UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

February 18, 2004

EPA-SAB-CASAC-04-004

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

The Honorable Michael O. Leavitt
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 Ongoing Peer Review of the Agency's *Fourth External Review Draft of Air Quality Criteria for Particulate Matter* (June 2003); and Peer Review of the *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information (OAQPS Staff Paper – First Draft)* (August 2003) and a Related Draft Technical Report, *Particulate Matter Health Risk Assessment for Selected Urban Areas (Draft Report)* (August 2003)

Dear Administrator Leavitt:

EPA's Clean Air Scientific Advisory Committee (CASAC), supplemented by expert consultants — collectively referred to as the CASAC Particulate Matter (PM) Review Panel ("Panel") — met in a public meeting held in Research Triangle Park (RTP), NC, on November 12-13, 2003 to: (1) discuss follow-on matters related to its ongoing peer review of the two-volume, June 2003 draft document, *Fourth External Review Draft EPA Air Quality Criteria for Particulate Matter* (EPA/600/P-99/002, aD, bD); and (2) conduct a peer review of the *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information (OAQPS Staff Paper – First Draft)* (EPA-452/D-01-001, August 2003) and a related draft technical report, *Particulate Matter Health Risk Assessment for Selected Urban Areas (Draft Report)* (August 2003).

This meeting was, in part, a continuation of the CASAC PM Review Panel's review of the Fourth External Review Draft of the Air Quality Criteria Document (AQCD) for PM in the current cycle for reviewing the National Ambient Air Quality Standards (NAAQS) for PM. As noted below, the Panel held extended discussions with EPA staff members on the plans for the completion of the AQCD for PM. The revised draft Chapters 7 and 8 of the Fourth External Review Draft of the PM AQCD were provided to the Panel and the public on December 30, 2003. A CASAC PM Review Panel teleconference to discuss these two revised draft chapters was held on February 3, 2004. The draft version of the integrative synthesis chapter (Chapter 9) is expected to be available shortly. A subsequent meeting of the Panel is planned when the remaining issues related to Chapters 7 to 9 will be reviewed.

In addition, the Panel reviewed the first draft of the Staff Paper (SP) for PM. This version of the staff paper was a preliminary version since the Panel has not yet closed on its review of the PM AQCD. In addition, further risk analyses and analyses of alternative forms of the PM standards are planned and will be included in the next version of the staff paper that will be presented to the CASAC PM Review Panel following the completion of the review of the AQCD for PM. However, the Panel felt it was very useful to be able to review the SP in it current form and to raise issues that are seen to need addressing while the air quality criteria document is being finalized.

1. Background

The CASAC 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 national ambient air quality standards (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. EPA is in the process of updating, and revising where appropriate, the AQCD for PM as issued in 1996. The roster of the CASAC PM Review Panel is found in Appendix A.

The CASAC PM Review Panel reviewed the October 1999 First External Review Draft of the AQCD for PM in December 1999, focusing primarily on the organization, structure, and presentation of material in the draft document. This was an early, incomplete draft of the PM AQCD, and it was understood that additional information would be incorporated in subsequent drafts. Accordingly, there was no expectation that the Panel would close on the draft document at this stage of its development. Nevertheless, the Panel was generally complimentary about the content and quality of this draft AQCD, while noting the need for considerable development both in structure and content.

The Agency revised the document in response to CASAC PM Review Panel and public comments, as well as to reflect additional new studies on PM effects that were not available in time to be referenced in the First External Review Draft. In July 2001, the Panel met again in a public meeting to review the March 2001 draft document, *Air Quality Criteria for Particulate Matter – Second External Review Draft*. Despite the fact that this version of the document was substantially revised and expanded, the Panel could not come to closure on that draft document and requested that the Agency further revise the draft PM AQCD.

EPA again revised the document in response to comments from the Panel and the public, and to reflect more new PM studies that had become available. The CASAC PM Review Panel met again in a public meeting in July 2002 to review the two-volume, April 2002 Third External Review Draft of the AQCD for PM. Following that third CASAC meeting, EPA again revised the document in response to CASAC PM Review Panel and public comments, and also to take into account peer-reviewed reanalyses of a number of epidemiological studies conducted to address statistical modeling issues that were identified after release of the latest draft PM AQCD.

On June 30, 2003, the Agency made available for public review and comment a Fourth External Review Draft of the revised AQCD for PM. The CASAC PM Review Panel met again in a public meeting on August 25-26, 2003. In summary, the Panel felt that this version of the draft document, while substantially improved over the Third External Review Draft, still required additional revisions — to include a completely rewritten integrative synthesis (Chapter 9) — before it could be deemed to represent an acceptable assessment of the current science on particulate matter. Dr. Les Grant, Director of EPA's National Center for Environmental Assessment (NCEA)/RTP, committed to draft a set of "framework questions" to be used to guide the restructuring of Chapter 9 ("Integrative Synthesis") of the PM AQCD. A teleconference was held on October 3, 2003 for the Panel to discuss follow-on matters related to its review of the Fourth External Review Draft of the AQCD for PM, and specifically, the discussion of the 'framework questions" leading to the restructuring of Chapter 9.

Furthermore, on August 29, 2003, the Office of Air Quality Planning and Standards (OAQPS), within EPA's Office of Air and Radiation, made available for CASAC and public review and comment the *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information (OAQPS Staff Paper – First Draft)* (first draft PM Staff Paper) and a related technical report, *Particulate Matter Health Risk Assessment for Selected Urban Areas (Draft Report)* (draft PM Risk Assessment). The purpose of the Staff Paper is to evaluate the policy implications of the key scientific and technical information contained in the EPA's AQCD for PM, and to identify critical elements that EPA believes should be considered in the review of the PM NAAQS. In essence, the Staff Paper is intended to "bridge the gap" between the scientific review contained in the AQCD for PM and the public health and welfare policy judgments required of the EPA Administrator in reviewing the PM NAAQS. The draft Risk Assessment describes and presents the preliminary results from a PM health risk assessment for fine particles ($PM_{2.5}$), coarse fraction particles ($PM_{10-2.5}$), and PM_{10}. The risk assessment methodology and preliminary results also are summarized in the first draft Staff Paper. The general methodology used in the risk assessment had been previously discussed in an advisory teleconference in May 2002 and two consultations (February 2002, May 2003). In these discussions, the Panel discussed the selection of cities to be examined as well as the need to provide PM_{10} risk assessments as a basis of comparison with $PM_{2.5}$ and $PM_{10-2.5}$.

2. CASAC PM Review Panel's Ongoing Review of the *EPA Air Quality Criteria for Particulate Matter (Fourth External Review Draft)*

The CASAC Particulate Matter Review Panel held extended discussions with staff members from NCEA on the plans for the completion of the Air Quality Criteria Document for PM. There was an opportunity for the staff to obtain clarification on the comments provided in the August 25-26, 2003 Panel meeting and the October 3, 2003 teleconference. The revised draft Chapters 7 and 8 of the Fourth External Review Draft of the PM AQCD were provided to the Panel and the public on December 30, 2003. A CASAC PM Review Panel teleconference to discuss these two revised draft chapters was held on February 3, 2004. The draft version of the integrative synthesis chapter (Chapter 9) is expected to be available shortly. A subsequent meeting of the Panel is planned when the remaining issues related to Chapters 7 to 9 will be reviewed.

3. CASAC PM Review Panel's Initial Review of the EPA's *Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical* Information (OAQPS Staff Paper – First Draft)

Subsequently, the Panel reviewed the first draft of the Staff Paper for PM. This version of the staff paper was a preliminary version since the Panel has not yet closed on its review of the PM AQCD. In addition, further risk analyses and analyses of alternative forms of the PM standards are planned and will be included in the next version of the staff paper that will be presented to the CASAC PM Review Panel following the completion of the review of the AQCD for PM. However, the Panel felt it was very useful to be able to review the SP in it current form and to raise issues that are seen to need addressing while the air quality criteria document is being finalized.

The remainder of this report summarizes the Panel's collective comment of the current version of the SP document. At this time, we are primarily focusing on the methodologies and approaches being taken since the Panel recognizes that a revised draft of the PM Staff Paper will be forthcoming that will reflect the changes in the AQCD as well as providing the results of the additional risk analyses including those on alternative forms of the standard. The comments of the individual Panel members are provided in Appendix B to this report.

In general, there is particular concern with respect to the lack of adequate consideration for ecosystems and welfare effects such as urban visibility. We will return to this problem when Chapter 5 (*Characterization of PM-Related Welfare Effects*) is evaluated later in this report. The overall structure of the document and the approaches taken in the SP, with the exception of the welfare effects, are appropriate, although there are a number of problems and issues that are described in this report.

Chapter 2 (*Air Quality Characterization*) reviews the basic atmospheric behavior of PM, the current understanding of concentrations and measurements and the relationship of ambient concentrations to human exposure. In general, this chapter is well written and represents a comprehensive summary of information contained in Chapters 2, 3, and 5 of the Fourth External Review Draft of the AQCD for PM. Nevertheless, there are some issues that the Panel would like to bring to the attention of OAQPS.

The scientific information concerning coarse thoracic ($PM_{10-2.5}$) particles is rather limited. However, some specific properties of these particles that are important for establishing a standard should be emphasized. These include: a shorter atmospheric lifetime; significant differences in chemical and/or biological compositions of particles in this size range depending on a geographical location; and, most importantly, a limited penetration into indoor environments that can explain low correlation between personal exposure and outdoor concentrations (as measured by central monitors). There need to be clearer distinctions made in describing composition and aerodynamic properties among the various size fractions (ultrafine,

fine, and coarse) and additional discussion of how those differences then affect the patterns of human exposure, dose to the lung, and variability of potential effects.

Section 2.8 discusses PM exposure assessment issues. This is a well-written section and addresses some key findings. This section relies only upon older studies (PTEAM, for example) and does not report findings from the more recent exposure assessment studies. The more recent work provides important new information and should not be neglected. There should have been more emphasis on the differences in behavior between fine and coarse particles. There is also relatively low penetration of coarse particles in outdoor air (inversely related to particle size) into indoor environments. Results from the very limited existing exposure studies of coarse particles suggest no relationship between personal exposure and outdoor coarse particle concentrations. The implications of these findings with respect to coarse particles need to be evaluated as an important basis for the coarse particle standard.

Section 2.9 is not well-balanced between visibility and climate effects and does not provide the background for the welfare assessments to be subsequently made in Chapter 5.

Chapter 2 should provide a sufficient base of information for the assessments. There is a good amount of background material for the health effects assessment in Chapter 4 (*Characterization of Health Risks*), but there is not the parallel basis for the welfare effects assessments in Chapter 5. There may not be adequate information, but if that is the case, the document needs to reflect the lack of information. There needs to be a discussion of visibility effects in both Class 1 areas and in urban areas. An important welfare effect for the standard-setting process occurs when visibility in urban areas is reduced to low levels.

Chapter 3 (*Characterization of PM-Related Health Effects*) is a summary of the information in the AQCD. It may benefit from the final version of the air quality criteria document when the integrative synthesis provides a basis for a more cohesive presentation of the understanding of the adverse health effects arising from exposure to airborne PM. However, clearer distinctions need to be drawn between the strength of information that is available on PM_{10}, $PM_{2.5}$, and $PM_{10\text{-}2.5}$. Chapter 3 should only be a review of the state of knowledge regarding human health effects of PM within the broader context of exposure to air pollution. The current version was felt to present an objective view of the science without regards to its specific policy implications. The Panel felt that the current version tended to take too much of an advocacy view of the human health effects studies.

The PM Staff Paper, like the PM AQCD, lacks a clear set of criteria for the selection of studies that are to be included in the discussion. Without the introduction of a well-defined set of criteria, the question will continue to arise as to why some studies are included and others are excluded. This leads to uncertainty as to the nature of the evidence reviewed and to the potential for bias in the selection process, or at the least the perception that there may be bias.

The Staff Paper is characterized by the same fuzziness around critical concepts as the Air Quality Criteria Document, particularly in relation to confounding effect modification and causality. There is laxness in the language around these concepts that leads to ambiguity of interpretation. In particular, the document does not carefully separate the quality and extent of the evidence available from the conclusions that might be reached. Examples are highlighted in specific Panelists' comments provided in the appendix to this report. These issues have been raised in the review of the PM AQCD and the clarifications to that document can help to focus the discussion in the Staff Paper.

There is a considerable emphasis in the SP on "consistency and coherence." The demonstration of consistency of positive effects across time-series studies is in some sense the result of a process that may involve selection of positive effect estimates in any given study from the sometimes large number of estimates generated by data analysis; some of these estimates may have been consistent with "no effect." "Consistency," if defined as positive effects in multiple studies, is therefore a likely outcome of a selection and modeling process that may bias towards including positive effect estimates. Further, the Agency's discussion of the use of findings from studies involving multiple cities (p. 3-89), in particular the National Morbidity, Mortality and Air Pollution Study (NMMAPS), to argue for consistency of effects should acknowledge the limitations of this modeling approach with respect to heterogeneity of results.

There is a significant difference between heterogeneity and variability in assessing results across study sites as well as by study design. The fact that results do not appear to be uniformly consistent is probably a strength in the data rather than a weakness. (For a thorough discussion of this issue, see David A. Savitz. *Interpreting Epidemiologic Evidence. Strategies for Study Design and Analysis*. Oxford University Press Inc., NY, 2003.) Because these studies are done with variable degrees of like data sources, one would expect variable results. It is also inappropriate to selectively assess studies involving multiple cities to select only those specific cities having positive or negative estimates, as these studies were not intended to be analyzed or interpreted in this fashion

The arguments put forward against considering the gaseous pollutants as confounding factors have already been questioned in our comments on chapter 8 of the PM AQCD. The SP incorrectly states that neither ozone nor SO_2 can be considered to cause cardiac effects (p. 3-73), whereas both have been shown to have cardiac effects in experimental studies (e.g., Tunnicliffe WS, Hilton MF, Harrison RM, Ayres JG, The effect of sulphur dioxide exposure on indices of heart rate variability in normal and asthmatic adults, *European Respiratory Journal* 17 (4): 604-608 APR 2001; Gong H, Wong R, Sarma RJ, Linn WS, Sullivan ED, Shamoo DA, Anderson KR, Prasad SB, Cardiovascular effects of ozone exposure in human volunteers, *American Journal Of Respiratory And Critical Care Medicine* 158 (2): 538-546 AUG 1998). The suspicion that air pollutants can cause cardiac effects is relatively new, so that there are very few data on cardiac effects of pollutants other than PM.

The Staff Paper repeats the argument in the PM AQCD in support of the notion that the gaseous pollutants are merely surrogate measures of ambient PM, and, interestingly, that CO and NO_2 are markers of vehicle-generated PM, and that SO_2 and ozone are markers of sulfate (p. 3-74). The Panel has raised questions regarding this material in Chapter 8 of the AQCD and thus, the same concerns prevail here. The ozone-sulfate correlations are often weak, so that ozone does not appear to be an appropriate indicator of sulfate. On the other hand, we are developing a better understanding that some pollutants are useful source indicators and of the complicated relationships among the concentrations of some key pollutants as they co-exist and interact in the air pollution mixture. We need to acknowledge the possibility that PM itself is simply a surrogate as well for the air pollution mixture and that the effects attributed to PM largely reflect exposure to the urban air pollution mixture more generally. The main disagreement is whether PM itself is immune from such considerations, that is, whether (1) gaseous pollutants are surrogate measures of PM; or (2) all of the pollutants, including PM, are surrogate measures of aspects of the atmospheric pollutant-meteorology mix.

The figure on p 3-96 and the corresponding discussion in the text (p. 3-94) attempt to address the plausibility of confounding by the gaseous pollutants by plotting effect estimate size, *i.e.*, relative risk (RR) against gaseous pollutant concentration for several studies. The fact that RR does not increase with increases in gaseous pollutant concentrations is taken as evidence that confounding by gaseous pollutants is unlikely. While this information is informative, this conclusion does not follow. Joel Schwartz introduced the approach of plotting effect size against the temporal correlation between PM and the gaseous pollutants, and this approach should also be cited.

The Panel continues to have concerns regarding the reporting of the "best lag" approach that was used in the AQCD and we again suggest that there is literature on this issue, particularly the work of Lumley and Sheppard (Assessing seasonal confounding and model selection bias in air pollution epidemiology using positive and negative control analyses. *Environmetrics* 11: 705-717, 200) that provides clear guidance on this issue.

The Panel agrees that the multi-city studies should be given the most weight. However, not all multi-city studies should be given equal weight. Not only are multi-city studies characterized by more precise estimates of effect, but some also use an unselected sample of cities and theoretically avoid publication bias. Only the NMMAPS and the Canadian studies, of the studies listed in Table 3-2 (p. 3-17), are unselected. The NMMAPS estimates of effect are the lowest, and the Canadian effects are sensitive to model specification. Thus, care needs to be taken in the interpretation of other multi-city studies.

There is a suggestion that generalized additive models (GAM) are preferable over generalized linear model (GLM) approaches and it is not clear to the Panel that this choice is appropriate. There are advantages and disadvantages to each and the application also depends on the actual model being fitted as well as the fitting technique.

An important issue that is not adequately addressed is the nature of the exposure-response relationship, as characterized in the epidemiological studies. The data from most studies, including the various time-series studies and the cohort studies, have been analyzed using linear models, mostly without a threshold. These models estimate the increment in relative risk per unit exposure, generally without consideration of the actual levels of exposure. Such linear models, while indicating an adverse effect do not explicitly consider the levels at which the effects were estimated. Thus, these results provide little guidance as to where a standard could be set to provide "an adequate margin of safety." This issue needs more thoughtful discussion in the PM staff paper.

Chapter 4 (*Characterization of Health Risks*) reports the results of the risk assessments for $PM_{2.5}$, PM_{10}, and $PM_{10-2.5}$ based on the current forms of the PM NAAQS, using the methodology that had been employed in the last round of the development of the NAAQS for particulate matter. The exposition of the risk assessment would benefit from a clear discussion of the near-linearity of the concentration-response function and its implications for time-averaging. The "effective threshold" sensitivity analyses may be the only calculations in which nonlinearity plays an essential role.

To provide a perspective on the risks being estimated in this analysis, the PM SP should provide coverage of the baseline morbidity and mortality statistics for at least the cities for which the risk assessments are being applied as well as more general regional and national values with special reference to cardiovascular and respiratory morbidity and mortality statistics. It is important that the SP include such statistics in order to: (a) provide perspective on PM-associated health responses; and (b) emphasize their central role in interpreting relative risk models for PM-associated health responses. There also needs to be additional sensitivity analyses for "effective threshold", particularly since, on page 130 of the Abt document, the statement is made "Different choices of slope adjustment methods can yield substantially different results." The document should provide a perspective on the range of these different results.

The Panel was disappointed with the ecological portion of Chapter 5 (*Characterization of PM-Related Welfare Effects*) which does not move toward a risk-based approach to evaluating the ecosystem effects. This SP was an initial opportunity to begin to frame ecosystem risk, although it is likely that there was insufficient data to permit a full risk assessment to be made. The Ecological Processes and Effects Committee (EPEC) of the EPA Science Advisory Board (SAB) prepared a document describing a framework for performing a risk assessment. This document appeared in the middle of the PM review process and thus, a full risk assessment process beginning in the PM Air Quality Criteria Document was not practical. However, it would have been useful in the SP to begin the move toward the risk assessment approach. This same protocol underpins all of the human health section in the Staff Paper, and it provides a structure and framework for the analysis. It would have been useful to use the PM SP to begin to develop the framework such that it could be more effectively employed in the future when other criteria pollutants are being assessed.

The Staff initial efforts in addressing PM-related effects on vegetation and ecosystems is based on the overriding assumption that, for the most part, one can attribute the response or responses of a receptor to a given air quality stressor within a given short time frame. This approach simply does not work in the case of PM. This is very unfortunate from the standpoint of environmental protection, especially in light of the fact that there are some forested ecosystems in the U.S. which are showing clear evidence of "nitrogen saturation," a portion of which is due to particulate nitrate deposition. The problem here is that this "nitrogen saturation" has been brought about by chronic long-term exposure to elevated nitrogen deposition. It is the cumulative load of nitrogen over time that has resulted in some forested ecosystems being nitrogen saturated. Some would say that the fact that we do not know the exact contribution of "particulate nitrate" deposition to the nitrogen saturation evidenced in some forest ecosystems prevents us from doing anything. This is not true. What is needed is a philosophical change in the way one approaches ecosystem protection. The European concept of "critical loads" is suggested as one possible scientific approach. This approach would more readily lend itself to risk assessment than the current information.

EPA appears to (again) be avoiding or postponing any serious consideration of a short-term secondary $PM_{2.5}$ standard to address adverse visibility effects. The Agency previously reported to Congress in 1979 that "Recently initiated research efforts in monitoring of fine-particles, transformation and transport studies, and progress in evaluating visibility values could provide support for a decision on the desirability of such an air quality standard by 1982 or 1983." "New" materials presented in the staff paper — including a preliminary comparison of Automated Surface Observing System (ASOS) visibility data and nearby $PM_{2.5}$ data and a proposed photographic evaluation method to determine public judgments of "adverse" visibility levels — are cited as approaches that could be employed in a <u>future</u> round of PM AQCD review and standard setting. These efforts indeed confirm and extend findings that have been well-established for 20 years, and raise the question why there should not be serious consideration of a secondary standard at the present time. EPA is moving rapidly to include continuous $PM_{2.5}$ mass monitors in the compliance monitoring network which is critical to provide better information for health studies, but at the same time, provides a near-term opportunity for applying a secondary standard for urban visibility protection. As noted in the Agency's 1979 Report to Congress, "such a standard would accelerate progress toward improved visibility throughout the Eastern United States and might also increase the efforts for visibility improvements in major urban areas of the Western United States. Thus, a secondary air quality standard for fine particles could effectively complement visibility protection programs in class I areas." A sub-daily averaging time, for example 6 or 8 daylight hours, would be especially relevant for addressing the most perceptible adverse visibility effects of $PM_{2.5}$ in non-Class 1 (urban and suburban) areas, would tighten the (dry) $PM_{2.5}$/(ambient) visibility relationship by reducing the influence of ambient aerosol water, and would substantially minimize the differences between Eastern and Western conditions.

