especially in the warmer seasons when natural VOCs will react to form organic PM. I estimated the natural extinction in the east to be 58 – 90 Mm$^{-1}$ (see Ferman, Wolff and Kelly, *J. Air Pollut. Control Assoc.* vol. 31, pp. 1074-1082, 1981). A 150 km visual range would not allow for a natural blue haze.
15. p. 3-1, line 18 – It should be stated that these are EPA's conclusions because not all CASAC members, at the time, agreed with conclusion #3.
16. p. 3-3, lines 9-34 – For balance another bullet or two should be added on the new issue of model selectivity. I recommend that the last bullet on page 269 of the HEI Special Report, "Revised Analyses of Time-Series Studies of Air Pollution and Health" be paraphrased.
17. p. 3-14, line 1 – It should be mentioned that many of the more than 80 new time-series studies are not interpretable because of the GAM issue.
18. p. 3-17, line 2 – How does EPA get ~1.0 as the lower limit when Figure 3-1 shows some estimates below zero?
19. p. 3-19, lines 25-30 – If the effects of individual pollutants could not be distinguished, how does this support PM time-series studies?
20. p. 3-21, lines 13-15 – As written, this implies that effect on higher educated people was just a little smaller and a little less significant, when in fact it was not significant at al.
21. p. 3-33, lines 22-26 – This is in conflict with the HEI commentary in "Revised Analyses of Time-Series Studies of Air Pollution and Health."
22. p. 3-35 – lines 9-11 – This is misleading because the HEI report (Revised Analyses of Time-Series Studies of Air Pollution and Health) on page 69 says: "This does not prove its (heterogeneity) absence, however, because power of the test to detect heterogeneity is limited…."
23. p. 3-35, lines 11-14 – This sentence is also misleading because it implies the initial study findings have credibility when, in fact, they used an inappropriate convergence criterion.
24. p. 3-42, lines 22-26 – This paraphrase of the 1996 CD is stated much more strongly than the original statement on page 13-92. For accuracy, it should be restated to reflect what was said in the 1996 CD.
25. p. 3-50, lines 1-7 – It should be mentioned that inclusion of $SO_2$ made the PM become insignificant.

26. p. 3-51, lines 23-30 – It should be mentioned that the NMMAPS multi-pollutant modeling approach was not definitive because of the different number of cities used in each of the multipollutant comparisons. It was apples to oranges comparisons.
27. p. 3-52, lines 1-5 – Staff fails to mentions that the majority of the studies shown in Figures 8-16 to 8-19 in the PM CD are not statistically significant.
28. p. 4-46 and 4-47, Figures 4-5 and 4-6– These figures need to include the AHSMOG and VA studies. EPA can explain why they give them less weight, but they should not be ignored. They should be part of the risk assessment.
29. Comment 27 holds for the entire risk assessment.
30. p. 5-10, lines 1-4 – This is cherry picking without justification.
31. p. 5-11, lines 28-29, and first 3 words on p. 512 – That is not what the HEI Review Panel said. They said they do not know how to select the most appropriate model.
32. p. 5-25, lines 14-15 – Cherry picking again.
33. p. 5-33, line 25 – This statement about being robust to alternative modeling is not supported since most of the studies only tried one or two approaches. Most did not carry out a systematic evaluation. Those that did found some significant differences.
34. p. 5-52, lines 5-7 – It is precisely for this reason that the 99$^{th}$ percentile should not be considered because it is an extreme value statistic that is unstable and is not robust.
35. p. 5-67, lines 21-23 – It is inappropriate to use $PM_{10-2.5}$ data from Windsor to estimate exposure in Southeast Michigan. During the summer of 1981, I operated four dichotomous samplers with size cuts at 15 and 3.5 μm in Southeast Michigan (Wolff et al., Atmos. *Environ.*, vol. 19, pp. 305-313, 1982). They were located in downtown Detroit, Dearborn, Warren and Ann Arbor. The best intersite $r^2$ for $PM_{10-2.5}$ was only 0.17. The $PM_{15-3.5}$ concentrations at all of the sites were dominated by very local sources, and one would expect similar behavior for $PM_{10-2.5}$. The measurement error using Windsor PM should make anyone suspicious of any epidemiological study reporting relationship between coarse Windsor PM and hospital admissions in SE Michigan.
36. p. 6-66, line 26 – I would replace “sulfate” with “fine” since CN is not restricted to sulfates.
37. p. 6-67, line 3 – I would delete “major” because it does not need to be major to be an important effect.
38. p. 7-5, line 25 – As I said in comment 13, I think 150 km is much too high for the East especially the photochemical season.
39. p. 7-12, lines 3-24 – Using 24-hour background levels to characterize 4- or 8-hour background levels is as inappropriate as using annual mean background levels to characterize 24-hour levels. The shorter the sampling period, the greater the variability of the data. In the absence of any 4 to 8-hour data, it is going to be extremely difficult to estimate reasonable distributions of appropriate levels.