The current short-term 24-hour primary standard of 65 Fg/m^3 (which is also the secondary standard) offers no protection against adverse visibility effects. At a minimum, EPA should dispense with the pretense that this is a secondary standard which offers any protection at its current level or even if a primary standard were set toward the upper end (50 Fg/m3) suggested for a revised short-term primary standard. Severely impaired visibility can be seen on days when $PM_{2.5}$ does not exceed 50 Fg/m^3, illustrating that lowering the primary standard to the upper end of the (30-50 Fg/m^3) range suggested would also offer no protection against adverse visibility effects. EPA's recent practice of setting secondary standards equal to the level of primary standards has no logical justification, and presumes that human health is always more sensitive to pollution effects than any other component of the environment or public welfare. This is simply not true for visibility effects. The human eye may be more sensitive to short-term $PM_{2.5}$ variations than is the human cardiopulmonary system, and as concentrations approach zero, perceptible visibility effects can be detected at concentrations less than a few Fg/m^3.

The majority of the Panel concluded it is premature to provide a detailed review of Chapter 6 (*Staff Conclusions and Recommendations on PM NAAQS*) since significant changes are still needed in the earlier chapters providing the review of the science. We have debated whether or not to provide any comments and the Panel has decided that we want to wait until the

PM AQCD is complete and the revised version of the PM Staff Paper based on the completed AQCD is provided to us. However, there are members of the Panel who have provided their individual comments on the draft version of this chapter. There is clearly a diversity of views on the Staff Paper in its current form. Careful attention to the revisions will be needed to resolve the issues that have been raised. We hope the comments in this report and the attached individual comments help to improve the next version of the Staff Paper. As always, the Panel wishes the Agency well in this important endeavor.

Sincerely,

Dr. Philip K. Hopke, Chair
Clean Air Scientific Advisory Committee

Appendix A – Roster of the CASAC Particulate Matter Review Panel

Appendix B – 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 Particulate Matter Review Panel*

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

Also Member: SAB Board

CASAC MEMBERS
Dr. Frederick J. Miller, Vice President for Research, CIIT Centers for Health Research, Research Triangle Park, NC

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

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

Dr. George E. Taylor, Jr., Professor and Assistant Dean, School of Computational Sciences, George Mason University, Fairfax, VA

Dr. Sverre Vedal, Professor of Medicine, National Jewish Medical and Research Center, Denver, CO

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

CONSULTANTS
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. Günter Oberdörster, Professor of Toxicology, Department of Environmental Medicine, School of Medicine and Dentistry, University of Rochester, Rochester, NY

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

Mr. Ronald H. 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

SCIENCE ADVISORY BOARD STAFF

Mr. Fred Butterfield, CASAC Designated Federal Officer, 1200 Pennsylvania Avenue, NW, Washington, DC, 20460, Phone: 202-564-4561, Fax: 202-501-0582, (butterfield.fred@epa.gov) (FedEx: Fred A. Butterfield, III, EPA Science Advisory Board (1400A), Ariel Rios Federal Building North, Suite 6450, 1200 Pennsylvania Ave., NW, Washington, DC, 20004, Tel.: 202-564-4561)

* Members of this CASAC Panel consist of:

a. CASAC Members: Experts appointed to the statutory Clean Air Scientific Advisory Committee by the EPA Administrator; and

b. CASAC Consultants: Experts appointed by the SAB Staff Director to serve on one of the CASAC's National Ambient Air Quality Standards (NAAQS) Panels for a particular criteria air pollutant.

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

This appendix contains the preliminary and final written comments of 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 the all suggested edits, a full perspective, and range of individual views expressed by Subcommittee 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 consensus views of the CASAC PM Review Panel and the CASAC are contained in the text of the report to which this appendix is attached. Panelists providing comments are listed on the next page, and their individual comments follow.

Panelist	Page #

Dr. Frederick J. Miller

Review Comments: OAQPS Staff Paper-First Draft **11_20_2003**
Fred J. Miller, Ph.D.

2. Air Quality Characterization

General Comments

Overall, this chapter does a good job of providing the background information to understand how PM is characterized, what the emission sources are, the measurement methods used to determine PM levels, and the nature of the trends and spatial patterns of PM in the United States.

Specific Comments

p. 2-20 What is magical about requiring 11 observations per quarter to use the data from a monitoring site, particularly as the various standards have now resulted in sampling schedules in monitoring programs that are different from what led to Appendix N of the July 18, 1987 Federal Register notice specifying at least 11 observations are required for a site. Staff should examine if there would be an impact on the risk modeling that is undertaken in subsequent sections of the staff paper should a different specification of number of samples be used?

p. 2-25, l. 1 Using the monitor with the highest concentration in each monitored county to represent the value for that county is a biased accounting of the PM concentrations present in any given geographical area. The CD shows that some cities have extensive variability among monitors, and the variability for PM 10-2.5 should be substantial in large metropolitan areas. Some weighting of the site monitoring values by the size of the surrounding population could easily be done and would result in a much more representative and far less biased concentration for use in the risk analyses. The tremendous influence on risk of which way the monitors are used in an area is borne out in Table 4-12.

p. 2-28 The % Diff reported in Table 2-4 is based upon using the maximum site value in the denominator. It seems like the % Diff should be based on the minimum site value. This would show that there is even more variability among sites that what staff have chosen to represent as the amount of variability.

3. Characterization of PM-Related Health Effects

General Comments

Throughout this chapter, the staff paper fails to acknowledge the difficulty of ascribing effects only to PM as opposed to PM as a reflection of the ambient mix of pollutants. Multi-pollutant models presented in the CD typically show substantial decreases in the magnitude of effect estimates for PM and have other pollutants also statistically significant and yet the CD typically portrays gaseous pollutants as only representing surrogate measures of ambient PM. This makes one wonder -- when the Ozone CD comes around, will PM be presented as merely a surrogate for ozone? Staff appear to fixate on PM, to not fully describe deficiencies in current studies in treating weather, and to not acknowledge the sensitivity of effect estimates to model specification and fitting procedures.

Here, I repeat the comments I made using data in Chapter 9 of the CD (The Integrative Synthesis chapter) relative to presenting an analysis using all of the studies in Chapter 8 of the CD that would identify where the strongest case can be made for the need for PM standards in light of the various studies that have been conducted. The staff paper authors have been quite selective in their use of studies to identify the appropriate indicator variable be that PM 10, PM 2.5, or PM 10-2.5. However, if all the studies in Chapter 8 were used, one would quickly see that PM 10 provides the most consistent indicator of various types of effects ranging from mortality to respiratory morbidity despite what has obviously been an *a priori* science policy and political decision that the Agency will move forward with promulgation of separate PM 2.5 and PM 10-2.5 standards.

Using only the data in Chapter 9 of the CD, one sees that PM 2.5 also does a reasonably good job for mortality and respiratory morbidity but is much poorer as an indicator variable for cardiovascular morbidity as is PM 10. In fact, for cardiovascular morbidity, PM 10-2.5 does almost as good a job as does PM 10. While the table below is a relatively simple one and does not account for various investigators analyzing the same city by different methods or over a different period of time, the point is that the data in Chapter 8 of the CD provide a wealth of information for attempting to identify the appropriate PM indicator variable and the level of that variable against which public health should be protected. The current development of the PM story in Chapter 3 of the Staff Paper reflects a more selective rather than a weight of the evidence analysis.

	Total No.	PM 10		PM 2.5		Pm 10-2.5	
Endpoint	Analyses	No.	% Positive	No.	% Positive	No.	% Positive
Mortality	39	26	62	29	59	19	12
Cardiovascular Morbidity	18	18	56	8	38	6	50
Respiratory Morbidity	24	23	78	13	62	7	29

<u>Specific Comments</u>

p. 3-7, l. 18 The statement of some of the toxicology studies using relatively low doses that are close to ambient concentrations needs to be documented with references. In my opinion, none of these studies are at low enough doses.

p. 3-10, l. 13 The staff paper asserts papers supporting new indices of morbidity such as low birth weights. The specific studies are limited, d0 not show anything other than an association, and are far from being in the category of supporting that PM is causative. Some of the caveats associated with these studies should be provided at a minimum. Moreover, later on (p. 3-41, l. 10) the staff paper quotes the draft CD as the results of these studies being suggestive of a causal relationship between PM exposure and infant mortality. This statement is too strong given the evidence and should be modified in both the CD and the staff paper.

p. 3-15, l. 4 Mortality results are cited from the multi-city and single-city studies as being, with only a few exceptions, generally positive with many being statistically significant. The NMMAPS analyses show only 2 out of 90 cities as being statistically significant for PM effects. Realizing that the NMMAPS was not designed to infer city specific effects, the staff paper still seems to be accepting values positive but not statistically different from zero as being supportive of PM effects. This is, in my opinion, out of line with the statistical rigor one would expect for setting of standards of such national importance.

p. 3-18, l. 6 The heterogeneity of PM effect estimates is commented on and then dismissed as attributing to the way the cities were selected and not necessarily individually having the power to compare with single-cities. That being said, I have to agree with Dr. Suresh Moolgavkar when he notes in his written comments that the individual city effects for the 15 northeast cities lie between 0.223 and o.271 and yet the regional effect estimate is 0.409 as a situation that flies in the face of common sense. The empirical Bayes estimation procedure with the hierarchial method appears suspect.

p. 3-29, l. 3 The statement that effect estimates between PM 10-2.5 and total mortality are generally positive though less likely to reach statistical significance compared to PM 2.5 and total mortality represents a stretch beyond reasonable interpretation. Only 3 out of 19 studies have PM 10-2.5 being statistically significant.

p. 3-31, l. 5 A range of 2-20% for cardiovascular or cardiorespiratory mortality represents a significant level of uncertainty about the potential effects of PM on these endpoints. Yet the general thrust throughout the staff paper is that the results are quite coherent for PM effects. Staff need to acknowledge that there is more uncertainty in effect estimates that what is frequently currently stated.

p. 3-41, l. 9 The discussion about PM exposure and infant mortality is overstated in the conclusion of the paragraph where the CD is cited as concluding the results of these studies are “suggestive of a causal relationship between PM exposure and infant mortality”. Discussions at the November 12-13 CASAC meeting reinforced my belief that this is too strong a statement in the CD and is not warranted in the Staff paper. These 4 studies show association at best and clearly do not warrant a statement implying causality.

p. 3-42, l. 25 The life shortening study of Brunekreef is puzzling. If the population life expectancy was truly reduced by 1.31 years with an exposure difference of 10 µg per cubic meter, epidemiology studies conducted in underdeveloped countries should clearly detect this. To my knowledge they do not. This makes the CD comment that the potential loss of population life expectancy might be even greater than Brunekreef’s estimate even harder to believe to be plausible.

p. 3-50, l. 20 This sentence and the next one are gross overstatements of the strength of the associations between PM 10-2.5 and admissions for both respiratory and cardiovascular diseases. Only 1 of 7 respiratory and 2 of 10 cardiovascular disease admission studies shown in Figure 3-9 are statistically significant and about 7 of these studies have a positive value of less than about 2%. Staff's presentation of the strength of these results is not warranted.

p. 3-60, l. 1 The statement "The results of U.S. and Canadian studies, presented in Table 3-4, generally show increased symptoms and decreased lung function with increases in PM exposure" is an incorrect representation of the data by staff. For the PM indicators present in the table, only 3 of the 14 are statistically significant. Moreover, of the 3 that are statistically significant, 2 relate to PM 15/10 indicators and one to sulfates. How staff could conclude these findings support the need for PM 2.5 or PM 10-2.5 standards is beyond me. The data contained in Figure 3-11a and Figure 3-11b are far more useful for any arguments relative to these endpoints for PM 2.5 and PM 10-2.5, respectively.

p. 3-78, The section on lag periods should make a clearer statement that the potential additivity of effects and the usefulness of distributed lags is more likely to be a possibility for alveolar level effects. Given the rapid clearance of particles from the head and conducting airways, effects from day to day are less likely to be cumulative.

4. Characterization of Health Risks

General Comments

Overall, the types of analyses and assumptions used are adequately described in this chapter. As a first draft, the chapter gives one a flavor or what the final chapter will need to contain. In this regard, there will be a need to be much more comprehensive in examining the validity of the assumptions, in presenting various sensitivity analyses, and illustrating the important findings without overwhelming the reader in minutia. As noted in specific comments below, more attention to thresholds and selection method for monitors to use in the risk analyses are needed.

Specific Comments

p. 4-11 There are two figures labeled as Figure 4.1. They appear on pages 4-11 and 4-18.

p. 4-30, l. 22 "Once it had been determined that a health endpoint was to be included in the assessment, inclusion of a study on that health endpoint was not based on the existence of a statistically significant result." This approach appears to impart a selection bias for the risk calculations because Staff did not adequately explain how the calculations were done, so the text should be clarified. However, it would still be very informative to compare the risk estimates from this approach with those generated only from statistically significant studies on the given endpoint.

p. 4-31 The section on hypothetical thresholds is a good start but falls short of rigorously examining what is potentially the most critical factor for the overall risk estimates and the implication of appropriate levels for PM standards to protect adequately the public health. The current approach evaluates specific “effective biological thresholds” and determines the risk for exposures that exceed the threshold. However, that is not the same as fitting alternative models that have a threshold and see how well these models fit the data. The fact that most investigators have used only log-linear models further amplifies the inadequacy of the current analyses to be definitive on the subject of thresholds.

p. 4-35, l. 22 The statement is made that a sensitivity analysis was conducted to examine the potential impact of using a distributed lag approach for short-term mortality associated with PM 2.5. However, no results of that sensitivity analysis are presented. And the reader should not have to go to the document entitled “Particulate Matter Health Risk Assessment for Selected Urban Areas: Draft Report” to dig out the results of this analysis.

p. 4-36, l. 17 The point is made that the draft CD concludes that the “lack of consistent findings in the AHSMOG study and negative results of the VA study, do not negate the findings of the Six Cities and ACS studies”. Factually this is a correct statement but the reasons given for excluding the AHSMOG and VA studies are superficial. The important point that is missed is that the failure to include all of these studies potentially overestimates the true PM risk.

p. 4-40, l. 10 I strongly object to the rationale provided for why the confidence intervals were truncated at zero. Not allowing the confidence interval to be negative misleads readers into thinking there is always an effect. This truncation violates one of the basic tenants of accurate statistical presentation of data. Moreover, uncertainty analysis methods applied to truncated intervals would lead to under reflecting the amount of uncertainty present in the effect estimates.

p. 4-44 Figures 4-4 through 4-8 would be much easier to understand if a line was inserted to create panels and the specific biological response was identified at the top of the panel. Having to go to the complicated figure legend is a serious detraction.

p. 4-47 Figure 4-7 is the most compelling figure in the entire staff paper as to the need for an annual standard for PM 2.5. (Provided the truncation at zero is corrected.)

p. 4-61, l. 3 Why not refit the C-R with the same original model but subtracting the “hypothetical threshold”? As pointed out in the supporting documentation (Abt, 2003b) the weighted average of the two hockey stick slopes is only one of many ways the slope adjustment could be made and that substantially different results would be possible. Later on in this paragraph, staff note that these sensitivity analyses are intend only to be illustrative of the possible impact on the risk estimates of alternative hypothetical thresholds (I prefer “effective biological thresholds”) and that a more thorough evaluation would require re-analysis of the original health and air quality data. <u>Given the billions of dollars that will be</u>

required to meet current or tighter PM standards, I think OAQPS is obligated to pursue these re-analyses. The need for a better examination of thresholds is reinforced on page 4-66 of the Staff Paper where the statement is made "Based on the results from the sensitivity analyses, the single most important factor influencing the risk estimates is whether or not a hypothetical threshold exists below which PM-related health effects do not occur". From a teleological and practical scientific perspective, I am of the opinion that "effective biological thresholds" exist for PM effects and that the thresholds are themselves different depending upon the biological endpoint under consideration. Most epidemiologists have used log-linear models in their analyses. This model specification basically ascribes Habers Law relative to the relationship between PM exposure levels and time and does not allow the model to identify a threshold. Habers Law is merely a special case of the more general power law family and has been shown not to apply to most biological data sets (Miller et al., Toxicology 149: 21-34, 2000). Given all of this, at a minimum, the Staff paper should explore additional sensitivity analyses for the "effective threshold", particularly since on page 130 of the Abt document the statement is made "Different choices of slope adjustment methods can yield substantially different results".

p. 4-68 Clarify the figure legends. Are the estimates over the course of a year or what?

p. 4-70 The sensitivity of the percent rollback needed to just meet the annual PM 2.5 standard to whether the maximum of monitor-specific annual averages is use or the average of monitor-specific annual averages is clearly demonstrated in Table 4-12. The use of the maximum represents a force fit that produce maximum risk estimates and clearly is not an unbiased representation of the likely exposure values and therefore likely risk for most of the population living within any of the major urban areas. Discussion at the November 12-13 meeting brought out that states and localities can pick either the maximum or the average across the area for use in their compliance and implementation programs – a situation that I believe needs to be changed. There is no reason that census data and population weighting methods should not be used in conjunction with specific monitor values to determine the extent of rollback needed to meet PM standards.

6. Staff Conclusions and Recommendations on PM NAAQS

General Comments

Staff are quick to hone in on a narrow range for consideration of the potential range of concentrations that could be considered for 24 hour and annual standards. The text comes across currently as a proscribed directive to hold the line at the current annual average for PM 2.5 at a minimum and to fill in where the Court said the Agency had to go to cover coarse-fraction particles. There is not much acknowledgement that the data base in support of and the case for a PM 10-2.5 standard is not a particularly strong one. In addition, the uncertainties in risk estimates due to the sensitivity analyses developed in Chapter 4 receive little attention here. The aspect of an "effective biological threshold" for some PM effects does not appear to be reflected in any of the ranges of values for the annual and short-term standards that are discussed. I would expect the next draft of the Staff Paper to include more analyses of the implications of various options and concentration ranges taking into account the uncertainties explored via sensitivity analyses.

Specific Comments

p. 6-5, l. 22 The life shortening new studies, infant mortality and other effects that are quoted from the CD, if further substantiated, as implying a significantly larger life shortening that previously estimated is a statement that comes across as "grasping at straws". Given the uncertainties already in the data base, this is a weak argument in favor of tightening the current PM 2.5 standards.

p. 6-8, l. 9 EPA staff now conclude the Court was correct in arguing that PM 10 is not an appropriate indicator for coarse-fraction particles. Some of the Court's basis for this decision was the double counting contained in PM 2.5 and PM 10 standards. I submit that double counting still exists, although at a diminished level, given the overlap that occurs if the PM 2.5 sampler has a 50% cut point at 2.5 μm and the PM 10-2.5 sample is obtained by differencing values from PM 2.5 and PM 10 samplers. This issue will need to be resolved because even a couple of micrograms of mass will have a tremendous impact on compliance monitoring.

p. 6-8, l. 14 The argument is presented that we now have a lot more information about PM 10-2.5 and there is no need to rely on PM 10 as a coarse fraction indicator. The footnote provides a big caveat in that almost all of the PM 10-2.5 values are estimated from collocated PM 2.5 and PM 10 monitors. Moreover, the argument put forth is strictly an aerosol science and not a biological one because < 2.5 μm compared to 2.5 μm $< X < 10$ μm has extensive overlap in regional deposition in the respiratory tract.

p. 6-18, l. 13 The authors of the Canadian studies that were reanalyzed noted the sensitivity of their analyses to temporal smoothing methods. The text should be clarified to note that the Phoenix and Santa Clara County studies were also reanalyzed and examined for temporal smoothing effects given that EPA Staff are placing great weight on these studies as lending support for considering an annual standard lower than the current one for PM 2.5.

p. 6-21, l. 6 The topic of monitoring values and the way that they are averaged is raised as a concern about sufficient uniformity in public health protection across the country, and staff indicate they are going to explore this issue further in the next draft. This is most appropriate, especially for any proposed PM 10-2.5 standards because monitoring values for coarse-fraction particles vary considerably across an area.

Mr. Richard L. Poirot

OAQPS PM Staff Paper, First Draft, EPA-452/D-03-001 (August 2003)
CASAC Review Comments, R. Poirot, November 2003

General
Overall, this is an excellent 1st draft, which provides a good deal of the kind of "integrative synthesis" which seemed to be lacking in the CD. The general recommendations in chapter 6 for retaining and tightening the primary $PM_{2.5}$ standards, especially the 24-hr standard, is well justified. The justification for specific annual or 24-hour levels of a PM_{10-25} is not as strong, although EPA staff (Karen Martin) made a good argument that taking no action on coarse particles was equivalent to retaining the existing PM_{10} standards. Mort Lippmann's suggestion that a relatively lenient PM_{10-25} standard might be an appropriate near-term "compromise" (which would also inspire the collection of better quality measurement data at more locations) is worth considering.

Considering that a substantial fraction of the "new information" presented in the CD clarifies and emphasizes the health effects of short-term $PM_{2.5}$ exposures, the "logic" of a "controlling" annual standard with a substantially weaker 24-hour "backup" standard has been weakened. If there are different effects from exposures over different averaging times, there can and should be different standards to protect against those effects, without need or consideration of which averaging time should be a more stringent "controlling standard" and which should be a more lenient "backup". As a practical consideration there are few imaginable control strategies that would reduce episodic exposures by merely "shaving peaks" that don't tend to shift the entire distribution downward. Conversely, intermittent or episodic controls (no-burn days, carpool incentives, tele-commuting, etc.) have proven effective at reducing ozone concentrations and (in the case of the Denver visibility standards - at reducing PM emissions) and have added benefits in terms of improved forecasting, more accurate health advisories and increased public awareness. Benefits from short-term emissions controls in specific urban areas would also have benefits over large downwind areas where reduced concentrations would be experienced over various averaging times.

As with previous comments on the CD, I was disappointed that EPA appears to be (again) avoiding or postponing any serious consideration of a short-term secondary $PM_{2.5}$ standard to address adverse visibility effects. "New" material presented in the staff paper – including a preliminary comparison of ASOS visibility and nearby $PM_{2.5}$ data and a proposed photographic evaluation method to determine public judgments of "adverse" visibility levels – are cited as approaches that could be employed in a <u>future</u> round of PM CD review and standard setting. I agree that these methods could indeed be helpful in future refinements of a secondary $PM_{2.5}$ standard, but <u>strongly disagree</u> that consideration of a secondary standard would be premature at the present time.

The 1969 CD for PM and sulfur oxides (predating EPA) includes 3 different graphic depictions of the strong, quantitative relationship between PM and light extinction. In the (outstanding) 1979 "Protecting Visibility: an EPA Report to Congress", (predating both regional haze protection for class I areas and primary standards for PM_{10}), EPA indicated that "a secondary air quality standard for fine particles could effectively complement visibility programs in class I areas" which would "accelerate progress toward improved visibility throughout the eastern United States…and also increase the efforts for visibility improvements in major urban areas of the Western United States." This report also indicated that "[R]ecently initiated research efforts in monitoring fine particles, transformation and transport studies, and progress in evaluating visibility values could provide support for a decision on the desirability of such an air quality standard by 1982 or 1983." Twenty-five years later, EPA is still not quite ready to consider such a standard, but with a few more years of research and analysis…

Visibility impairment is an instantaneous effect, and therefore best addressed by standards applied to short averaging times. The current short-term 24-hour primary standard of 65 ug/m3 (which is also the secondary standard) offers no protection against adverse visibility effects. At a minimum, EPA should dispense with the pretense that this is a secondary standard at its current level or if a primary standard were set toward the upper end (50 ug/m3) suggested for a revised short-term primary standard. Note from my previous "supplemental visibility" comments on the 3rd Draft of the PM CD (attached here for reference), the severely impaired visibility photo for Burlington, VT on 7/7/02 at 62 ug/m^3 $PM_{2.5}$ was not even an exceedance day, and 21 worse days per 3 years are permitted under the current "secondary" standard. The severely impaired visibility illustrated in the Boston photo for 7/16/99 was on a day when $PM_{2.5}$ did not exceed 50 ug/m^3, illustrating that lowering the primary standard to the upper end of the (30-50 ug/m^3) range suggested would also offer no protection against adverse visibility effects. EPA's recent practice of setting secondary standards equal to the level of primary standards has no logical justification, and presumes that human health is always more sensitive to pollution effects than any other component of the environment or public welfare. This is simply not true for visibility effects. <u>The human eye is more sensitive to $PM_{2.5}$ effects than the human lung is</u>, and as concentrations approach zero, perceptible effects can be detected at concentrations less than a few ug/m^3.

Visibility impairment is <u>caused by</u> fine particles. The $PM_{2.5}$ mass / visibility relationship can be described, nationally, in terms of an empirically (or theoretically) derived extinction efficiency of about 6 $\pm$ 3 m^2/g. The spatial and temporal variability in this strong, causal relationship is due almost entirely to the effects of aerosol water present in the ambient aerosol and removed by drying in the FRM definition of (dry) $PM_{2.5}$ mass (from filters weighed at about 40% RH). Regardless of chemical composition, if visibility effects are considered under relatively dry conditions, typical of western areas but also often present at eastern areas at mid-day, the extinction efficiency will strongly converge on a value of 4 m^2/g.

The examples of local judgments of adverse visibility effects presented in the staff paper include short-term (4 to 8-hour, mid-day) visual range limits of 40 km (Phoenix), 50 km (Denver) and 40 to 60 km in the Fraser Valley. Vermont also established a state visibility standard in 1985 expressed in terms of a summer seasonal sulfate concentration of 2 ug/m^3. This compared to a current level at that time of 4 ug/m^3, and was intended to reflect the "reasonable progress required under Section 169a of the 1977 CAA. Had it been attained (we're about half way there) the average visual range would have increased from 40 to 50 km. Under relatively dry mid-day conditions, where an extinction efficiency of 4 to 5 m^2/g would be appropriate, a visual range of 50 km would correspond to a $PM_{2.5}$ concentration of about 15 $\pm$ 2 ug/m^3. While such low concentrations may be economically unfeasible to attain on a short-term basis, a standard set at more lenient levels – in the range of 25 to 35 ug/m^3, reflecting a visual range of about 25 to 35 km under dry mid-day conditions, might be a reasonable compromise. It would be twice as stringent as the current secondary standard and twice as lenient as the local standards cited in the staff paper. Additional flexibility, if needed, is provided by the less urgent nature of a secondary standard, or could be achieved in the usual manner of fiddling with the compliance metrics or compliance dates.

I would also strongly encourage EPA to carefully consider a sub-daily averaging time for this secondary standard, and will provide additional thoughts on this and associated supporting analyses in subsequent more detailed comments.

Chapter 2 is a good summary, clearly written. No major comments – but one significant issue is raised by p. 2-53, lines 18,19 mention of EPA "natural events" policy (which is not entirely clear to me for $PM_{2.5}$ and), which may (especially as methods are rapidly improving to identify natural events) start to complicate determination of the 98th percentile, which already allows exclusion of worst 7 days/year. Some additional discussion of this might be warranted here.

P. 2-9, L. 15-19: Suggest revising to "Potassium in coarse particles comes primarily from soil, with additional contributions from sea salt in coastal areas. Potassium in fine particles comes mainly from emissions of burning wood, with large but infrequent contributions from fireworks, and significant contributions from the fine tail of coarse mode soil particles in areas with high soil concentrations." (July 5th is often the highest fine K day of the year, and at some sites there's more $PM_{2.5}$ K from soil than from smoke).

P. 2-9, L 23-25: Suggest revising to "The amount of particle-bound water will vary with the particle composition and the ambient relative humidity. Sulfates, nitrates ad some secondary organic compounds are much more hygroscopic than…"

P. 2-10, L. 16: Suggest revising to "…ranging from minutes to days …" (the fine tail of coarse mode African & Asian dust can last a week or more, else we would never see it here).

P. 2-12, L. 16: Could add "Episodic emissions from dust storms and forest fires are difficult to quantify and to allocate accurately in space and time, and discerning between natural and anthropogenic "causality" for these source categories is especially challenging."

P. 2-16, L. 21: Suggest revising to "Smoke particles composed primarily of carbon, including black carbon (BC),…"

P. 2-18, L. 25: Suggest adding "ambient" before "PM".

P. 2-27, L. 7: Change "meteorological" to "meteorology".

P. 2-32, L. 2,3: Change to "…consistency of these $PM_{10-2.5}$ is relatively uncertain, and they are…"

P. 2-37, L. 24: Change “was” to “were”.

P. 2-48: Figure 2-21 has “=” instead of dashes in x-axis labels.
P. 2-54: “Background” is a difficult concept. You might consider something like the “IMPROVE minus sulfates” metric that Warren White suggested.

P. 2-61, L. 20-21: “Soil dust” is not “fine mode”. Could rephrase to “…black carbon, and the fine tail of coarse mode soil dust.”

P. 2-62, L. 1-3: Not really correct as stated. Could rephrase to “…of a given mass, dry particle size distribution, and composition…” (the relative humidity has already influenced the size distribution of the ambient aerosols).

Chapter 3 is primarily beyond my expertise, but appears to be logically presented and clearly written. I note the page 3-37, line 4 indication that visibility data were used as $PM_{2.5}$ surrogate in ASHMOG cohort study. Ironic that the $PM_{2.5}$ / visibility relationship is plenty clear enough to use as a quantitative indicator in health studies upon which the primary standards are based, but later (Chapter 5) the PM/vis relationship becomes hopelessly unclear when EPA considers (and rejects) a secondary standard…

Chapter 4 risk assessment methodology is clearly stated but very complex (requires 40 pages to describe methods). I assume this complex approach is necessary, but wish there were a simpler, more straightforward alternative.

P. 4-49, L. 13: Add “as” between “well” and “the”.

P. 4-49, L. 23-25: Rephrase “…Phoenix, it was only possible to develop cardiovascular mortality risk estimates, and for Seattle, only risk estimates for asthma hospital admissions were possible.”

P. 4-58, L. 6-11: In discussing the (sometimes inconsistent) results from multi-pollutant models, it might be important to emphasize that effects of confounding pollutants are not necessary causal. Certain gaseous pollutants are likely strong surrogates for certain source-related fractions of PM mass – for example CO or NOx for automotive PM, O3 for secondary PM, SO2 for SO4, etc. If data are available, it might be informative to show a few correlation matrices – perhaps with seasonal stratification - which include $PM_{2.5}$, PM_{10}, $PM_{10-2.5}$ and gaseous pollutants. I suspect that at some sites & seasons, some of the gasses might correlate with $PM_{2.5}$ as well as or better than PM_{10} does. I’m not quite sure what would cause the PM coefficient to increase when a second pollutant is added to the mix. Is there some logical explanation for why this might occur?

P. 4-60, L. 7-11: I don’t have a suggested alternate approach but also don’t think it’s really “likely” that background correlates (perfectly) with “as is” levels at any sites. Emissions modulations, locations and transport patterns are quite different between most natural, transcontinental and continental anthropogenic sources. On the other hand, I don’t think its likely that this (matching of percentiles) would interject a specific directional bias to the results, as it would tend to both understate and overstate background (or human) influences, and by equal proportions. This might logically have an effect of increasing the uncertainty bounds, but should not shift the means or signs of the results. Maybe you could just state this differently: "Background levels were assumed to correlate with as is concentrations. We assigned….in its distribution. While this procedure would unavoidably result in mismatched combinations of estimated background and non-background influences on a daily basis, these mismatches would not tend to introduce an overall directional bias to the results."

P. 4-64, L. 5: change "are" to "is".

Chapter 5 is clearly written and generally follows the material presented in the CD. The discussion of visibility includes several kinds of new material not presented in the CD, including: EPA staff analysis of ASOS visibility vs. continuous $PM_{2.5}$ (provides some additional indication of quantitative PM visibility relationships), summarization of various economic studies and local visibility standards (provides information on perceived adversity of visibility effects), and a series of photographs showing urban visibility at a range of $PM_{2.5}$ concentrations (which illustrate effects and might subsequently be employed in some sort of future "public" determination of "unacceptable" visibility levels – and might therefore be employed to justify a future visibility or PM secondary standard). Some of the older cited references for the valuation studies are not referenced in the CD, and perhaps a brief summary section could be added to the CD for consistency.

As indicated in previous comments on the 3rd and 4th drafts of the CD (attached here for clarity), I disagree with EPA staff position that pending new data and analyses provide a logical rationale for postponing consideration of a secondary $PM_{2.5}$ standard (or "pretending" that the current secondary standard of 65 ug/m^3 is a reasonable threshold for adverse visibility effects).

P. 5-5, L. 4-6: While this is likely a true statement, it is not really demonstrated anywhere in the staff paper or CD. Many previous analyses of airport visibility data have demonstrated excellent correlations with fine particle mass or species data, especially when constrained to daylight hours and when adjusted for RH effects. A good example is provided in Figure 4-37 of the 4th draft CD, based on human observer visibility, constrained to daylight hours, showing extinction efficiency of 4 m^2/g and published 25 years ago. The ASOS data are likely of better quality and more consistent across all hours of day and from site to site, but do not provide a logical reason for indefinite future delay.

P. 5-11, L. 3: add "rural" before "West". Mane-made contributions are much higher than 1/3 of total in western urban areas.

P. 5-12, L. 2: Direct optical measurements are not in fact used in "implementing air quality management programs to improve visibility" under the regional haze regulations and guidance – and for good reason. I wonder if EPA is trying to set up a whole new program of required transmissometer measurements (the last thing the states want or need to hear) as an excuse to avoid considering a fine particle standard? We already have nephelometers deployed as continuous $PM_{2.5}$ monitors and don't need new optical standards. Fine mass is an excellent indicator; we just need secondary standards.

P. 5-13, L. 1-18: The discussion of East vs. West differences in the IMPROVE results is accurate and useful. However, since these large longitudinal differences in reconstructed extinction derived for Eastern and Western IMPROVE sites are subsequently repeated in Chapter 6 (P. 6-43, lines 20-27) as partial justification for avoiding or postponing consideration of a secondary $PM_{2.5}$ standard (with a statement in the following paragraph (line 29) that "urban visibility remains poorly characterized at this time", I 'd like to offer several comments and observations. Use of IMPROVE data to draw inference about differences in Eastern and Western urban areas can be misleading, for the following reasons:

Class 1 areas are inherently remote and IMPROVE sites are intentionally located at high elevations within these remote areas. The elevational distribution of Western IMPROVE sites (west of 100 degrees long.) is substantially higher than at Eastern sites as shown in Figure 1 (based on IMPROVE sites and data for 2000 through 2002 extracted from VIEWS). Two thirds of Eastern sites are below 500 meters, while a similar fraction of Western sites are above 1000 meters. Thus, at least some of the East/West differences in reconstructed extinction at these sites may be a function of elevation rather than longitude.

Figure 1. Cumulative distributions of Eastern and Western IMPROVE sites by Elevation

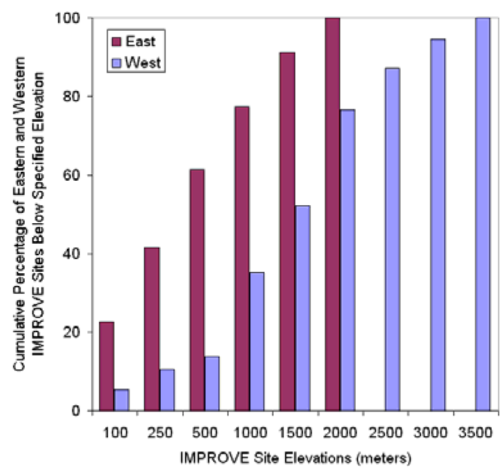
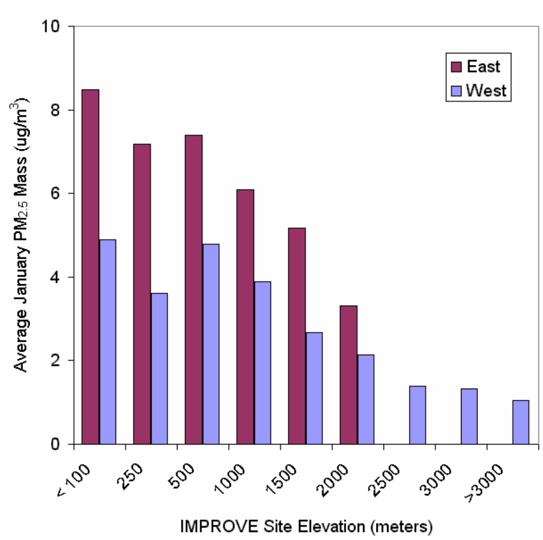
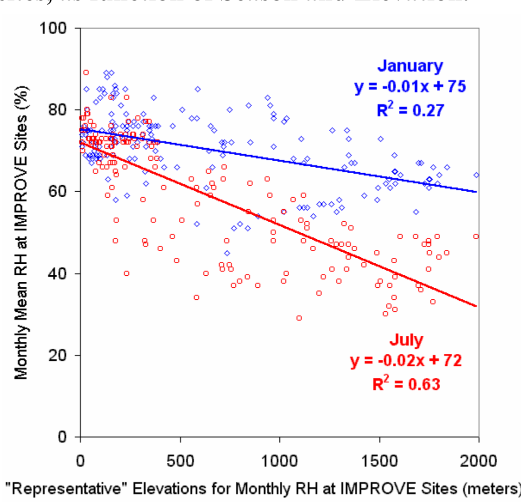


Figure 2. Winter PM 2.5 Mass at Eastern and Western IMPROVE sites by Elevation



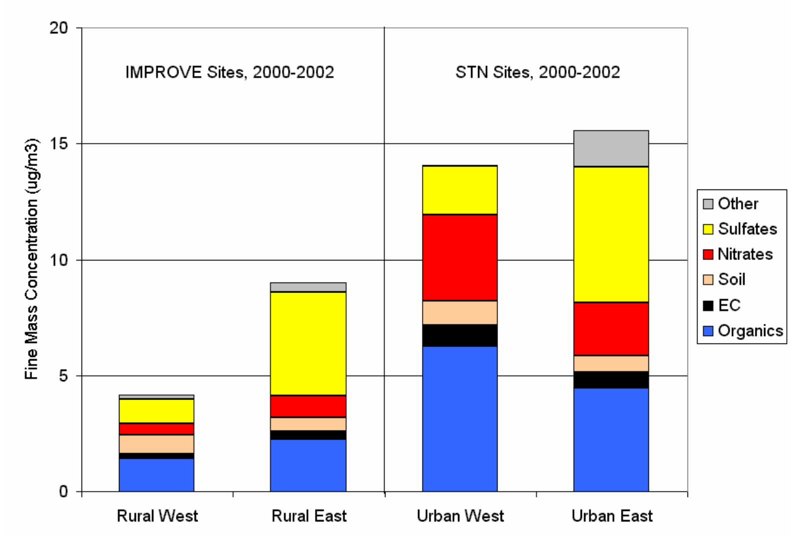
Higher elevation sites are more frequently above the boundary layer, especially during the winter months, and so will tend to experience lower aerosol concentrations, as indicated for both Eastern and Western sites in Figure 2. There is also a general decrease in relative humidity with elevation, as indicated in Figure 3, especially during the summer when both mixing depths aerosol sulfate concentrations are highest.

Figure 3. Monthly Mean RH at IMPROVE sites, as function of Season and Elevation.



Thus, Western IMPROVE "reconstructed extinction" estimates will tend to be lower than Eastern estimates purely as a function of increasing elevation and the associated decreases in aerosol concentration and RH.

Figure 4. East/West Rural/Urban differences in $PM_{2.5}$ Mass and Composition based on IMPROVE and EPA STN Data 2000-2003



Descending from IMPROVE mountaintops to the lower elevation valleys where human populations (and cars, trucks, woodstoves and industries) are concentrated, the East/West differences are substantially diminished. Figure 4 is based on averages of all available IMPROVE and (urban) EPA STN data extracted from VIEWS for the 3-year period 2000 through 2002. These data were screened to include only days when $PM_{2.5}$ mass and all major species were present, and further limited (as a QA screen) to samples where the reconstructed mass was within $\pm$ 50% of the measured mass. While the remote IMPROVE data indicate Eastern concentrations more than double those in the West, the urban STN data suggest much more similar concentrations. Urban concentrations of sulfates (assumed $(NH_4)_2SO_4$) and "other" (unspeciated mass – most likely particle bound water) are higher in the East, but organic matter, EC, soil and nitrates are higher, on average, at the western urban sites. Thus, the effects of fine particles on urban visibility are likely to exhibit much less variation from East to West than that indicated by the remote IMPROVE data.

Additional exaggeration of East/West differences are related to the way in which the hygroscopic f(RH) growth functions are applied according to the EPA regional haze regulations and associated guidance. A key feature of the haze regulations is that visibility protection is considered important at all times of day in these otherwise pristine parks and wilderness areas. Thus the hygroscopic growth functions that enhance reconstructed extinction attributed to sulfate and nitrate compounds are based on (climatologically derived) distributions of all RH conditions (below 95%) that are encountered at these sites. Because of the strongly non-linear nature of the hygroscopic growth curve, the monthly mean f(RH) growth curves, based on a linear average of the skewed individual combinations of RH and f(RH) ends up heavily weighted by the most humid hours, which tend to occur predominately ay night. This effect is illustrated in Figure 5, where the plotted data points compare the monthly mean RH at each IMPROVE site for the months of January and July with the monthly mean f(RH) functions for these sites. The solid curve is the hygroscopic growth function upon which these monthly estimates are based. The higher f(RH) values for the monthly mean reflect the effect of this non-linear averaging. At the Lye Brook, VT IMPROVE site, for example, the average summer (JAS) f(RH) is 3.24 and the average RH is 74%. However, from the f(RH) growth curve, an f(RH) of 3.24 implies an actual RH of 84%. Figure 6 shows the long-term summer diurnal mean patterns in RH and (human observer) visual range from the Burlington, VT airport for summer (JAS, also limited to <95% RH). Average summer RH in VT (during periods for which precipitation and fog are eliminated by the <95% RH screen, as they are in the haze regulations) reaches a maximum of about 80% at night, but decreases to less than 60% during mid-day daylight hours, and the visual range approximately doubles accordingly. The f(RH) for IMPROVE reconstructed extinction, based on 24-hour protection is thus about twice as high as that which might be employed if visibility protection (such as from a secondary $PM_{2.5}$ standard) were limited to daylight hours only (as is the case for the Denver, Phoenix, Fraser Valley and California standards/guidelines).