**Responses to CASAC Charges**

Chapter 2

1. The characterizations have been clearly communicated, appropriately characterized (except for the background levels), and relevant to the review.
2. Appropriate distinctions have been made.
3. The only deficiencies I see is the in the definition of policy relevant background (PRB) and in the estimation of the PRB levels. See my specific comments 7 through 10. The biggest deficiency is that the chapter provides no guidance on what 4-hour or 8-hour PRB levels are.

Chapter 3, 4, and 5

1. I was very pleased to see that the data are no longer truncated at zero. This in itself contributes to a more balanced presentation. I have numerous additional comments on how to make it more balanced in my specific comments.
2. In reviewing the $PM_{2.5}$ and the $PM_{10-2.5}$ results in Figures 3-1 and 3-2, I cannot perceive significant differences. Yet the SP concludes differently. For mortality, 5 out of 28 studies are statistically significant for $PM_{2.5}$ while only 1 out of 18. For admissions and ER visits 7 of 16 are statistically significant for $PM_{2.5}$ while 2 of 10 are significant for $PM_{10-2.5}$. Under the circumstances, I think the statement that "$PM_{2.5}$ is likely causally associated with mortality" is too strong. Evidence for any causal relationship for $PM_{10-2.5}$ is even weaker.

   The influence of alternative model specification is inadequately addressed in this chapter.

3. I agree with the admission that we cannot prove or disprove the existence of a threshold and I applaud the inclusion of thresholds in the subsequent risk assessment.
4. A major shortcoming in chapter 4 is the exclusion of the VA and AHSMOG studies from Figures 4-5 and 4-6. A "balanced" presentation would include them.

   Staff's decision to not include mortality as a short-term endpoint of $PM_{10-2.5}$ is a sound one.

   Uncertainties associated with alternative model specifications are not adequately considered. Uncertainties associated with the long-term studies are ignored by excluding the VA and AHSMOG results.

   I thought the sensitivity analyses were fairly comprehensive with the exception of the long-term studies because of the above-mentioned exclusion.

5. I agree with the approach. Placing less reliance on the $PM_{10-2.5}$ risk assessment results is justified.
6. I agree with Staff's selection of the indicators and averaging times. I agree with the form of the annual standard, but I do not agree that a 99$^{th}$ percentile should be considered. The

99th percentile is an extreme value and it is not robust. The reasons cited in the 1996 for the selection of the 98th percentile are still valid today.

I do not agree with any of the ranges selected. The uncertainty of the assumption of causality is underscored by the fact that the vast majority of the risk estimates for $PM_{2.5}$ and $PM_{10\text{-}2.5}$ shown in Figures 3-1 and 3-2 are not statistically significant. In addition, 7-years of toxicology studies have failed to produce any plausible biological mechanisms only speculations. Consequently, I do not even think an annual standard of $15\mu g/m^3$ is defensible.

7. For the reasons stated in 6, the lower end of the range is not justified.

Chapters 6 and 7

1. I feel the welfare effects evidence is technically sound with the exception of the background visibility for the eastern U.S.