Figure 5. Regional Haze Regs Hygroscopic Growth Functions

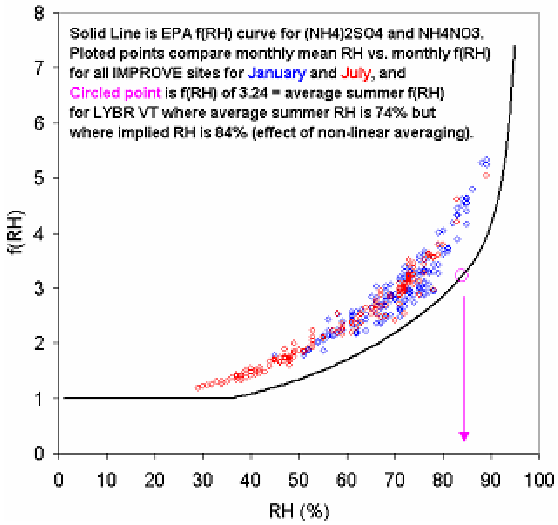
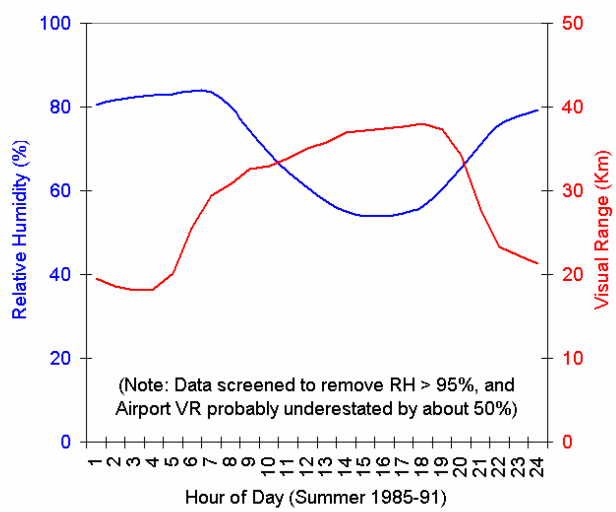


Figure 6. Burlington, VT Summer RH and Visual Range



If a secondary fine particle standard for visibility protection were limited to (6 or 8) daylight hours only, the net effect is to basically dry out the aerosol – or rather consider it only during the driest period of the day. This would tend to substantially reduce the East/West differences, and would also substantially tighten the $PM_{2.5}$ / visibility relationship. <u>The difference between $PM_{2.5}$ mass and light extinction lies almost entirely in the water present in ambient aerosols which is deliberately removed when we dry our FRM filters to 40% RH</u>. This effect can be illustrated by comparing the fine mass/ extinction relationships in Figures 6 and 7. Figure 6 is based on all IMPROVE data, 2000-2003, screened to include samples where reconstructed mass was within ± 50% of measured mass, and using the EPA monthly mean f(RH) functions. Figure 7 is based on the same aerosol data but includes f(RH) for assumed daylight-only conditions.

Figure 6. IMPROVE Reconstructed extinction vs. mass using 24-hr f(RH)

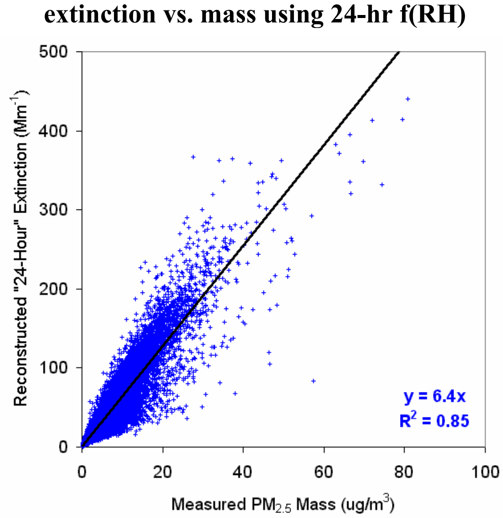
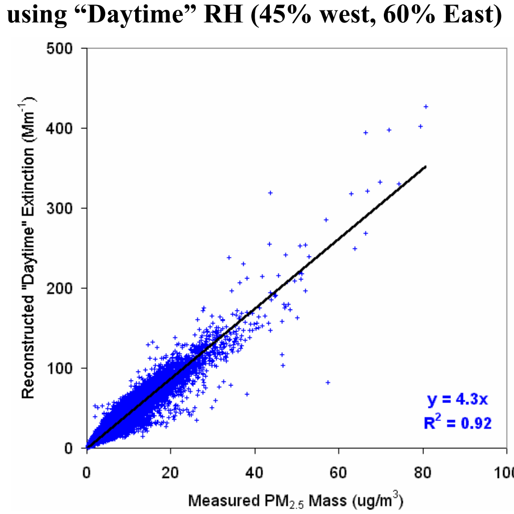


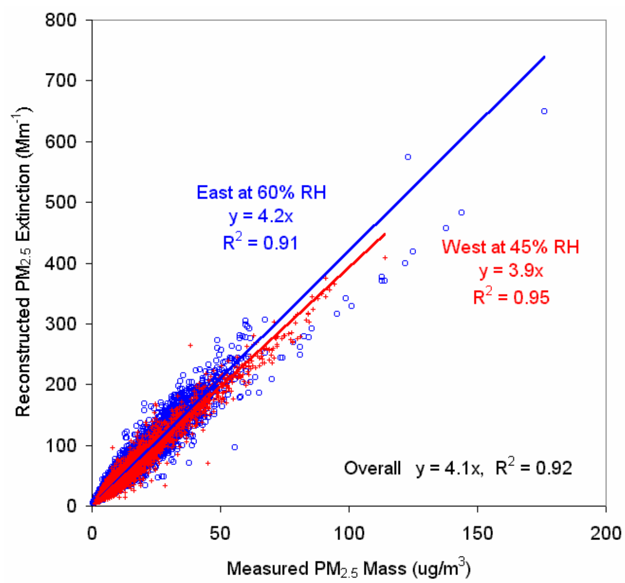
Figure 7. Reconstructed extinction vs. mass using "Daytime" RH (45% west, 60% East)



Even with the EPA monthly mean (and heavily nighttime influenced) f(RH) growth functions, a strong, fundamental extinction/mass relationship is clearly evident, and could be expressed as $b_{ext} = 6.4 \pm 3$ x $PM_{2.5}$ mass, where 6.4 is the implied extinction efficiency of dry fine particles to wet ambient fine (and coarse) particle extinction in units of m^2/g, and would be decreased to 6.0 if the effect of coarse particles were removed). The higher extinction efficiencies (>8 m^2/g) are predominately from humid, high sulfate Eastern sites, and the lowest efficiencies (< 4 m^2/g) are predominately from arid Southwestern sites. Using "daytime-only" f(RH), based on assumed mid-day RH of 45% in the West and 60% in the East, the extinction/mass ratio decreases and tightens to a more constant value of about 4 m^2/g and the R^2 increases from 0.85 to 0.92.

A similar fine mass extinction efficiency of about 4 m^2/g is obtained when IMPROVE reconstructed extinction formulae are applied to the urban EPA STN data, and when f(RH) functions are based on assumed mid-day humidity levels of 45% in the West and 60% in the East. With this daytime-only constraint, the differences between Eastern and Western urban sites are substantially reduced and the $b_{ext}/PM_{2.5}$ relationship tightens. It might also be noted that this relationship of 4 m^2/g was also indicated in Figure 4-37 of the 4th draft CD, based on daytime-only data, published 30 years ago. Another 5 to 10 years to further study the new ASOS data is not likely to change or improve this relationship.

Figure 8. Reconstructed extinction vs. mass at STN sites using assumed daytime RH of 45% West and 60% East



As indicated earlier, the 3 local visibility standards summarized in the CD are all based on daytime-only observations. In addition to minimizing effects of fog, mist (heavily hydrated hygroscopic aerosol) and other natural influences that may cause or interact with pollutants to impair visibility, there are other logical reasons to consider a daytime-only averaging time for a secondary visibility standard. Nighttime visibility is less important in urban areas. Unlike wilderness areas, there are fewer campers sleeping out under the stars and urban light pollution substantially diminishes urban views of distant objects. Its dark, there's not much to see except lights, and most of us are indoors with our eyes closed and asleep. A sub-daily secondary standard would also focus more attention on the quality of continuous hourly data, which are currently adjusted to be FRM-like only on the basis of their aggregated 24-hour totals, while the quality of the hourly data remains unclear. If "boiling off" of volatile organics and nitrates results in a substantial FRM adjustment during the cooler seasons, it is probable that a seasonal adjustment would tend to under-adjust the hourly data at night and during the AM rush hour (maximum volatiles) and over-adjust the mid-day data when volatile losses are least (and/or when non-volatile sulfate concentrations are highest). Such errors could have direct implications for mischaracterizing short-term health effects from specific source categories. Thus a sub-daily secondary standard would both focus on the time periods when hourly data are most accurate and also cause needed scrutiny of the continuous data during other parts of the 24-hour day.

P. 5-14, L. 5: "diesel" could be changed to "motor vehicle" (as urban gasoline vehicles also contribute to all of the above).

P. 5-14, L. 16: could change to "…these truncated data are not ideal…" (they are still quite useful during severe events when VR < 10 miles.

P 5-15, L. 20-22: I would think comparisons between ASOS and continuous $PM_{2.5}$ would be a much more useful future exercise than proposed comparison with (daily) STN (assuming EPA actually had an interest in a secondary $PM_{2.5}$ standard). Such comparisons might also employ some of the ASOS QA, RH screening and RH adjustment methods used by Husar in the NOAA report I cited in last CD comments (Husar, R.B. (2002) Evaluation of the ASOS Light Scattering Network, Final Report to J. F. Meagher, NOAA Aeronomy Laboratory R/AL, Boulder Colorado.). Also (as evidenced in attached figure from that report) the need for RH screening and adjustment would be substantially minimized if the comparison were limited to the hourly data from daylight hours only (not that this future analysis is a necessary precondition to considering a secondary standard).

Figure 9. RH screening and RH adjustments of ASOS Visibility data from Husar (2002)

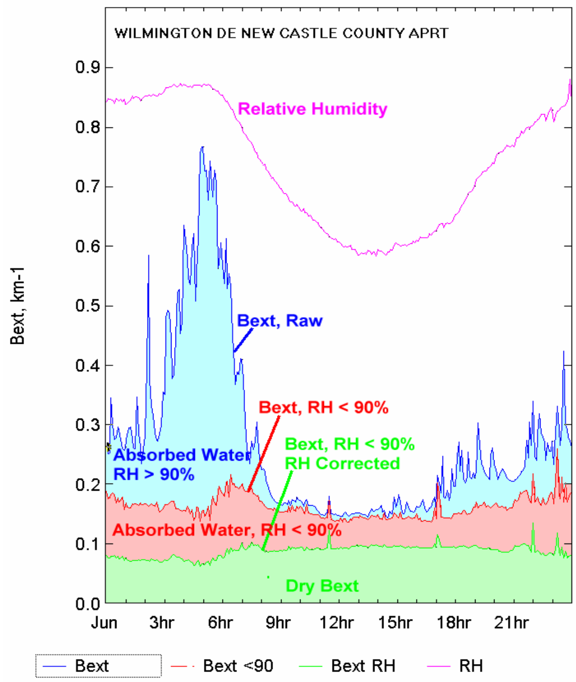
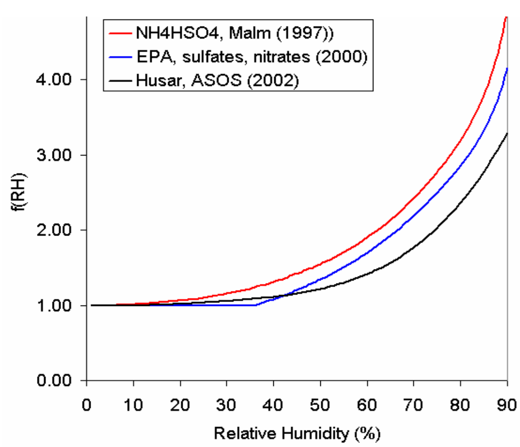


Figure 10. Comparative Hygroscopic Growth Functions for sulfate compounds & fine mass



The "RH correction factor" employed in the Husar (2002) evaluation of ASOS data, empirically derived for the specific purpose of relating ASOS bext to $PM_{2.5}$ mass is reproduced in Figure 10, along with the EPA f(RH) curve recommended (for ammonium sulfate and ammonium nitrate) in the regional haze guidance, as well as a similar curve from Malm (1997) for more acidic ammonium bisulfate. Compared to the EPA curve, the water uptake is slightly more rapid for the more acidic species and slightly less for Husar's generic ASOS curve. But Husar's curve was developed for the specific purpose of "drying" the ASOS visibility data for comparison with nearby continuous fine mass data (also artificially dried), without regard to chemical aerosol composition. This provided the "best fit", and in light of the excellent adjusted bext vs fine mass relationships at sites throughout the country (submitted in my comments on the 4th draft CD) provides additional evidence of the strength of the PM $b_{ext}/PM_{2.5}$ relationship. When (or if) future EPA ASOS analyses are conducted, some of the things we will learn are that many of the ASOS sensors are poorly calibrated and maintained, that there are numerous errors in the data archival process, and that the ASOS "Belfort model 6220 forward scatter visibility meter" does not respond to effects from back-scattered light – and so differs, sometimes strongly from what the human eye might perceive. Hence, NOAA warns pilots that "If conditions are bright enough for a pilot or a controller to use sunglasses, you can expect the automated systems to report visibility approximately twice what the human eye perceives. If an ASOS observation reports a 4 mile visibility, you can expect a report of around 2 miles by a human observer." (http://www.nws.noaa.gov/asos/vsby.htm). What we will not learn is that there is some new, previously unknown relationship between fine particles and visibility. That relationship was well established > 30 years ago.

P. 5-16, Fig. 5-2: Note that the relative absence of any relationship with longitude provides additional support to my argument that east/west differences in urban VR/PM relationships are not as large as for remote sites.

P. 5-17, Fig. 5-4: I suspect that the "Diagonal line indicates (the 1:1 line, not) the regression line".

P. 5-21, L. 13-19: In addition to these summarized visibility standards, you could also add a Vermont State visibility standard, which was adopted in 1985 and expressed as a summer seasonal sulfate standard of 2 ug/m^3. This standard was established to represent "reasonable progress toward attaining the congressional visibility goal for the class 1 Lye Brook National Wilderness Area, and applied there and also to all other areas of the state with elevations > 2500 ft. At that time, average VT summer sulfate levels were about 4 ug/m^3, and attaining that standard (assuming other pollutants remained the same) would have increased visual range from 40 to 50 km. You might also mention here that the Lake Tahoe standard is an 8-hour standard, constrained to RH < 70%, and equal to 30 miles (48 km). Thus there are 2 additional areas (in addition to Denver, Phoenix and Fraser Valley) where visual range < about 50 km has been locally judged to be unacceptably adverse. At a (daytime) fine particle extinction efficiency of 4m^2/g, this visual range translates to a $PM_{2.5}$ concentration of 17 ug/m^3 (or a bit lower if small effects from coarse particles and NO2 absorption were also considered). Thus there is very strong convergence about a judged level of adverse visibility at about the level of the current annual $PM_{2.5}$ standard, but applied to much shorter 8-hour (6-hour in Phoenix) daylight averaging times.

P. 5-29, Figs. 5-26 and 5-28: I note that there is no Fig 5-27 and suggest that you include one with (much) lower PM concentration. Else you tend to give the impression that (EPA's arbitrary judgment of) a potential secondary standard might lie somewhere between 30 and 65 ug/m^3 (2 to 4 times higher than the equivalent levels already determined in Denver, Phoenix, Fraser Valley, Lake Tahoe and VT).

P. 5-29, L. 15-16: "…EPA hopes to pursue [a more extensive photographic visibility valuation survey] in the future [although it has committed no resources] to help inform the next periodic review of the PM secondary standards". This (and the future ASOS research excuse) sounds mighty similar to EPA's optimistic claim of 25 years ago that "recently initiated research efforts…and progress in evaluating visibility values could provide support for a decision on the desirability of [a secondary fine particle] standard by 1982 or 1983" (EPA 1979 Report to

Congress on Visibility). If or when EPA commits the resources to such future studies, I think they could be informative – and might lead to refinements to a secondary standard, but the nature of scientific enquiry (even for "the dismal science" of economics) is such that there will always be ideas for future research. There is no air pollution/effect relationship that is currently understood nearly so well as the relationship between fine particles and visibility, and there is no standard nearly as inadequate as the current secondary short-term $PM_{2.5}$ standard to protect against effects which are clearly adverse at the level of that standard.

Rather than another 25 years of (unfunded) research, I strongly recommend EPA consider proposing a short-term, daylight-only, secondary $PM_{2.5}$ standard. Evoking (from 1st page of my general comments) Karen Martin's logic for the PM coarse standard (that no action is equivalent to endorsing the PM_{10} status quo) and Mort Lippmann's suggestion that in consideration of the (PM-coarse/health effect) uncertainties a standard might be set at a somewhat lenient level, I suggest considering a secondary $PM_{2.5}$ standard in the range of 25 to 35 ug/m3. This would equate to a visual range of about 25 to 35 km and would be about twice as stringent as the current 24-hour standard (and/or the 50 ug/m^3 upper range of proposed lowered primary short-term standard) and twice as lenient as the implied level of $PM_{2.5}$ at the many cited locations where a visual range of less than 50 km has been judged to be adverse (and 10 or more times more lenient than the implied level required over time for protected class 1 areas). It would also be consistent with EPA's observation on P. 6-45, lines 1-3 that "appreciable improvement in the visual clarity of the scenic views…occurs at concentrations toward the lower end [30 ug/m3, 24-hour] of the staff-recommended range of consideration for the 24-hour $PM_{2.5}$ standards" and (lines 7-10) that "revisions to the primary standards… would afford greater visibility protection, especially toward the lower end of the staff-recommended changes for the primary standards." In other words, staff accepts that a standard toward the upper end (45-50 ug/m^3) of the proposed short-term primary standard would not afford much in the way of visibility protection.

Uncertainties may (and will likely always) persist regarding a precise level of visibility impairment considered adverse by different observers under different lighting conditions for different scenes, at different locations, but those uncertainties are of a much different nature than those which relate to primary health effects, in that a direct causal mechanism is clearly understood and, for all practical purposes, there is (almost) no lower bound $PM_{2.5}$ concentration at which a PM change will not result in a perceptible change in visual air quality. But this true "no threshold" is not a logical reason for no standard. Substantially greater uncertainties (or flexibilities) are introduced by (EPA discretionary) variables like the compliance metrics, compliance dates, and implementation schedules associated with any secondary (or primary) standards.

One final point (whew) and associated recommendation that I want to raise in regard to a suggested focus on daylight-only (6 or 8 hours), is that it may well be a more stable metric, and not necessarily a more stringent metric than a 24-hour average (see Figure 11 for example), as it would tend to avoid the extreme and variably high concentrations that tend most frequently to occur over night and in the early morning hours. I think it would be a relatively simple data processing exercise for EPA (especially since it currently has relatively unique access to AQS data) to develop some comparative evaluations based on available continuous $PM_{2.5}$ data. A few different definitions of daylight-only could be selected, and calculations could be made at multiple sites for the frequency at which various thresholds (say 15, 20, 25,…65ug/m3) would be exceeded. I'm not sure whether these should be based on "FRM-adjusted" or "non-adjusted" data (maybe both). An example of this kind of analysis (from an older batch of continuous data I extracted a few years ago, quality or adjustment status unclear to me) is displayed as Figure 12. I think this kind of data exercise could be helpful to EPA (and others) in considering an appropriate level of a secondary standard (and for other reasons), and that it could also be done fairly quickly.

Figure 11. Example Average Diurnal $PM_{2.5}$ concentrations at selected urban sites

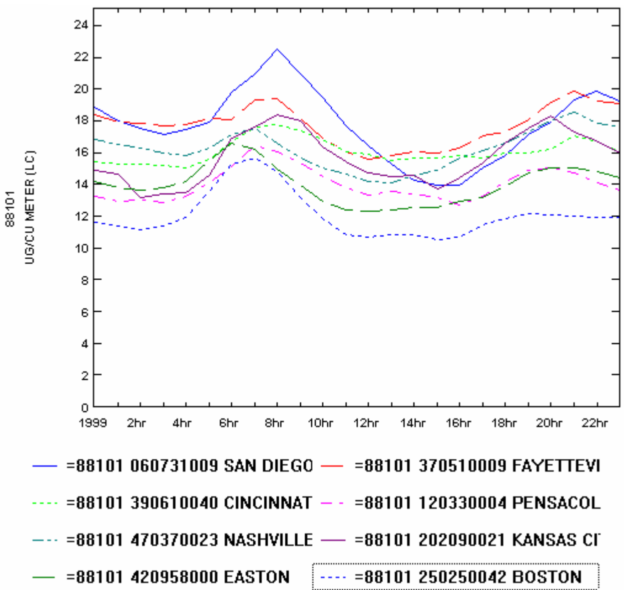
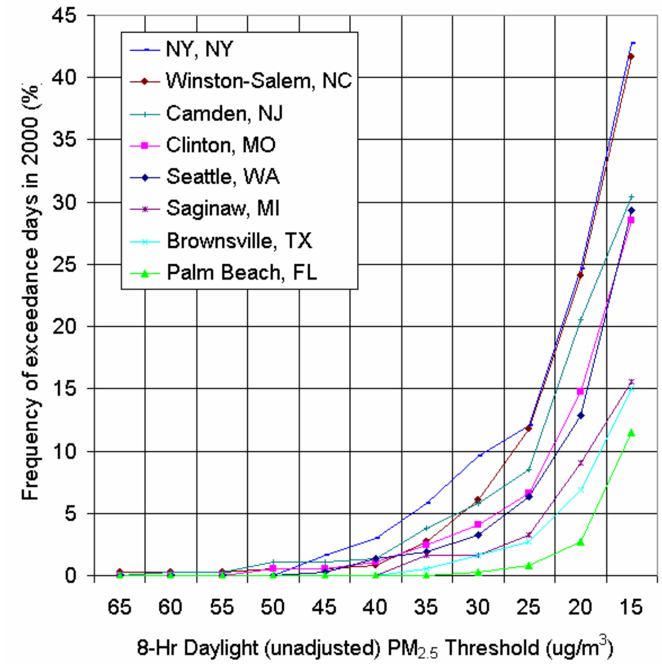


Figure 12. Example suggested EPA Analysis of Alternate 8-hour daylight PM2.5



Note that there appears to be a fairly distinct "inflection point" or shift in the curves in Figure 12 in the range of 25 to 35 ug/m^3 where a change in threshold represents a large change in the frequency of exceedance. A more detailed evaluation of the currently available continuous $PM_{2.5}$ data would put EPA in a good position to carefully consider options for both the level and compliance metrics for a secondary standard.

Chapter 6 is generally clearly written and the recommended ranges for revised standards seem well justified. A few exceptions include:

P. 6-4, L. 27-29: "…newly available short-term exposure studies that provide evidence of … associations with $PM_{2.5}$ in areas with air quality below the annual standard…" Yet despite indications of short-term effects, EPA continues to advocate the annual standard as a "controlling standard". There is bound to be conflict between annual and short-term standards if it is also required that the annual standard always be more stringent, and this "logic" would appear to preclude the possibility of a short-term standard set toward the lower end of the recommended

30-50 ug/m3 range (and therefore also preclude revisions to the primary short-term standard that would offer any protection against adverse visibility effects).

P. 6-21, L. 6-14: Although there may be some benefits to spatial averaging, providing the (currently available) option of using it or not for compliance determination is equivalent to offering a choice among alternative standards. Possibly EPA could conduct some analyses that would allow spatial averaging but with a somewhat stricter compliance metric (say 99^{th} percentile for 24-hr, or 15 ± 1 for annual) that would make a spatially averaged compliance metric more "equivalent" to the highest monitor alternative.

P. 6-40, L. 15-16: This recommended range of 30 to 13 (and subsequent range of 75-30 short-term) for $PM_{10-2.5}$ is so broad that it may fail to provide useful information to the Administrator. Possibly in future drafts this range could be described as one for which some justification could be provided within these extremes. But for a staff "recommendation", a smaller range would seem appropriate. I have no opinion on what that range should be, but don't think the justification for the low end(s) is especially strong.

PP 6-41-44: As previously suggested (a few times) I disagree with the avoidance of considering a secondary $PM_{2.5}$ standard, and hope EPA will reconsider. Two statements that seem illogical or contradictory (or I don't understand them) are:

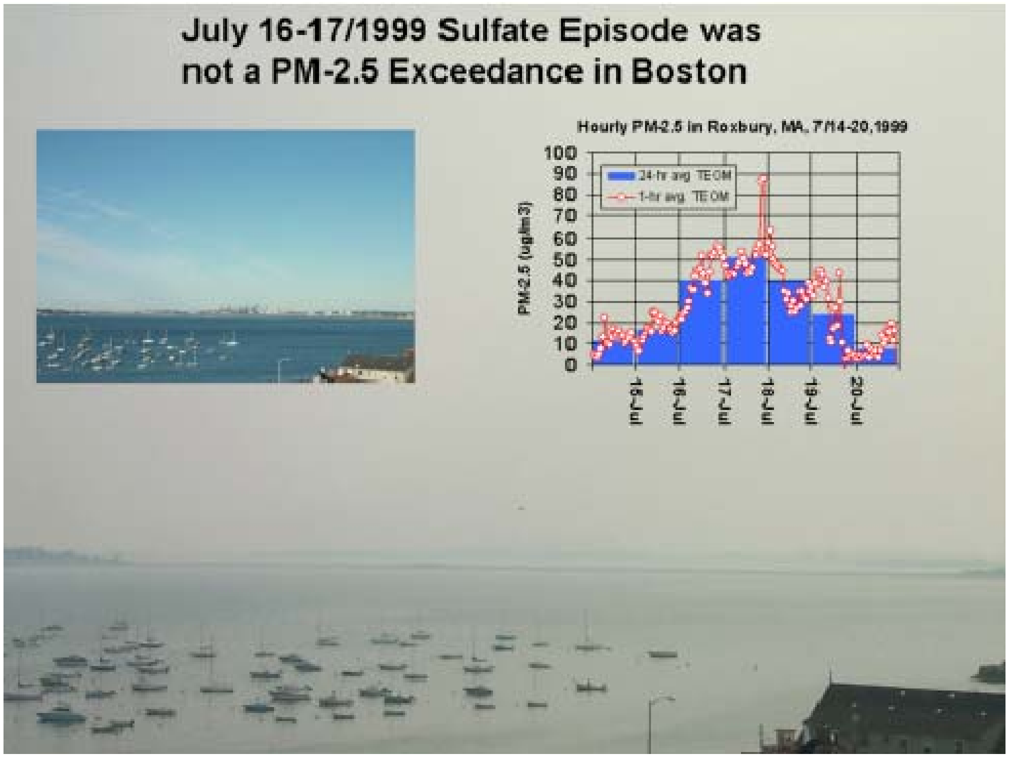
P. 6-41, L. 26-28: (In 1977), "EPA determined that an approach that combined national secondary standards with a regional haze program was the most appropriate way to address visibility impairment". Good idea; we now (finally as of 1999) have a regional haze program but not the secondary standards. In fact, the promulgation of the regional haze regulations removes an important previous obstacle in establishing a secondary standard, as it "covers" the class I area requirements and therefore "frees" the secondary standard to be applicable (or "controlling") only in non-class I areas.

P. 6-43, L. 11-14: "staff continues to conclude that PM, especially in the fine fraction, produces adverse visibility effects in various locations across the country, and that addressing visibility impairment solely through setting more stringent national secondary standards would not be appropriate." Is staff suggesting that a new secondary standard to address some of these "adverse visibility effects in various locations across the country" would somehow require a repeal of the regional haze regulations? If not, then where does the "solely" come from? I don't get this logic, and have also offered previous comments on what I think are flaws in the "justification for no action" bullets that follow on lines 20-27 and 28-34 and P. 6-45, L. 1-3. I also question the P-6-45, L. 13-14 position (excuse) that "local programs continue(s) to be an effective and appropriate approach…" Such local programs are always an option for additional levels of visibility improvement, but only in locations like Denver, Phoenix, Fraser Valley and Lake Tahoe where local emissions cause a substantial fraction of the local impairment.

Attachments: Figures and Associated Text from Comments on 3rd and 4th Draft PM CD

1. Illustrations of Impaired Visibility from Supplemental Comments on 3rd Draft PM CD: Since visibility impairment is an instantaneous effect of PM (and gases and weather) and since the anthropogenic effects are dominated by effects on light scattering and absorption by fine particles, short-term relationships between $PM_{2.5}$ and visibility would appear to provide the most logical basis for considering a secondary PM standard. A 24-hour averaging time might be appropriate (especially since few comparative PM and extinction data are currently available to consider shorter averaging times). Since the current primary 24-hour $PM_{2.5}$ standard of 65 ug/m^3, 98th percentile is both rarely exceeded and extremely hazy, it can be argued that virtually the entire distribution of adverse $PM_{2.5}$ effects on visibility lies beneath the level of "protection" provided by the current standard. An example of this is shown below in the 3 PM CAMNET photo from Burlington, VT during the recent 7/7/02 Quebec forest fire transport event. Peak hourly $PM_{2.5}$ exceeded 100 ug/m^3 and minimum airport ASOS visibility was in the range of 2 to 3 miles, but 24-hour $PM_{2.5}$ concentrations were less than 65 ug/m^3.

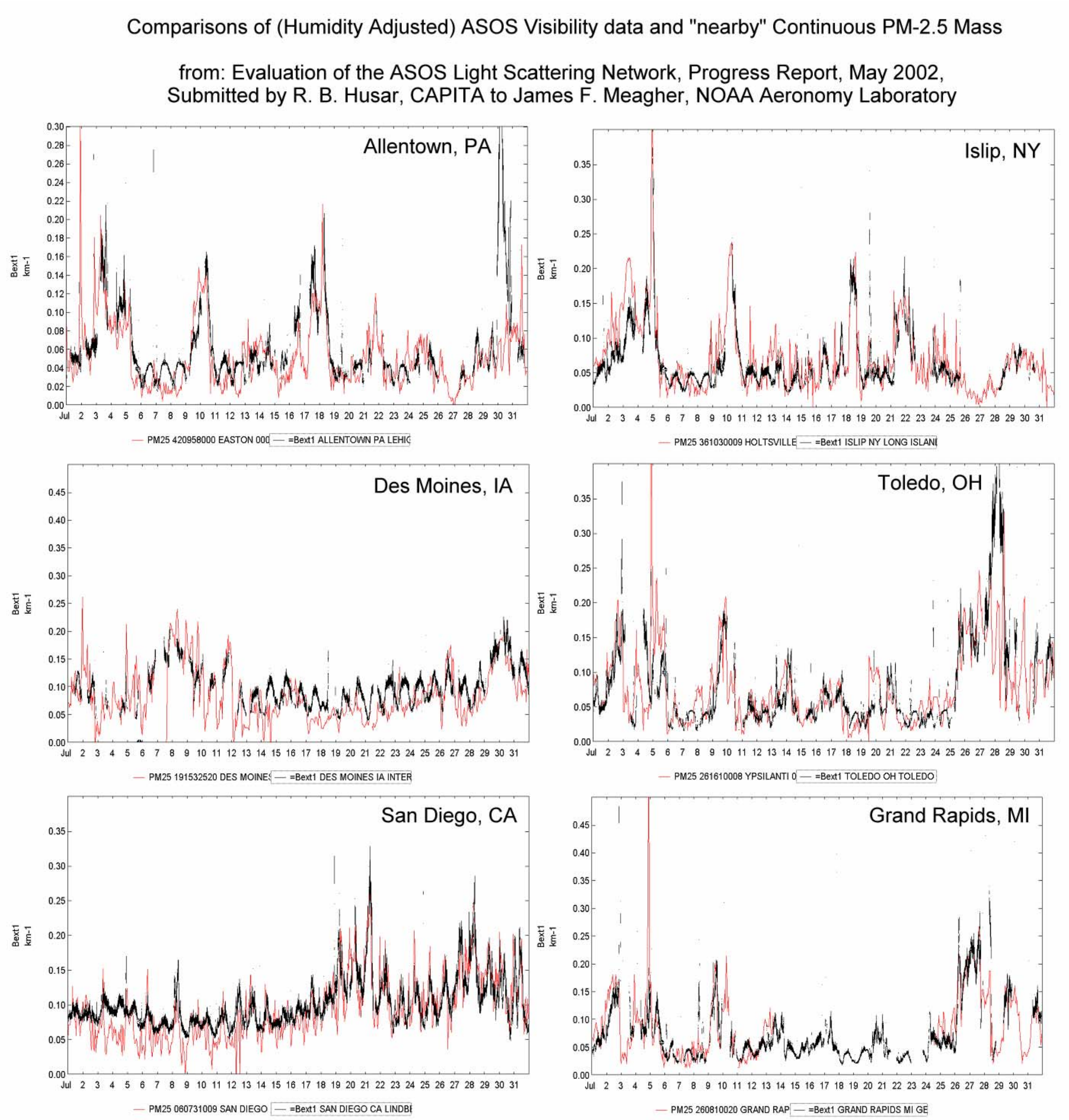
A second illustration below shows a similar effect on visual air quality in Boston, MA during a regional fine particle sulfate episode centered on July 16-17, 1999. The CAMNET photo is from 5 PM on 7/16/99 when hourly $PM_{2.5}$ concentrations from nearby Roxbury, MA and HSPS were averaging about 50 ug/m^3, with daily mean concentrations of about 40 ug/m^3, increasing to about 50 ug/m^3 on the following day. No exceedances of the 24-hour $PM_{2.5}$ were recorded in New England during this event, yet peak hourly extinction data from sites like the Great Gulf, NH IMPROVE nephelometer, and Burlington, VT and Martha's Vineyard, MA Airport ASOS were in the range of 400 to 800 Mm^{-1} - an implied visual range of about 3 to 6 miles.



Beyond the aesthetic effects of such haze episodes, there are also potential impacts on airport operations and aviation safety. NWS ASOS currently report visual range of 10 miles (roughly equivalent to 30 deciviews and quite hazy) or greater as a single "unlimited visibility" category, but provide higher resolution information when visibility is "limited" to less than 10 miles. "Haze" is reported at 7 miles or less. "Visual flight rules" (VFR) apply at 5 miles or greater, "marginal visual flight rules" (MVFR) at 3 to 5 miles, and "instrument meteorological conditions" apply at less than 3 miles. In an analysis of 1990-1997 general aviation accidents in the NTSB database, Goh and Wiegmann (1991) noted that while only 2.5% of reported accidents (about 50 per year) involved transition from VFR to IFC conditions, 72% of those VFR to IFC accidents were fatal, compared to a 17% fatality rate for aviation accidents overall. The 7/16/99 fatal crash which took the lives of JFK Jr and passengers occurred 4 hours after the above Boston CAMNET photo was taken. Other pilots in the area reported extremely poor visibilities in the range of 2 to 4 miles, although ASOS reports all along the flight path indicated "visual flight rule" conditions (> 5 miles), and the weather observer at nearby Martha's Vineyard estimated > 10 mile visibility a few hours earlier. Hence impaired visibility (identified as a key contributing factor in the NTSB review of this accident), can have adverse consequences even when "visual flight rules" apply.

2. Illustrations of RH-Adjusted bext vs. $PM_{2.5}$ Mass from Comments on 4th Draft PM CD

The following figure, taken from a recent evaluation of (raw, uncensored) ASOS visibility data by R. B.Husar compares "humidity-adjusted" light extinction with continuous $PM_{2.5}$ mass measurements for 6 sites in different regions (which presumably have different aerosol compositions). The RH adjustment involved screening out periods of humidity > 90% and then applying a generic (inverse) f(RH) function – based on an assumption of consistent hygroscopicity in the aerosol mix at all sites. The $PM_{2.5}$ data are not collocated, merely in the same urban areas, yet the correspondence is remarkably strong at all sites.



Dr. Frank Speizer

MEMORANDUM

To: Fred Butterfield

From: Frank Speizer

Subject: Comments on First Draft of Staff Paper dated August 2003

Date: November 21, 2003

Although I was one of the first to comment that putting out a first draft before the CD was finished was a mistake, I must say that I was pleasantly pleased as to how well the discussion of this draft went at our meeting in RTP 10 days or so ago. In general I thought the discussion was helpful and I certainly came away with the thought that the next revision will be all the better for it.

I have several specific comments, many of which were handled at the meeting, which are presented below, but I wanted to make some general comments first that Staff will need to take into account as they proceed to respond to both the comment made by committee members as well as from the public commentary.

First there is a significant difference between heterogeneity and variability in assessing results across study sites as well as by study design. The fact that results do not appear to be uniformly consistent is probably a strength in the data rather than a weakness. (For a thorough discussion of this see Savitz's new book Interpreting Epidemiological Data, 2003). Because these studies are done with variable degrees of like data sources one expects variable results. It is also inappropriate to selectively assess studies of multiple cities to pull out communities that are positive or negative to make specific points when these studies were not designed to be analyzed in this fashion.

I also think it is important not to simple forget all that came before the last 5 years, but also not to be suggesting that we have learned nothing new in this last cycle of 5 years and forcibly be wedded to explaining where we were 5 years ago. There may be a bit of revisionist history to say that we had effectively summarized the committee position 5 years ago with the George Wolff chart, when in fact that chart was imposed on the committee by the chairman's leadership, and as I recall there was considerable uncertainty in first completing the chart and certainly there was not a uniform agreement as to how it would be used. Finally, significant progress in both methods and analyses have been completed in the last 5 years, which overall have leant further credence to the previously findings that were criticized and thus have move the bar, albeit not all the way, along the continuum toward less uncertainty, particularly for PM 2.5. As discussed there were acceptable data for PM 10 five years ago, and therefore in moving to a PM 10-2.5 we must keep in mind that it must take into account the previous data and not be judged solely on the basis of what all would agree is a limited measured 10-2.5 data base.

Specific Comments

Chapter 2

Additional discussion might also be included in table 2.2 with regard to course fraction in that by sizing the course fraction as >1.0 um may lead to confusion when discussion turns to PM 10-2.5.

I think additional discussion of ultrafine particles is needed. This needs to be described where measurements have been made, particularly in association with ambient PM measurements. This is particularly important as background to the discussion on pp 2.59-61 where discussion of relative penetration indoors of PM of ambient origin takes place. The lack of data, if this is the case, should be specified.

A good example of where this is necessary is on page 2.61, line 24 where it is stated light absorption from black carbon is relatively small component of PM …. This might be true by mass but may not be true by number and I suspect we just do not have sufficient data to be sure.

Chapter 3

Table 3.1==Consider expanding to include column which identifies cellular, animal, and human data with indication of multiple sources. Alternatively, consider columns that expand to include sources of data as basic science, toxicological (animal and human), and epidemiological.

Although on page 3.14 the authors indicate that some 80 time-series studies are reported the numbers alone are not important. Enough have been done. However, looking at table 3.2 it is clear that none of the PM10-2.5 are significantly different from null and this will become a problem. Argument on page 3.29 and Figure 3.6 is weak for PM 10-2.5. Conclusion from data not justified. Will need to see revisions in CD.
(Mort Lippmann's suggestion of dealing with this lack of data by making regional specific adjustments to PM 10 is a good one, and certainly if carried out will make me more comfortable that there will be sufficient data to come up with a range of effects for course particles). Section 3.31 essentially discusses total mortality, covers particle size and concludes on life shortening. Seems weak mostly because introduces role of infant mortality, the data of which are weak at best.
A RANDOM THOUGHT: Discussion points out the lack of consistency in the PM10-2.5 data but really it are a lack of data mostly. Facts are that we see consistency and coherent data for PM10 and PM2.5 and only consistent data for respiratory effects for PM10-2.5. What does it mean? Lack of studies or all really effects of fine, since PM10 contains fine. PM2.5 contains fine and PM10-2.5 does not. Not prepared to conclude on PM10-2.5 at this point in time. However, by following thorough on Mort's suggestion may actually have sufficient data. Need for a better meta analysis on page 3.52 rather than vague statement "less frequently statistically significant …."

Chapter 4

Page 4.13, lines 8-15: In Chapter 4 discussion of important parameters of outcomes not clear where asthma exacerbations are counted. ER visits may b e more important (and more easily counted) than respiratory symptoms. Not clear that should leave out AHSMOG and Veterans' Study. May need to specify limitations.

Concern that presentation of sensitivity analysis uses too high levels of PM2.5 for background. Given the range of measured as is. Where as the presentation of up to 30ug/m3 results in substantial reductions in risk estimates and narrower (and more realistic range of backgrounds) would suggest not much savings over getting to background and would negate the argument as to whether there is a threshold level.

Will be interesting as to how next draft of SP deals with this for PM10-2.5, given little real data.

Chapter 5

Very little discussion of size distribution of particles. This needs to be explored by those in the know as I can imagine that source of particles may make some more corrosive than others.

Chapter 6

Page 6.17 Need to consider adding the changing demography of the population. With increasing numbers of older people more people with health problems and therefore greater numbers of "susceptible groups" that will mean greater numbers of ER visits, admission and eventually greater morbidity and mortality attributable to pollutants at existing levels.

Page 6.27 I do not understand why "staff now concludes that PM10 is not an appropriate indicator for such a standard" This seems to have been a court decision rather than a scientific or risk assessment decision.

Need to have some data on proportion of PM10 is really PM10-2.5 and just how regional this proportion is. Would like to have figure that follows table 6.2 that gives proportion of PM10 that is PM10-2.5.

Need to raise discussion on page 6.51 on what happens to soiling from PM10 if we drop to PM10-2.5.

Summary table in appendix impressive in the lack of significant effects in the PM10-2.5. However. This does not mean that we do not need a course particle standard.

I suspect the actual number range will become clearer in next draft.

Dr. George E. Taylor, Jr.

Review Comments
OAQPS Staff Paper – First Draft

George Taylor
Professor and Associate Dean
School of Computational Sciences
George Mason University
Fairfax, Virginia

16 November 2003

This review is limited to major issues of concern. There are numerous issues that will be self correcting with editorial attention in the future, and none of these are outlined.

Introduction (Chapter 1)

1.3 Approach. This section is not adequate from my perspective. The discussion is really on objectives rather than a lucid and articulate statement of the risk assessment approach. This should be a major part of the game plan to lay out the approach to risk assessment for human health and ecology (common ground of a general risk assessment) and then to allow the separate sections to discuss the specifics of the approach for each (human health and ecology/natural resources). This omission is a major concern.

There is no mention of the parallel section to that of human health risk assessment to that for ecology and natural resources; I trust this was an oversight. I would recommend that a separate statement be added to emphasize this approach.

Air Quality Characterization (Chapter 2)

This chapter presents information on air quality with respect to human health, climate and visibility.

It seems odd that there is no effort to link the air quality herein with ecology and natural resources. Is it that there is no interest in doing so within the Agency or is it that the ecological community at the Agency does not converse with the air chemistry group. As above, this is a serious omission.

In section 2.5, there is discussion of the utility of using "regions" for analysis of trends and these are shown in Figure 2.3. There is a statement made about the value of using these regions. What data do you have that argues that the selection of boundaries is any better or worse than another? Many of the boundaries are strictly geometrical and do not appear to have a consistent biogeographical or climatological underpinning. For example, is the Northwest really inclusive of the broad geographical and climatologically dissimilar landscapes? Certainly the onshore flow on the Washington coast produces a far different background PM than that in Montana.

Characterization of Health Risks (Chapter 4)

This section provides a nice "game plan" for how the risk to human health is to be approached from a methodological standpoint. The chapter has generality and specificity, and the combination of the two is very helpful in helping the reader appreciate how the analysis is to proceed. I commend the Agency on using this approach and clearly articulating the protocol.

This chapter stands in marked contrast to the effort devoted to ecology/natural resources. The same approached used two decades ago for ecology/natural resources is used in this draft. This fails to capture the quality of ecological risk assessment that is state-of-the-art in ecology and the Staff Paper suffers from this omission.

Characterization of PM Related Welfare Effects (Chapter 5)

This chapter has many shortcomings that follow from the deficiencies of the CD that underpins this chapter. Some of the shortcomings can not be rectified, but it would be best not to perpetuate some of the shortcomings that are liabilities. Most of these have been discussed previously, but most have been rejected by the Agency without comment.

<u>Lack of Focus</u>. The majority of the CD focused on issues that were not relevant to the issue of PM, and the document failed to clearly articulate this shortcoming. To the casual reader or someone without knowledge of the PM in ecology/natural resources, one might arrive at the conclusion that PM is a serious concern in the community of ecologists/natural resource managers. If the CD had been focused on truly PM, the effort would have been abbreviated, less sensational and absent of an environmental, philosophical basis for the argument.

Many of these issues continue to sacrifice the staff paper and again the uninformed is left without a sense of what the science is and what the uncertainties are.