2. I think the methodologies are technically sound.

3. No. I do not think the state and local programs are appropriate to use in setting national standards. This is especially true in the East where many urban areas have no spectacular vistas and the urban haze is dominated by regional haze that is being addressed by the Regional Haze rules.

4. I think it is scientifically justified to use $PM_{2.5}$ mass as a basis for protecting visibility. I think a "one size fits all" national urban visibility standard is unnecessary in the East because the Regional Haze rules will improve visibility in the urban areas in the East. The alternative times and the range of levels are inappropriate because no attention was given to 4 and 8-hour background distributions.

5. No comment.

6. No comment.

# Additional Comments on the January, 2005 PM Staff Paper and the April 6 & 7, 2005 CASAC Meeting

by

George T. Wolff
(4/15/05)

### Primary Standards

I agree that there should be annual and 24-hr primary NAAQS for $PM_{2.5}$. I agree that there should only be a 24-hr primary NAAQS for $PM_{10-2.5}$. The form of both 24-hr standards should be the 98$^{th}$ percentile because, as I articulated in my earlier comments, this is more robust than a 99$^{th}$ percentile. The 99$^{th}$ is a more extreme value, which has the property of bouncing in and out of compliance from year to year.

I was disappointed that the health effects experts on the PM Panel did not consider or discuss the carefully documented public comments prepared by Suresh Moolgavkar, Anne Smith, Jon Heuss, Allen Lafohn, Kenny Crump, and Paul Switzer. Collectively, these comments present a compelling case that model selection determines the outcome of any particular epidemiology study when the relative risk is barely above the noise level. These comments demonstrate that the PM issue is not about health effect studies, but it is solely about the design and interpretation of statistical models. Depending upon the design, one can make a case for implicating any measured pollutant or no pollutant at all. As a result, I cannot support the tightening of any of the PM NAAQS based on the epidemiology studies, and the toxicology studies provide no evidence of adverse health effects near the present levels of the NAAQS. For $PM_{10-2.5}$, where admittedly the evidence for a causal relationship is even weaker than for $PM_{2.5}$, I support creating a placeholder NAAQS that would force the agency to collect additional data without creating the economic hardships associated with nonattainment. A placeholder NAAQS around 100 $\mu g/m^3$ would accomplish those goals.

### Secondary Standards

It is inappropriate and unnecessary to establish a secondary NAAQS for urban visibility. When urban visibility is impaired in the Eastern U.S. it is because of regional haze for which we already have Regional Haze rules. The Eastern U.S. is unlike the Western U.S. where urban “brown clouds” occur in many cities. This should strictly be a local issue. In addition many Eastern cities have no scenic vistas to protect. Any urban haze rules should be voluntary and locally adopted.

In addition, Chapter 7 does not provide the information necessary to put this issue in proper perspective. For example, there is no information on what the 4 to 8 hour background concentrations are. We know that the variability of pollutant concentrations increases as the sampling time decreases. Therefore using 24-hr background concentrations to estimates the 4 to 8 hr distributions would be inappropriate. We need to analyze 4 to 8 hr data. In addition, there is no discussion on how many urban areas would be nonattainment. We need tables like Table 5-3 to see how many areas would be nonattainment at various levels within the ranges proposed by

EPA.  My sense is that a 4 or 8 hr standard near the low end of the range proposed by EPA would cause widespread nonattainment and result in the secondary NAAQS being the controlling NAAQS for $PM_{2.5}$.

## NOTICE

This report has been written as part of the activities of the U.S. Environmental Protection Agency's (EPA) Clean Air Scientific Advisory Committee (CASAC), a Federal advisory committee administratively located under the EPA Science Advisory Board (SAB) Staff Office that is chartered to provide extramural scientific information and advice to the Administrator and other officials of the EPA. The CASAC is structured to provide balanced, expert assessment of scientific matters related to issue and problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the EPA, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use. CASAC reports are posted on the SAB Web site at: http://www.epa.gov/sab.