<u>Risk Model/Protocol</u>. The new generation of ecology has adopted a risk based approach to assessing how stressors affect ecosystems (natural and human dominated). This same protocol underpins all of the human health section in the Staff Paper, and it provides a structure and framework for the analysis. There is no parallel structure in the CD for ecology/natural resources and that same shortcoming is perpetuated in the Staff Paper.

It is time for the Agency to step into the new science of ecology. Any further delay perpetuates the idea that ecology remains a staid science, unable to adopt new methodologies.

What is more frustrating is the commitment from the Agency to rectify this problem and to date the Agency has superficially addressed the issue or has done so in a patronizing manner (see later concern)

References. This is a serious problem with the CD and now with the staff paper. It is recommended that the authors adopt the norms for the professional scientist. There are no other acceptable norms that should be considered. This issue has been raised before in the CD but is not recurrent in the Staff Paper.

There is a new concern that comes solely from the Staff paper. There is a common method to simply reference the CD on conclusions. These conclusions are appropriate but oftentimes the conclusions are from the open literature. The indirect citation approach fails to recognize the contribution of scientists whose idea was original and published in the open literature.. I would encourage recognition of key conclusions that are the unique contributions of scientists outside of the Agency. For those that are solely from the CD, it is fine to show the CD as the source of the conclusion.

In a parallel vein, there are some conclusions that are ascribed to the CD and then others are ascribed to the open literature. Is there a distinction to this dissimilar system of referencing?

Philosophy. As an outgrowth of the CD, this chapter perpetuates the idea that there is a concern for PM effects on ecology/natural resources. If you were to counsel the community of ecologists/natural resource managers, it is doubtful if many would even have any concern for PM. I do not follow that argument and the staff paper is ill equipped to address the issue of risk. Is it possible that the CD and Staff Paper are both developed from a philosophical position rather than a position of science? This might explain why the CD and Staff Paper seem to misrepresent the science. I would encourage the Agency to adopt an open science basis for its assessment; and if philosophy is the underpinning, that philosophy needs to be expunged.

Opportunity Missed. It has been argued several times that the current effort on PM was an opportunity missed. As the science of risk has marched forward in the discipline of ecology/natural resources, the Agency has remained stuck with its ad hoc approach for the CD and Staff Paper as it used three decades before. This round of the CD and Staff Paper were an opportunity missed to develop and apply and test the risk approach to air quality and ecology/natural resources.

Ingredients of a Risk Assessment. These are the essential ingredients of a risk based approach:

- Problem Formulation/Objective (clear articulation)
- Exposure Analysis
- Exposure-Response Analysis.

None of these aspects are addressed in the CD or the Staff Paper as it relates to ecology/natural resources. It is striking that the human health staff effort adopts this risk assessment approach and does a laudable job in developing the data, analysis and conclusions. The process and methodology is open to all.

Secondary Standard. The secondary standard is proposed to mirror that of the primary standard. There are no data to support that position. The Agency can not proceed with that recommendation without conducting a risk assessment. This conclusion is more based on the philosophy than the tenets of environmental science.

PM, Ecosystems and Vegetation. This is the basic focus of the Chapter. I am not quite sure that the exclusive focus on ecosystems and vegetation is appropriate as there are many other missing "components" in ecosystems. For example, where do microbes come into play? These are discussed but are not subsumed under this titling. Equally notable is the disparity in human health and the section on ecology. I would argue that if human health is having that much of an effect on human health that wildlife must also be affected, simply by first principles. A simple analogue approach would place exposure as being greater and the diversity of organisms would likely include some that are very sensitive. Ecosystems are composed of components other than producers.

Risk Assessment, Science and Patronization. There has been a repeated effort to encourage and cajole the Agency to adopt a risk assessment approach to the CD and Staff Paper as it relates to ecology/natural resources. The current documents are based on an undefined ad hoc approach that dates to the 1970's.

Promises have been made to adopt that approach, at a minimum in the Staff Paper. The commitment to that approach in the CD was summarily reject by the Agency. On page 5-36, the Agency offers a statement on a risk-based approach, and the statement is one paragraph. Thereafter, there is no effort to adopt the methodology of risk assessment. I regard this effort on behalf of the Agency to be one of patronization. I question the spirit and intent of this position and view the position as being a misrepresentation of the commitment from the Agency.

Environmental Sciences Versus Environmentalism. The aversion to a risk based approach with a quantitative framework that is formalized allows for the Staff Paper to evolve as an expression of the philosophy of environmentalism rather than one of environmental sciences. This is an unfortunate conclusion to offer but the phraseology and aversion to the risk assessment model allows for such a tangent to underpin the effort.

One of the principal reasons for the failure of scientists in the environmental sciences to obtain traction and respect in the larger community of natural sciences is the tendency for the environmental sciences to be "affected" by one's personal philosophical positions. This problem was one of the reasons that a formal risk assessment framework was initially develop, to allow all parties to follow the objective, methodology, data, argument and conclusion. This approach was summarily dismissed by the Agency and the consequence is that a philosophy drives many of the sections on ecology/natural resources.

Section 5.2. Visibility

Figure 5-3 and 5-4. Is the legend correct in this figure? Is the diagonal line the regression line (as indicated) or is it the 1:1 line? In these same figures, it is difficult to read the legends and units. If the diagonal line is the 1:1 line (and not the regression line), how does that affect your analysis (page 5-15)?

Visibility on the National Mall. (pp 5-28+). I am not convinced that this section should be included. Whereas the pictures have some value in some audiences, I am not convinced that the Staff Paper is the right place. Keep the focus on sciences as much as possible.

Section 5.4 Effects on Vegetation and Ecosystems

Perhaps this section should follow that of the human health and focus on the risk rather than solely effects. Is the title appropriately vegetation and ecosystems or are other components included as indicated in the first sentence?

Is there an objective and organization to this section or is it ad hoc in its approach? I would recommend that the outline used in human health be adopted here, that being a risk based approach.

It is important in any risk assessment that the following be articulated:

- Problem statement
- Exposure
- Exposure-response functions.

Are these to be addressed in this section or is the approach ad hoc?

Page 5-37, line 14. This is an untrue statement as PM also affects light attenuation and climate.

Page 5-37, line 28. This is a place for science. It is important to be clear that most of the discussion in the CD is irrelevant as it addresses deposition unrelated to PM. The role of PM in most of the constituents is a minority function and in many cases trivial.

Page 5-39, lines 8-9. This contrasts with others sections. Is the focus solely plants and ecosystems or are wildlife and microbes included?

Page 5-40, lines 9-10. The statement that exposure-response functions not being available places this chapter in the category of “why bother”. If a risk model were adopted early on, this shortcoming would have surfaced early and precluded a tremendous amount of work. Herein lies the value of adopting a risk approach.

Page 5-41, lines 9+. Why discuss acidic precipitation in a PM document? Again, the adoption of a risk based approach would have excluded this discussion. In the absence of the risk based approach

Page 5-41, line 17. Is it true that acid rain destroys plant cuticles? This again is more sensationalism. A risk based approach would be of value in reigning in this position, forcing one to be science based.

Page 5-42, line7. What constitutes an “impressive burden of particulates”? This is not a scientific statement.

Page 5-42, lines 27+. Again, this opening sentence and the next ones are unsupported by the science. I would encourage more of a science-based approach rather than a philosophical position. A risk based approach is advised.

Page 5-43, line 21. I would recommends a section on wildlife even if it is first principles.

Page 5-44, lines 23+. This is an example of inaccurate statements that tend to reflect a philosophical position in lieu of a science-based position. To link acid deposition and nitrate deposition to PM without any caveats is not scientifically sound.

Page 5-45, lines 3+. This section has some serious referencing problems and I encourage the Agency to adopt the reference model used by the remainder of the scientific community.

Figure 5-35. If this figure originates from the CD (as in an original contribution) the citation is appropriate. If the citation is otherwise, perhaps the Agency should give the author his/her due credit.

Page 5-48, lines11+. The paragraph begins with a statement that some forests in the US are "showing severe symptoms of nitrogen saturation…". I would argue that this statement is ill-advised in this document (which is a <u>PM</u> Staff paper). The uninformed will link this statement to PM, and perhaps that is the purpose. But, that is sensationalism, not science. If the CD had been focused on PM per se, this philosophy would not have emerged. Is this paragraph a unique statement herein, or is it a citation of someone else's conclusions?

Page 5-49, lines 8+. Unless I am misinformed, the cited study has no bearing on the discussion here and again suggests that the intent is more philosophically based. The study in Minnesota was not designed to mimic nitrogen deposition as it was a long-term <u>fertilization</u> study. The analogy to N deposition is poor at best. The implied analogy to PM is not appropriate science. Again, perhaps a good dose of science is in order.

Page 5-50. This entire page relates to how nitrogen inputs may change biodiversity and ecosystem structure and function. To the uninformed, the linkage to PM is not evident. The fact is there is a very loose or even tenuous linkage. This is more sensationalism than science. Perhaps a good dose of a risk assessment approach is in order. All of this section focuses on nitrogen but what is missing is the critique that clearly states that PM and nitrogen are not synonymous.

Page 5-52, lines 16+. This section is not relevant to a PM Staff paper and it needs to be either heavily caveated or deleted. The associated discussion of base cations is unwarranted as well. To the uninformed, this section links PM to all of these effects.

Page 5-56, lines 16+. The discussion of critical loads is ill advised. The basis has not be set for this analysis. It is interesting to note that there is more discussion of using critical loads than a discussion of risk assessment methodology in ecology.

Page 5-58, lines 19+. Here is a discussion of invertebrate in the section on vegetation. Not sure the logic of that is clear.

Page 5-59, lines 6+. Biodegradation is an active field called "phytoremediation". I would argue that this section does not reflect the literature.

Section 5.4.6. Rural PM Air Quality Network (page 5-62)

I would recommend that this section be expanded. If a risk based approach were adopted as the framework, this would have been a highly developed section and it would have been up front in the document. In the Staff Paper, a discussion about a monitoring site without any discussion about the data is nonsensical.

Lines 23-25. This reveals a significant lack of commitment to the risk based approach

Section 5.4.7. Summary

Since there is no objective, other than an ad hoc one, a conclusion is difficult to contemplate. What was the objective?

Page 5-63, lines 12-13. Where are beneficial effects discussed in the staff paper?

Page 5-63, lines 13-15. This is a conclusion that can only be reached through a risk assessment so it is inappropriate to raise that herein.

Page 5-63, lines 16-19. Where do PM-related effects clearly exist in ecology/natural resources? Is this is a philosophical conclusion/tenet rather than a scientific one.

This section is not a summary as there is no objective.

Section 5.5. Climate Change and Solar Radiation

Page 5-67, lines 12-16. Are you certain that you want to end a paragraph on a tenuous PAH statement?

Page 5-68, lines 5-8. This sentence is nonsensical.

Chapter 6. Staff Conclusions

It is recommended that the Agency make sure that conclusions flow from your objectives, that you have articulated your methods, and that you have data to follow through on your objectives. In the case of ecology/natural resources, there are:

- no objectives,

- no methods and
- no conclusion from the above.

Section 6.6 Secondary Standard Options

These notes related to discussion on pages 6-46 and the pages that follows.

Page 6-46, lines 14-17. I disagree with the conclusion. There are no data linking PM and effects. There are data relating air quality but certainly not PM. The Staff simply ignored addressing the issue of exposure and this precludes any assessment of risk.

Page 6-47, lines 17-20. This conclusion is not consistent with the data analysis (exposure and exposure-response), and I would argue that the phraseology is more a philosophical position than a scientific one.

Page 6-48, lines 10-12. Again, in the absence of a risk analysis, this statement is not supported.

Page 6-48, Lines 22+. This discussion addresses acidic deposition. The CD and Staff paper are for PM only.

Page 6-50., Lines 22+. In the absence of a risk assessment, there is no support for setting the secondary standard equal to the primary standard. There are no data to support that position.

Page 6-55, Lines 22+. I am not convinced this conclusion is warranted.

Particulate Matter Health Risk Assessment for Selected Urban Areas: Draft Report

This section is highly valued in the meaningful effort to be open in the procedure and methodology. On pages 3+ (general) and pages 7+ (specific), there is a clear statement on the ingredients for risk assessment for human health, with considerable specificity of the goals, end points and protocols. The discussion includes a rationale approach to assessing risk given several different scenarios.

The staff conducting the assessment of risk to ecology/natural resources might find some value in adopting a similar approach.

Dr. Sverre Vedal

November 2003
Critique of PM Staff Paper draft
Sverre Vedal

Chapter 3 (Health effects)

1. Consistency and coherence.

Both consistency of effects, and therefore coherence (which to some extent assumes consistency), are a matter of degree and hence subject to interpretation. The demonstration of consistency of positive effects across time series studies is in some sense the result of a process that involves selection of positive effect estimates in any given study from a large number of arguably equally viable possible estimates from which to choose. "Consistency", if defined as positive effects in multiple studies, is therefore nearly unavoidable, or a foregone conclusion. Further, the use of multiple city studies (p.3.89), in particular NMMAPS, to argue for consistency of effects, is ingenuous. The formal tests of heterogeneity in NMMAPS likely lack statistical power. Also, consistency was only deemed to be present in NMMAPS after application of Bayesian hierarchical modeling.

It is unclear to me why a stepwise increase in the size of effect estimates from the most adverse outcomes to the least adverse is support for coherence (p.3.98 and 3.100). It is clear that the population impact of a given size of effect estimate increases as the proportion of the population affected increases: that is, as the adversity of the outcome decreases. This does not require that the size of the effect estimate needs to increase. Is population impact being confused here with size of effect estimate?

One alternative approach to addressing coherence is to demonstrate similar effects on mortality and morbidity within a given city. This arguably preferable approach is not considered.

2. Confounding and other biases.

Although I believe that the arguments put forward against considering the gaseous pollutants as legitimate confounders is misguided (see my comments on chapter 8 of the CD), I will address one factual point of the argument. It is incorrectly stated that neither ozone nor SO2 can be considered to cause cardiac effects (p.3.73), whereas both have been shown to have cardiac effects in experimental studies (Tunnicliffe). The suspicion that air pollutants can cause cardiac effects is relatively new, so that there are very few data on cardiac effects of pollutants other than PM.

The Staff Paper repeats the argument in the CD in support of the notion that the gaseous pollutants are merely surrogate measures of ambient PM, and, interestingly, that CO and NO2 are markers of vehicle-generated PM, and that SO2 and ozone are markers of sulfate (p.3.74). The ozone-sulfate correlations are often weak, so this seems an unlikely role for ozone. I believe there is now that there is beginning to be some consensus that ambient concentrations of pollutants are in fact surrogate measures. The main disagreement is whether PM itself is immune from such considerations: that is, whether gaseous pollutants are surrogate measures of

PM, versus whether all of the pollutants, including PM, are surrogate measures of aspects of the atmospheric pollutant-meteorology mix.

The figure on p.3.96 (no figure number is supplied for this figure) and the corresponding discussion in the text (p.3.94) attempt to address the plausibility of confounding by the gaseous pollutants by plotting effect estimate size (RR) against gaseous pollutant concentration for several studies. The fact that RR does not increase with increases in gaseous pollutant concentrations is taken as evidence that confounding by gaseous pollutants is unlikely. This does not follow. Joel Schwartz introduced the approach of plotting effect size against the temporal correlation between PM and the gaseous pollutants, which at least makes some sense. The approach presented here, however, is entirely unconvincing.

The reporting of "best lag" (p.3.78) is again defended as in the CD, whereas I find this practice difficult to defend. The practice of reporting different lags for different cities (since best lags differ from city to city) is defended (p.3.79).

3. <u>Multicity studies</u>.

I agree that the multi-city studies should be given the most weight. However, not all multi-city studies should be given equal weight. Not only are multi-city studies characterized by more precise estimates of effect, but some also use an unselected sample of cities and theoretically avoid publication bias. Only the NMMAPS and the Canadian studies, of the studies listed in Table 3-2 (p.3.17), are unselected. The NMMAPS estimates of effect are the lowest, and the Canadian effects are sensitive to model specification.

4. <u>Statistical modeling</u>.

It is not clear to me why GAM is preferred over GLM at this time for more valid effect estimation (p.3.24, line 10).

5. <u>Chronic exposure studies</u>.

It is unclear to me what is intended by the phrase, "… do not negate the findings of the 6 Cities and ACS studies" when referring to the findings of the AHSMOG and Veteran's cohort studies (p.3.39). It seems to me that one needs to decide either that the findings of these other studies should be ignored (providing a sound rationale), or that the findings indicate that there is some uncertainty and inconsistency in the cohort study findings.

<u>Relatively minor comments</u>.

1. p.3.22, line 7. While this is true in APHEA when using a stricter GAM, the GLM estimates were substantially lower than either GAM estimate.
2. p.3.23, line 8. This should be "…<u>statistically</u> significant…"
3. p.3.64, lines 11-64. Most human experimental and toxicologic studies have not found any changes in peripheral white cell count, hemoglobin concentration, or platelet count. To point out that some did is misleading.
4. p.3.7 (Table 3.1) Autonomic effects are not strictly "direct effects on the heart", but are pulmonary reflexes, as is indicated on p. 3.9 (line 2).
5. p.3.82 (line 14). This should be "<u>3.5</u>.3.1".
6. p.3.94, line 27). As I indicated in my comments on chapter 8 of the CD, of all potential outcomes in the Utah Valley, the steel mill closure design only used respiratory hospitalizations; school absences were examined only in a time series design. The correct references should be to

Pope, 1989 (hospitalizations) and Pope, 1992 (school absenteeism) and should be in the reference list.

Chapter 4 (Risk assessment)

1. Study areas.

As noted in my comments on chapter 8 of the CD, although the cities in NMMAPS with more power have more precise estimates of effect, these effects are not necessarily more "homogeneously positive" (p.4.15, line 3).

It is not clear to me why Provo, UT findings are presented (p.5.21,23,25), given that Provo is unlikely to meet any of the statistical power criteria.

2. Concentration-response models.

The contention that corrected GAM estimates provide more valid effect estimates than do GLM estimates, and are therefore preferred, is not justified (p.4.31, line 19).

Regarding thresholds, it is not clear that any single study has examined whether thresholds are present "in a statistically significant manner" (p.4.32, line 8). These have largely been descriptive presentations, as have the observations of linearity.

Again, an argument against confounding by gaseous pollutants based on lack of association between effect estimates and gaseous pollutant concentrations (p.4.34, line 7), as pointed out in my comments on chapter 3, is not sound.

The sensitivity analyses outlined in Table 4-9 (p.4.41) are a beginning, but are in no way comprehensive.

3. Results.

I do not understand why there is so much variability in the width of the 95% confidence intervals in Figure 4-7 (p.4.47).

What is the justification, and merit, to doubling effect estimates to simulate distributed lag estimates (p.4.62, line 3)?

Minor points.
1. What are the units for mortality (y-axis) for figures 4-15b (p.4.68) and 16b (p.4.69)?
2. Legend to Figure 4-5 (p.4.45), item 15. Should this be NO2?

Chapter 6 (Conclusions/recommendations)

1. Arbitrariness.

There is a sense of arbitrariness in the process of narrowing attention to a given range of annual and 24-hour concentrations.

2. Uncertainties/limitations.

Although uncertainties and limitations in the PM health effects findings are acknowledged, it is not clear how these are incorporated into the recommendations. Formal incorporation of uncertainty is limited to that reflecting sampling variability in estimates of effect (indicated by their 95% confidence intervals). There are many other sources of uncertainty in the time series studies, including: 1) adequacy of control for temporal effects and meteorology, 2) selection of "best" lags, 3) model selection, and 4) selection of studies. There are formal methods for incorporating other sources of uncertainty.

3. "Controlling" and "backup" standards

I agree with the approach of using an annual concentration as the "controlling" standard, and the 24-hour average concentration as a "backup" standard. This is partly motivated by the inability to identify any justifiable 24-hour average threshold concentration below which effects are not detectable (although one could theoretically take a stance [probably difficult to defend] that effects below a given concentration are so trivial that ignoring them still allows one to protect public health with "an adequate margin of safety").

Once having determined that the annual standard will have primacy, support for a certain range of annual concentrations rests almost entirely on the short-term (time series) concentration studies (pp.6.14-15). I find this approach to be forced. The mean 24-hour concentration in a given study (or city) is used to determine whether effects are present at a given annual concentration. Yet, effects in these studies could be determined by concentrations above or below this concentration, or both. The mean 24-hour concentration provides little, or no, insight into this. It therefore seems difficult, if not impossible, to use these data in focusing on an annual concentration range. Without determining the concentrations below which effects are not detectable in these studies, I cannot see that they provide usable information. I therefore favor using the long-term exposure (cohort) studies for the purpose of identifying a range of annual average concentrations.

4. The "backup" standards

The two approaches used for focusing attention on the range of 24-hour average PM concentrations (pp.6.22-24 and 6.34-35) essentially rely on measured distributions of PM concentrations and estimates of the percentage of US counties that would be out of compliance at selected concentrations. This determination should instead be based on health effects at given concentrations.

5. The coarse PM standard

It is my opinion that proposing a coarse PM standard is premature at this time. Observational findings are based on time series studies about which, to my mind, there is sufficient uncertainty to preclude setting a standard. The PM2.5 observational findings, which include the time series studies and their attendant uncertainties, also include cohort study data, which while not entirely consistent and not immune to concerns regarding confounding, are now a main pillar. The extensive toxicological and human experimental data are also enhancing the plausibility of the observational findings on PM2.5. In contrast, the coarse PM data do not include positive findings from cohort studies, and the toxicologic data are slim. Further, there is good evidence that PM of crustal origin, as a subset of particles included in the coarse fraction, are not particularly toxic, as opposed to some other components in the coarse fraction. In many settings, the coarse fraction is dominated by crustal PM.

Minor and/or editorial comments

1. The evidence for specific toxicity of PM due to sulfate or acid aerosol (p.6.10, line7) is meager.
2. The concept of "weight of evidence" (p.6.17, line1) is not readily applicable to a single study; rather, it finds its utility in looking at the combined evidence from multiple studies. In fact, the "weight of evidence" from a single time series study is relatively low, given the multiple potential comparisons that can be made (multiple lags, outcomes, models, etc.).
3. I would make clear that this (p.6.17, line16) is referring to short-term concentration studies.

Dr. Barbara Zielinska

November 2003
Comments on the 1st Draft of PM Staff Paper
Barbara Zielinska

Chapter 2

In general, this section is well written and represents comprehensive summary of information contained in Chapter 2, 3, and 5 of the 4th Draft CD. I have a few minor comments listed below:

1. It is true that the scientific information concerning coarse (PM10-2.5) particles is rather limited. However, some specific properties of these particles that are important for establishing a standard should be emphasized. This include a shorter atmospheric lifetime, significant differences in chemical compositions depending on a geographical location, and most importantly a limited penetration into indoor environments that explain low correlation between personal exposure and outdoor concentrations (as measured by central monitors)
2. Page 2-25, line 10. It is surprising that 98th percentile 24-hr average PM2.5 concentrations above 65 ug/m3 appear in Montana. What is the reason for these high concentrations?
3. Figure 2-14, page 2-39 is very difficult to read.

Chapter 3

1. Section 3.5.3.2 (PM Components and Source Related Particles) has several inaccurate statements. Page 3-84, line 17-22 lists elemental and organic carbon (OC/EC) as indicators of motor vehicle emissions and similarly page 3-85, line 21-25 lists COH, fine PM, NO2 and CO in addition to OC/EC, as mobile sources related pollutants. These statements are not accurate, since OC/EC, fine PM, NO2 and COH are related to many combustion sources, not necessarily motor vehicle emissions.
2. The same Section 3.5.3.2, p. 3-87 discusses bioaerosols, including endotoxin. The treatment of this subject is rather weak. There are evidences of bioaerosol present in PM2.5 fraction, not only in the coarse fraction. As the recent CASAC review pointed out, the discussion concerning bioaerosols in the 4th draft of CD needs to be improved as well.
3. Chapter 3 discusses in several places the so-called "intervention" experiment in the Utah Valley (p. 3-56-57, 3-94). To me, this experiment doesn't indicate the toxic effect of ambient particles, it only indicate the toxic effect of emissions from the very specific source, i.e. the steel mill. The fact that the oxidant activity, inflammatory responses, etc., of the ambient PM were greatly reduced after the steel mill closure, indicates that not all ambient PM is created equal and that the PM health effect depends greatly on its sources and chemical composition. In

fact, the transition metal content of the Utah Valley PM during the still mill operation was more closely linked to health effects than mass of the particles. This "intervention" experiment would rather support the concept of the source control, but not necessarily a general NAAQ PM standard.

4. Section 3.5.1, page 3-74, line 1-8. The statement that gaseous pollutants can serve as surrogates for ambient PM exposure, and that CO and NO2 are markers for vehicle-generated PM, and SO2 and ozone are markers of sulfate is not correct. What ozone has to do with sulfates? For example sulfate concentrations are low in the South Coast Air Basin, CA, but ozone is often high. Also, the correlations between CO, NO2 and PM emissions from motor vehicles are not straightforward.

Chapter 6.

I have some general concerns regarding this chapter:

1. The proposed more stringent standard levels (especially 24-hr PM2.5) seem to be rather arbitrarily established. Although Staff Paper lists many uncertainties and limitations of PM- health relationship, there is no explanation how these uncertainties are incorporated into the new proposed levels.
2. Coarse particle standards do not seem to be adequately justified. The draft Staff Paper acknowledges that the crustal material, often important fraction of coarse particles, is not toxic. Coarse particles that originate from traffic, i.e. road dusts with tire and break debris, deposited motor vehicle exhaust, etc. may show health effects, but geological material is rather insignificant. Furthermore, coarse particles have shorter atmospheric lifetime, are not uniformly distributed, and their penetration into indoor environments is low. Thus, the correlation between ambient concentrations as measured by central monitors and personal exposures is rather limited. I'm concerned that the method for establishing the coarse particle standard, proposed during the Nov. 12-13 meeting, which takes into account an average ratio of PM2.5/PM10, is not justified. This ratio would be very different for different settings (i.e. rural versus urban, midwest versus northeast, etc.), as pointed out in Section 2.5.6 of the draft Staff Paper.

Dr. Jane Q. Koenig

September 28, 2003

Comments on the OAQPS Staff Paper for the PM criteria document
Jane Q Koenig

In response to the issues raised by Les Grant, here are my responses.

In my opinion, both ch 2 and 3 contain adequate air quality information to be judged complete.

Regarding ch 4, I get a feeling of déjà vu. I thought we had approved a method of approach and were awaiting results.

Regarding Ch 6.

I agree with the authors that selecting a range of primary standards is largely a public health policy judgment. I also agree with their decision to continue to use undifferentiated particle mass as the basis for the indicator for fine PM standards.
Probably things would be cleaner if PM1.0 had been chosen as the indicator as PM1.0 more certainly has different sources than PM10.

Regarding the averaging time I would like to see a suggestion of what concentration would be considered for a one-hour standard and some recognition of the problems communities have that are impacted by episodic smoke (agricultural burning and forest fires).

I do not believe the data exist to allow a conclusion that peak 24 hr PM2.5 concentrations contribute a relatively small amount to the total health risk. (page 6-14).

I disagree with the Staff recommendation that the 98^{th} percentile be retained (page 6-23). As I understand it, this allows a community 7 days above the standard before action is taken. I think seven PM episodes are too many to protect the public health.

Regarding the CF standard;

This is a much more difficult decision due to the paucity of data. I do judge that 75 ug/m3 is too high. Seventy five would be more than double the concentration at which effects are seen.

Recommendations;

In my judgment the recommended range for the 24 hr PM2.5 standard should not be as high as 50 ug/m3. In 1997, when Carol Browner sent a recommended 2.5 standard to the White House, she selected 50 ug/m3. I believe we now have evidence that this concentration would not be protective of public health. Including 50 ug/m3 in the recommended range would allow selection of a standard that actually does not reflect the wealth of new data published since 1997.

I also do not support a range for the annual standard that includes 15 ug/m3 for much the same reasons as stated above. If we believe, as I do, that research since 1997 has shown a wider range of health effects associated with fine particles and at lower concentrations, then we should support a stricter standard not the current one. If in fact, individuals are still dying and being made ill from PM exposure in the US, public health demands a stricter standard for protection of sensitive populations.

I commend the authors of the staff paper on a well written document.

Jane Koenig

Dr. Petros Koutrakis

Staff paper review by Petros Koutrakis,
Harvard University, School of Public Health,
Boston, MA
November 10, 2003

PARTICLE PROPERTIES

Section 2.3, page 2-12; lines 11-12: I think the $\pm$10 uncertainty is too optimistic.

Same page, line 26; Considering that atmospheric lifetimes of coarse particles are much shorter than those of fine particles, comparing fine and coarse emissions may be misleading. For instance, for the same emission rates the resulting ambient concentrations will be higher for fine particles.

Page 2-40, Figure 2-15; I wonder what happened after 1995. Did they change the sampling or analysis method? Something must have happened around this period.

Page 2-41, line 21; Of course, PM_{10} will be a suitable indicator for fine or coarse particles since it encompasses both fractions!

Page 2-42, line1; I suggest using "diurnal" instead "temporal".

Page 2-51, line 13: do you want to say "Is North America supposed to include Mexico and Canada?"

Section 2.8 discusses PM exposure assessment issues. This is a well-written section and addresses some key findings. The section focuses only on the relationship between personal exposures and outdoor concentrations. However, the scope of this section could be expanded to address additional important PM exposure assessment issues such as: exposure to specific sources; differences between acute and chronic exposures; exposure characterization at greater time resolution, and; the implications of varying indoor/outdoor PM ratios for epidemiological studies. In addition, this section relies only upon older studies (PTEAM) and does not report findings from the recent exposure assessment studies. Finally, there should have been more emphasis on the relatively low penetration of coarse particles (inversely related to particle size) into indoor environments. Results from the very limited existing exposure studies of coarse particles suggest no relationship between personal and outdoor coarse particle concentrations. The implications of these findings for the proposed coarse particle standard needs to be evaluated.

Page 2-60, lines 1-4; This sentence needs editing because it is not very clear.

Section 2.9; I fear that this section will not have any impact on setting up a new standard. This information is of academic importance, but is not particularly suitable for influencing decision-making. I know we need to write something about everything, but I am not sure whether this is correct. It would be more appropriate to just report that particles may impact visibility and the radiative balance of the atmosphere and that we have no idea about the quantitative relationships between concentrations and these effects. Tutoring the administrator on Physics 101 is not necessary.

HEALTH EFFECTS

The Health effects section is very well-written. The integration of epidemiology and toxicology is commendable. The implications of the recent exposure assessment findings to epidemiology are commendable as well. I do not remember seeing anything about threshold and the shape of the dose-response relationship. Maybe these issues are discussed in the risk section. Finally, the information presented for coarse particle health effects is very sparse. Although I believe that it is a good idea to replace the PM10 standard by a standard for coarse particles, I am not sure that there is strong scientific evidence for this decision.

Page 3-12, lines 9-11; I am not sure that I completely agree with this statement. Animal studies have been quite valuable in our efforts to investigate particle health effects. Many epidemiological findings have been reproduced by animal studies. These studies have used animal models of cardiopulmonary and vascular decease exposed at relatively low doses and observed outcomes similar to those previously reported by epidemiological studies. Subsequently, these findings were used to design the new generation of epidemiological studies. This synergy between human and animal studies is one of the major advancements made since 1997.

Section 3.4.1; Individuals exposed to high PM concentrations are also at high risk. Often individuals with low income live near busy streets and industrial facilities (this is a feature of environmental inequity). Also genetic factors can induce susceptibility. These two factors should be included in this section.

Page 3-76, lines 12-14; This is a quote from the HEI report which is not clear to me.

Section 3.5.2.3; One of the most important findings of the recent exposure assessment studies is the varying impact of outdoor sources on indoor environments and thus exposure. Homes with high air exchange rates are less protected from outdoor sources. The opposite is true for homes with low rates. Homes located at areas with harsh winters or very hot summers exhibit low air exchange rates, such as Boston in the winter and Atlanta in the summer. In contrast, homes in California are well-ventilated and present high indoor/outdoor PM ratios. Therefore, one would expect that for the same outdoor PM levels individuals living in areas with moderate weather receive higher relative exposure. This may explain the results of the APHEA study which found higher risk factors for South Europe as compared to Northern European countries participating in the study. This varying impact of outdoor sources was only briefly discussed in the exposure section. Also, the results by Janssen et al showed that use of air conditioning (a surrogate for low home ventilation) explained some of the heterogeneity in risk factors among cities in the NMAP study.

Section 3.5.3.1; Only a small fraction of outdoor ultrafine particles are found indoors. Also, since concentrations of ultrafine particles present considerable spatial variability, one would expect to find only a weak or negligible relationship between personal and outdoor exposures. Sometimes we need to be reminded that exposure is necessary in order to produce an effect.

An alternate way for humans to be affected by ultrafine particles is the following: First, ultrafine particles coagulate (stick) onto fine particles, so the fine particles can act as vectors for the ultrafines. Subsequently, once fine particles are deposited inside the pulmonary system, these ultrafines can be released by de-coagulation, perhaps by way of interactions with lung surfactants.

Page 3-84, lines 23-30; There is too much emphasis on aerosol acidity, that in my opinion is not justified. If there is any clear evidence for toxicity of specific types of particles, it is for particles associated with traffic, but this not at all stressed in this chapter. Results from the California children's study and the Harvard Six City study certainly suggest that traffic particles are quite toxic. Some European studies support this as well. Furthermore, the Harvard animal CAPs studies have found strong associations between several cardiac outcomes and fine road-dust in more than one experiment (Batalha et al, 2003).

Page 3-90, lines 4-6; This is a very strong statement. The evidence for coarse particle health effects is not sufficient to derive such a conclusion. The Staff Paper should also report studies that have not found health effects associated with coarse particles. It is interesting that in page 3-99, lines 20-23, the Staff Paper reports "…results are not as consistent as those for fine particles". In the 1997 CD, the Philadelphia study and the Six Cities study were used to show that there are no coarse particle effects, in order to strengthen the case for setting up the fine particle standard. I personally, think that not all coarse particles are toxic and that only road-dust and coarse particles from industrial activities can be toxic. Road-dust encompasses many toxic components, deposited vehicular exhaust emissions, brake materials, tire debris (which includes latex and many metals), biological materials (such as pollen, endotoxin and spores, among others), and nutrients for microorganisms such as sulfates and nitrates. A fraction of road-dust can be found in the fine size range (below 2.5 μm), but its majority is present in the coarse mode.

Regulating alumino-silicates, calcium carbonates and other benign crustal materials may not be the best and most cost-effective approach.

Page 3-94, line 29; The Dublin study should also be included among these studies. Also the California children's study and the recent Harvard Six Cities study re-analysis have reported some very intriguing findings on this issue.

RISK ASSESSMENT

Page 4-16, paragraph starting on line 9-; When assessing risks, it is important to keep in mind that the penetration of both coarse and ultrafine particles from outdoors to indoors is less efficient than for fine particles.

Table 4-7 on page 4-25; It is important to take into account the different relationships between personal exposures and outdoor concentrations, when comparing cities with different climatic conditions.

Table 4-8 on page 4-26; It is surprising that coarse particle background concentrations in the West are similar to those in the East. One would expect these concentrations to be higher in the drier West.

Page 4-32, line 25; Typo.

Figure 4-3, page 4-43; The short term fine particle related mortality "as is" is higher in Boston than Los Angeles. As we know LA is more polluted. Am I missing something here? In Figure 4-7 one can see that for long-term mortality LA is higher than Boston. Do we believe that acute effects produce sub-acute effects, which ultimately become chronic? If the answer is yes, how one can explain these results?

Figure 4-10, page 4-52. How one could explain that the percent confidence intervals for PM10 are smaller than for PM25? Similarly for figure 4-13.

Tabl3 4-13, page 4-72; What is the connection between the short and long-term particle exposure related mortality? For instance, are the 550 deaths in LA included in the group of 2730 deaths?

References:

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, (2003).

Dr. Allan Legge

December 10,2003 (revised January 24,2004)
TO: Dr. Phil Hopke/Mr. Fred Butterfield
FROM: Dr. Allan H. Legge

Review Comments: First Draft OAQPS PM Staff Paper entitled
"Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information"

Overall Comments: The following comments are directed at the 'PM-related effects' on vegetation and ecosystems (Chapter 5) and to a lesser extent 'Staff Conclusions and Recommendation' (Chapter 6) relating to vegetation and ecosystems as presented in the 'Staff Paper'.

Staff is to be commended for their initial efforts in attempting to address the matter of PM - related effects on vegetation and ecosystems. This is a difficult if not impossible task for PM, however, given the current required approach to setting, maintaining and/or revising an air quality standard. The overriding assumption that one can attribute, for the most part, the response or responses of a receptor to a given air quality stressor within a given short time frame simply does not work in the case of PM. This is very unfortunate from the standpoint of environmental protection especially in light of the fact that there are some forested ecosystems in the US which are showing clear evidence of 'nitrogen saturation' a portion of which is due to particulate nitrate deposition. The problem here is that this 'nitrogen saturation' has been brought about by chronic long-term exposure to elevated nitrogen deposition. It is the cumulative load of nitrogen over time which has resulted in some forested ecosystems being nitrogen saturated. Some would say that the fact that we do not know the exact contribution of 'particulate nitrate' deposition to the nitrogen saturation evidenced in some forest ecosystems prevents us from doing anything. This is not true. What is needed is a philosophical change in the way one approaches environmental protection. The European Concept of 'critical loads' was suggested as one possible scientific approach when reviewing the PMCD. This approach would more readily lend itself to risk assessment than the current information.

Specific Comments:
1. Page 5-50, line 10.
Spelling, should read, "-----; for example, in studies of the"
2. Page 5-52 line 25.
Suggest this read "----, a period that coincides with the increased emissions of"
3. Page 5-57, lines 20-25.
This paragraph needs to be rethought. First it is indicated that the critical loads concept has significant potential for the long-term protection of ecosystems but then goes onto say that the approach is too data intensive to be practical in the US to protect sensitive US ecosystems from adverse effects related to PM deposition. This does not

make sense and reflects a lack of understanding of the critical loads concept. PM would no longer be the focus but rather total deposition along with cumulative deposition of the parameters of concern.

Dr. Paul J. Lioy

Comments of Dr. Paul J. Lioy on: OAQPS Staff Paper for PM --- Submitted by E-mail 11-11-03

Overview:

The Staff has provided an important set of analyses from which to assess the risks associated with PM and its various size fractions, and should be commended for their efforts on a difficult and ever evolving environmental health issue. Further, the presentation was very clear and easy to follow. As stated in the text of the document, however, the results must be viewed with some caution as the CASAC has not yet closed on the Criteria Document for PM.

Major General Concerns:

The heart of the matter for the Staff Paper is Chapters 6 – the Staff Conclusions and Recommendations on PM NAAQS, which is supported by the exposure and effects characterizations and risk assessments in prior chapters.

The case for the long term PM standard – Annual Standard is compelling, and has been solidified by research (primarily epidemiological) that has been conducted and reported since 1996. The range identified by Staff for the Annual $PM_{2.5}$ standard will be debated, but I see no reason that precludes forward with an Annual NAAQS for $PM_{2.5}$.

Based upon current knowledge of exposure – response relationships, the case for a lower short term (24h) standard, beyond the 65 ug/m^3 standard that was promulgated in 1997 by EPA, has not been adequately made in the current draft of the Staff paper.. There are few studies to date that have focused clearly on this important issue. The suggested range is based upon the notion that if you attack and reduce the peak or near peak levels of $PM_{2.5}$ mass you will then reduce the Annual Mean. This approach would have a high degree of credibility if all PM that accumulated in the atmosphere was from primary emissions, and had the same or a consistent suite of sources. However, $PM_{2.5}$ levels are significantly affected by photochemical smog processes that produce secondary fine particles, and these particle are transported long distances. Thus, periodic smog events can contribute to levels above a 24 standard; but, would regional strategies be the most effective way to bring down the mean? -A scientific question that still requires an answer.

Reducing the annual $PM_{2.5}$ emissions from both stationary and mobile sources of primary PM particles would be the most effective approach for reducing the annual mean. Please note, on page 4.73 of the Draft Staff Paper there is a caution about using a rollback of the peaks as a method for achieving the annual mean. This caution would also be supported by the fact that a local or regional increase in $PM_{2.5}$ can be caused by unusual sources and or unusual events. For example, in 2002 the States of NY and NJ were affected by forest fires in Canada. These led to significant increases in $PM_{2.5}$ that resulted in violations of the 65 ug/m^3 standard at multiple sites for two or three consecutive days. I am sure that similar experiences occurred in California, and Texas and Florida over the past months and years, respectively. The question is: do these isolated events have any bearing on the reasons for achieving an Annual $PM_{2.5}$ standard? The evidence that currently exists in the draft Criteria Document and the Staff Paper for PM, do not support a "yes" answer at this time. Surely, improvements in forest management will help reduce the severity of these costly and deadly fires, but they are not the root cause of the long term exposure-response relationships that have been identified in many epidemiological studies completed over the past decade and a half. I recommend that the EPA revisit this issue in the next draft of the Staff Paper.

The suggested coarse particle standards for $PM_{10-2.5}$ are even more troubling than the short term fine particle standard. Again, l did not find any well established exposure – response relationship for this size fraction. An even more fundamental issue is based on the fact that the coarse particles discussed in the Staff Paper are limited to the material that exists only within the narrow size interval of 2.5 to 10 um in diameter. This definition of coarse particles totally ignores atmospheric contributions of coarse particles above 10 um in diameter. Without a substantive discussion and evaluation of the definition of coarse particles and the potential for yielding short term or long term effects the proposed standards are arbitrary. This point was raised at previous CASAC meetings and on conference calls, and deserves some action. Maybe the issue requires specific acknowledgment of this problem in the Staff Paper and a serious recommendation for a National Conference on the Coarse Particle NAAQS issue. I point the Staff and others to the tragic events of 9-11-01. At that time most of the dust and smoke released during the first week were above 10 um in diameter. There were no measuring devices available to quantify the levels of the mass above 10 um in diameter during the first week or subsequent weeks, and no reasoned standards to refer to for assessing the potential short term risks. A deficiency in our monitoring capability and standards. Clearly, the issue of coarse particles is work in progress, but I am concerned about arbitrarily defining coarse particles as $PM_{10-2.5}$ before there is adequate data and information to support limiting concerns about coarse particle health and welfare effects to this narrow range of particles.

Major Specific issues:

1. Pg 2.53, Pg. 4.26. I am puzzled about the range of background levels used for the risk assessment, especially for the coarse particle fraction. I find it somewhat difficult to understand how the background for coarse particles can be the same for east and west locations. However, this may be correct because of the definition of coarse particles used by the Staff is $PM_{10-2.5}$. In this narrow size range, the average background levels may be quite similar. I would expect much greater differences in average “background mass contributions” for coarse particles above 10 um in diameter.
2. Pg. 4.36 to 4.38. Discussion about uncertainties and sensitivity is qualitative. The presentation of quantitative values for the level of uncertainty would be useful for each size fraction considered in the risk assessment. This could help prioritize the variables of concern in the risk assessment.

Dr. Mort Lippmann

REVIEW COMMENTS
OAQPS PM STAFF PAPER – AUGUST 2003 DRAFT
by M. Lippmann

General Comments

This first draft of the PM Staff Paper has provided the CASAC PM Panel with a description of the OAQPS interpretations of the scientific, peer-reviewed literature in the fourth draft (June 2003) of the PM CD. It presents a straightforward description of its selection of the studies it finds most relevant too the setting of the next PM NAAQS, and how it has interpreted them. It also presents the results of the risk assessment performed by Abt Associates, and its preliminary recommendations on PM NAAQS. It also acknowledges that it is prepared to make revisions based on its reviews of the CASAC and public commentaries to the fourth draft of the PM CD, which were not available to OAQPS prior to the completion of this first PM Staff Paper draft.

I found this Staff Paper draft to provide a fair and balanced presentation of the relevant literature. It thoroughly and appropriately addressed the use of this literature in terms of defining its options on the index pollutants, the most appropriate averaging times, the statistical form(s) for the PM NAAQS, and the concentration ranges appropriate to the protection of the public health and welfare with an “adequate margin of safety”.

While adjustments will need to be made to reflect the further CASAC and public comments on the final chapters of the PM CD and this draft of the Staff Paper, I would not recommend any major changes in format or approach to this document, and commend the OAQPS staff for the work they have done in preparing this draft.

Specific Comments

Page, Line	Comments
2-3 (Table 2-1)	Under "Coarse Particles", line 3: change "usually" to "may" and add text to end of line as follows: "when resuspended dust is a major component of ambient air PM".
2-3 (Table 2-1), 2-7, line 2	Under "$PM_{10\text{-}2.5}$", line 3: change "inhalable" to "thoracic". By convention (ACGIH, ISO, CEN), inhalable refers to PM aspirated into the nose or mouth.
2-9, lines 1 & 2	change "droplets which react" to "vapor that reacts"
2-9, line 24	add "and humidity" after "particles"
2-10, line 12	change "as" to "in"
2-15, line 14 & elsewhere	change "COH" to "CoH"
2-16, lines 19, 23	change "impacting" to "collecting". Impaction is not the only mechanism for particle collection in filters.
2-18, line 5	insert "PM" before "sampler"
2-18, line 23	insert "and conversion of PM components to gas-phase chemicals" after "filter"
2-19, line 1	change "filters" to "impaction plates"
3-4, line 13	insert "portions of the" before "ultrafines"
3-59, lines 4-17	The Children's Health Study (CHS) findings are not properly discussed. The Peters et al (1999a) results were from a cross-sectional analysis with limited statistical power. The Gauderman et al. (2000 and 2002) papers described the results for two separate cohorts of 4^{th} graders followed over four years, and did find consistently statistically significant reductions in the growth of both MMFF and PEFR, albeit not generally for FEV_1 and FVC. The consistency of the results of Gauderman et al. (2000, 2002) and of Avol et al. (2001) on cohort children who moved is compelling evidence of PM-related decrements in lung development.
3-68, line 21	The words "older children: could easily be misconstrued. Change to "children studied from fourth grade to eighth grade".

3-85, lines 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.
3-85, line 19	insert “mass” before “indicators”
4-5, lines 27,28	delete “with some studies suggesting associations between $PM_{10-2.5}$ and mortality as well” and move “(CD, p. 8-57) to line 26 after “mortality”
5-9, line 12	change “PM” to “$PM_{2.5}$”
5-42, lines 4,5	What about Pb and As?
5-43, line 10	delete “While these substances are not criteria pollutants”. They are clearly part of PM, which is a criteria pollutant.
6-4, line 28	insert “between adverse health effects and” to replace “with”
6-20, line 6	delete “in 1994”
6-20, line 7	insert “In 1994, $PM_{2.5}$” before “ranged”
6-20, line 9	insert “over a four-year period” after “concentration”
6-28, line 27	insert “in children” after “growth”
6-31, lines 22, 23	This statement, while true, is misleading. It should be qualified by noting that the gaseous co-pollutants in the Six-Cities study were not significantly correlated with mortality.
6-32, lines 22-25	Once again (see my comments above on page 3-59, lines 4-17) the findings of the CHS study are not adequately discussed.

Dr. Joe Mauderly

Comments on OAQPS Draft Staff Paper on Particulate Matter

Joe L. Mauderly

General Comments:

Overall (and excepting the relatively minor factual and editorial points raised below), I find this first draft to contain a reasonable distillation of the current health data from the CD, and a reasonable range of recommendations regarding the PM standards. The key points of uncertainty regarding the current health data (epidemiology, clinical studies, and toxicology) seem to be appropriately described. A major issue for discussion (and differences of opinion) will be how strongly these uncertainties should restrict the proposal of more stringent standards. Given current information, however, the current <u>range</u> of recommend actions seams reasonable. Pending issues raised by others having different technical expertise, I'd say that for a first draft, the Staff Paper is well on the way.

Specific Comments:

<u>Chapter 3</u>

3-4, L 12: The sentence is not incorrect, but actually, fine and coarse fraction PM deposit in all three regions, not just in the tracheobronchial and alveolar regions, as the sentence might be taken to suggest.

P 3-6, L 11: It should read "—health outcomes—", not "health endpoints".

P 3-9, L 2: One could argue that PM-induced reflexes constitute an indirect, rather than a direct, mechanism.

P 3-9, L 15: I don't think that it's conceptually correct to state that "particles also may carry other substances with them" (i.e., non-particulate substances). Anything carried by a particle <u>is</u> the particle. Particles are complex with many "core" and "adsorbed" materials, but if it's on a particle, then the whole thing is a particle.

P 3-10, L 9: I believe that reduced lung growth rate should also be on the list.

P 3-13, Footnote No. 3: I believe it also excludes homicides.

P 3-44, L 2-4: This statement isn't clear. I suppose that you might mean that indications of mechanisms at work may support causality for development of a health outcome, although not a direct measure of the outcome itself, but that's just a guess.

P 3-52, L 15: The name "Dominici" is misspelled here as "Domenici". You need to do a universal search on the name in the text and references and make sure it's correct. Dominici is the statistician/epidemiologist at Hopkins. Domenici is the Senator from New Mexico.

P 3-57, L 1-4: First, it is not clear that there has been "controlled exposure of humans to diesel exhaust particles". There have been controlled exposures of humans to diesel exhaust, resulting in inflammation. Therefore, the fundamental point being made is valid. However, do not confuse exposures to exhaust with exposures to "particles". There have been some nasal instillations of diesel particles, but that's not what is suggested by the sentence. Second, CD p. 7-20 is cited as the reference for this statement. There is nothing at all on that page of the CD that refers to diesel particles.

P 3-57, L 22: Presumably, the "industrial PM source" referred to here is the Utah steel mill. That is referred to as a "steel mill" on the previous page. Being consistent would help avoid confusion.

P 3-65, L 9: The implication of the wording is that sometimes it is <u>not</u> difficult to separate effects of different pollutants. As far as I know, it is <u>always</u> difficult, and usually impossible to explicitly separate the effects of multiple pollutants given the present data with which epidemiologists have to work.

P 3-67, L 15-17: There is nothing about "genetic susceptibility" on page 7-52 of the CD, as indicated here. There is one citation on that page referring to hyperlipidemic rats as a susceptibility model, but no statements about genetic susceptibility per se. We know very little about genetic susceptibility among humans, although we all believe it is a factor. In the context used here, any transgenic or selected animal model of susceptibility could be called a study of "genetic susceptibility", but that's pretty circular evidence for genetic susceptibility among the human population.

P 3-73, L 20-21: I recall no evidence that environmental levels of NO_2 cause "irreversible alterations in lung structure". Certainly several studies of animals exposed chronically have not demonstrated such changes at much higher levels. Give a reference if you have one – modify the sentence if you don't.

<u>Chapter 6</u>

P 6-13, L 5-6: It's not clear what is meant by the statement that "a 24-hour averaging time is consistent with the majority of community epidemiologic studies". Just what is the "consistency" to which you refer? The epidemiology is largely based on 24-hour monitoring data, so how could the results be "<u>in</u>consistent" with a 24-hour averaging time? What's the point?

Dr. Roger O. McClellan

Comments on "Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information" (OAQPS Staff Paper – First Draft, EPA-452D-03-001, August 2003)

Roger O. McClellan
Advisor: Toxicology and Human Health Risk Analysis
13701 Quaking Aspen Place NE
Albuquerque, NM 87111
E-mail: roger.o.mcclellan@att.net
Telephone: 505-296-7083
Fax: 505-296-9573

December 1, 2003

A. EXPECTATIONS OF STAFF PAPER

The comments I offer on the draft Staff Paper (SP) are grounded in my view that the SP should serve as a "bridge" between the Criteria Document (CD), which is an encyclopedic exposition of all that is known about Particulate Matter (PM), and the regulatory decisions that must be made in setting the National Ambient Air Quality Standard (NAAQS) for PM. In my view the SP should be a critical science-based analysis of the evidence that bears on the setting of the NAAQS, namely, choices for (a) indicator(s), (b) averaging time(s), (c) numerical levels, and (d) statistical form(s) that will meet the statutory requirements of Section 109 of the Clean Air Act by proposing primary and secondary NAAQS that protect public health and public welfare, respectively. The present document does not meet this expectation.

B. OVER-ARCHING COMMENTS

I offer the following over-arching comments before proceeding to offer some specific comments on the various chapters.

1. Premature Release of Staff Paper

Several years ago I recommended that the Agency provide a draft outline for the PM SP that would lay out the decision-making process the staff intends to use in making recommendations on indicators, averaging times, ranges of numerical levels and statistical forms of the NAAQS for PM. My reasoning was that such a document would provide a basis for discussion and, indeed, debate on the decision-making process without engaging in debates over specific numerical values. I was told that the schedule did not allow time for the approach I recommended. In retrospect time was available to have followed the recommended approach.

The staff pushed ahead on the SP and, indeed, released it prematurely prior to CASAC closure on the PM CD. In my opinion, the result is a "mess." The ORD staff, CASAC and the public are still engaged in vigorous discussion as to what should be included in the CD and its interpretation. The OAQPS staff has issued a draft SP that is based on guesses as to what will be in the final CD and initiated a quantitative risk assessment for PM using concentration-response coefficients and forms that have not yet been agreed upon by CASAC. Perhaps the most serious flaw in what has developed is that OAQPS in the SP has provided draft recommendations for ranges for both $PM_{2.5}$ and $PM_{10-2.5}$ standards. In my opinion, the SP fails to provide a clear road map for the decisions that yielded the draft ranges. Not surprising, the appropriateness of the ranges is already the focus of much discussion and lobbying. In my opinion, the Staff has moved the "regulatory policy train" ahead of the "science train" on the way to revising/setting the PM NAAQS. The Staff, CASAC and the public should still be focusing on the science and an evidence-based decision-making process rather than arguing over the appropriateness of the indicators and ranges.

2. Background of PM Levels

The issue of what are the background levels for the various PM indicators is not adequately covered in the SP. This relates in part to the inadequate coverage of this topic in the CD. This matter requires Staff and CASAC attention and resolution.

3. Baseline Health Statistics

The SP is seriously deficient in not providing coverage of the baseline morbidity and mortality statistics for major cities, regions and the U.S. with special reference to cardiovascular and respiratory morbidity and mortality statistics. It is important that the SP include such statistics to (a) provide perspective on PM-associated health responses, and (b) emphasize their central role in estimate using relative risk models to PM-associated health responses. A cynical view is that the Agency does not want to present the baseline health statistics to avoid providing perspective on the very weak and variable PM-associated health response signal relative to the substantial burden of cardiovascular and respiratory morbidity and mortality from a multitude of risk factors. The wide variation in cardiovascular and respiratory morbidity and mortality across the U.S. emphasizes the need for caution in using concentration-response functions for one region in combination with baseline health statistics for a second region to estimate PM-associated health impacts.

In my opinion, it is becoming increasingly apparent that (a) air quality, including PM and its characteristics, (b) weather, (c) baseline health statistics, and (d) concentration-response functions must be treated as packages on a regional basis. Attempts to force the science to conform to a single national structure may be flawed – one size may not fit all the diverse regions of the United States. (When I refer to regions I do not automatically accept the artificial designation of regions used by EPA. It has some serious flaws, especially as regards the portion of the U.S. west of the Mississippi River.)

4. Concentration-Response Functions

The SP does not adequately address the issue of how concentration-response functions are derived and used. The use of log-linear functions is accepted by default without adequate discussion and considerations of limitations and advantages. Most importantly, the linkage to the underlying health baseline data is not adequately discussed. In my opinion, EPA has been negligent in not exploring in a more rigorous fashion alternative concentration-response functions over the range of contemporary ambient PM concentrations observed in the U.S.

5. Reconciling Use of Log Linear Concentration-Response Functions and Provision of an Adequate Margin of Safety

In failing to provide a clear science evidence-based strategy for setting the NAAQS for PM, the SP does not address how the Agency will reconcile the use of log linear concentration-response functions and the setting of standards with an adequate margin of safety. This is a glaring deficiency in the present document. In short, how low will be low enough? Is the use of log linear functions to set a standard that results in a calculated excess of 1, 10, 100, 1000, 10,000 deaths per year or some other number consistent with an adequate margin of safety? Or does the Agency propose to use some level of statistical certainty (or uncertainty) as an indicator of having achieved an adequate margin of safety? For example, one could argue that if a PM_{10} standard (24 hour average) were being set the NMMAP's data for the 86 cities not having a statistically significant PM effect could form the decision on the 24 hour PM_{10} standard.

At some point, the Agency has a responsibility to share with CASAC and the public how it intends to bridge from the available PM science to setting the NAAQS for PM. In my opinion, the current use of log linear concentration-response models that are highly uncertain and a myriad of mathematical calculations in the absence of a decision structure is not adequate.

6. Staff Paper Organization Does Not Support Decision-Making on PM NAAQS

The present SP is not organized in a manner that clearly presents the science that under-girds the setting of the PM NAAQS. Chapter 1 fails to provide a clear road map as to how the science will inform decisions on setting the PM NAAQS with an adequate margin of safety to protect public health. I have recommended inclusion of a separate chapter to provide this road map.

Chapter 2 is scientifically interesting but excessively long and fails to provide key atmospheric science information on background levels of PM germane to setting the PM NAAQS.

Chapter 3 has two key deficiencies. First, it fails to present key baseline health data. Second, it is not a balanced exposition of information on PM-associated health responses.

Studies with statistically significant or marginally significant responses are emphasized and studies that are not statistically significant are ignored. The influence of weather and co-pollutants as confounders has been understated.

Chapter 4 is a premature application, in view of unresolved issues concerning the underlying data base of risk assessment techniques to estimate excess PM-associated effects.

I will defer to others with regard to Chapter 5 on welfare effects.

As I have noted earlier, Chapter 6 on Staff Conclusions and Recommendations on PM NAAQS was released prematurely and is already serving to polarize several sectors of the public with debate focusing on their opinion on the ranges presented rather than the science under-girding the indicators and ranges. As noted earlier, the SP does not provide a clear exposition on the decision criteria to be used in selecting ranges (and statistical forms) for the several PM indicators.

C. <u>SPECIFIC COMMENTS</u>

1. <u>Chapter 1. Introduction</u>

This chapter provides an adequate introduction to the document with one glaring exception. It would be useful for the chapter to conclude by noting that a subsequent chapter in the document will provide the strategy for evaluating the evidence relevant to setting the NAAQS for PM. I will refer to this as the first missing chapter.

This missing chapter on "science evidence-based decision-making for the setting of the NAAQS for PM" is a serious deficiency. There is a need for a clear road map as to how the staff intends to integrate, summarize and use the available scientific evidence for recommending indicators, averaging times, ranges of numerical levels and statistical forms that will protect public health with an adequate margin of safety.

The failure to include such a chapter has already led to substantial misunderstandings over how the PM evidence is to be evaluated. This includes serious charges that the SP lacks clear ground rules for science-based decision-making and, instead, has engaged in a "cherry picking" approach to selecting and using literature (and, indeed, parts of studies) to support a position that the $PM_{2.5}$ standard should be "tightened" and a new $PM_{10-2.5}$ standard promulgated.

I urge the Agency to provide a new chapter in the revised SP that provides a strategy, with specific ground rules, for evaluating and integrating the scientific evidence relevant to the setting of the NAAQS for PM.

2. Chapter 2. Air Quality Characterization

In general, this is a well-written and informative chapter. However, it needs to be improved in two ways. First, the language that bridges from the scientific language (coarse, fine and ultra-fine and accumulation modes) used to describe particles and the monitoring language (PM_{10}, $PM_{10-2.5}$, and $PM_{2.5}$) needs to be reviewed to make certain the terms are properly used. Second, discussion of the critical issue of background levels of PM_{10}, $PM_{10-2.5}$, and $PM_{2.5}$ needs to be strengthened. Indeed, the CD coverage of background levels of PM should also be revisited and revised.

pg 2-14: It would be useful if some quantitative data could be provided on the contribution of precursor gaseous emissions to $PM_{2.5}$ levels to complement the data in the Primary Emissions, $PM_{2.5}$ column of Table 2-3.

pg 2-25, line 10: An explanation should be given as to the factors influencing the occurrence of the high 24-hour average $PM_{2.5}$ concentrations in California and Montana since they are different.

pg 2-41, line 22: It would be appropriate to add a sentence such as — "The substantial regional variation in the ratio of annual mean $PM_{2.5}$ to PM_{10} from about 0.35 to 0.65 indicates that it is not appropriate to use PM_{10} as an indicator for $PM_{2.5}$ by using national average value. By the same token, PM_{10} levels cannot be used as indicators of $PM_{10-2.5}$ levels."

The chapter could be improved by including a discussion, perhaps after the present Section 2.4, on how measurement techniques have changed over time. The discussion should emphasize the uncertainty in "translating" past PM measurements into current FRM measurement values. These uncertainties need to be considered in interpreting and using exposure-response coefficients from different studies.

3. Chapter 3. Characterization of PM-Related Health Effects

The chapter could be substantially improved if it were to summarize current knowledge in a more direct fashion without confusing the issue by referring to the 1996 CD and then what has been learned since 1996. In my opinion, this approach was confusing and inappropriate in the CD. The approach is even more inappropriate in the SP. The SP needs to make clear the scientific criteria that are relevant to setting the NAAQS for PM based on current knowledge, without regard to when it became available. The chronology of when the information was developed is irrelevant to how the information is evaluated *in toto* today.

pg 3-26, Figure 3-4: The figure should be modified to show separately (a) total mortality, (b) cardiovascular or circulatory mortality, (c) respiratory mortality, or (d) cardiorespiratory mortality. It is not appropriate to lump (d) with (b) as shown in the current figure.

pg 3-27, Figure 3-6: Same comment as above.

pg 3-28, Figure 3-6: Same comment as above.

All of the figures should be carefully reviewed to determine if the units used are adequately identified in the figure or legend, i.e., excess effect per 1, 10, or 25 µg of PM indicators. Indeed, one can argue that the most scientific approach would be to always state excess risk per µg of PM indicator.

Whenever laboratory animal or controlled human studies are cited care should be taken to clearly indicate the route of exposure, duration of exposure, quantity administered or concentration in the air and when health measurements were made.

pg 3-65: The discussion of sensitive groups is seriously deficient in failing to note the role of cigarette smoking as a major determinant of cardiovascular and respiratory morbidity and mortality. Smoking is the major determinant of "pre-existing respiratory and cardiovascular disease" (pg 3-65, line 17). It follows from this and the use of relative risk models that the majority of any excess PM-associated health responses will be in smokers. Why does the Agency not want to make this point clear?

pg 3-87: If the Laden *et al* (2000) study is to be included, then it will be appropriate to include Graham and Hidy (2003) that identifies some serious shortcomings in the Laden *et al* (2000) analysis.

The present organization of Chapter 3 does not present the information on PM-Related Health Effects in an optimum fashion to understand how it will be used in setting the NAAQS for PM. I urge the Staff to revise the structure of the chapter so it is aligned with critical issues in setting the NAAQS. Specifically, it would be useful to provide a brief introduction noting that the evidence available allows consideration of three potential indicators (PM_{10}, $PM_{2.5}$ and $PM_{10-2.5}$) and for each indicator consideration of two potential averaging times (annual and 24 hours).

This introduction would be followed by sections on each indicator. A suggested outline for a revised Chapter 3 is shown below.

1. Introduction
2. Nature of PM-Associated Health Effects
3, Baseline Health Effects Data
4. Exposure Concentration-Response Models
5. Epidemiological Evidence
    a. PM_{10} Indicator
        (1) Long-term Exposure (Annual Standard)
        (2) Short-term Exposure (24 hour Standard)
    b. $PM_{2.5}$ Indicator
        (1) Long-term Exposure (Annual Standard)
        (2) Short-term Exposure (24 hour Standard)
    c. $PM_{10\text{-}2.5}$ Indicator
        (1) Long-term Exposure (Annual Standard)
        (2) Short-term Exposure (24 hour Standard)
6. Supporting Evidence
    a. Controlled human exposures
    b. Laboratory Animal Studies
7. Coherence
8. Summary

The present chapter is seriously deficient in not providing background information on various indices of mortality and morbidity. Hence, the recommendation for section 3 above. The appropriate presentation of such indices should include statistics for the U.S. and selected cities and regions. The associated discussion should note the most important factors associated with differences in the regional statistics such as age and smoking. Presentation of the baseline data is important because of its role in using relative risk models.

A brief discussion is needed on exposure concentration-health response models. Hence, the recommendation for section 4 above. This section will follow naturally from the previously requested material on baseline health statistics.

The information on epidemiological evidence should be organized in a manner that considers the evidence relative to specific (a) indicators (PM_{10}, $PM_{10-2.5}$, and $PM_{2.5}$) and (b) averaging times (24 hr and annual) as suggested for Section 5.

Chapter 4

A major challenge in reviewing Chapter 4 is the frequent need to refer to the technical support document (Abt 2003). I urge the Staff to consider placing in the SP certain key data that is presented in the technical support document. Specifically, it would be useful to include in this chapter the baseline health statistics, population sizes, and concentration-response coefficients used in any quantitative analyses.

The approach to treatment of "thresholds" needs to be more clearly presented.

In many of the figures in Chapter 4, the lower bound values have been truncated at zero. The result is to seriously misrepresent the results to a casual reader. The calculated values below zero are as real as the calculated values greater than zero.

In the captions for Figures 4-15 a and 4-16A, it would be useful to include the annual averages for the $PM_{2.5}$ and $PM_{10-2.5}$ measurements, respectively. In Figure 4-15b, it would be useful to relate the under-lying non-accidental mortality to provide perspective to the calculated excess PM-related mortality.

Dr. Günter Oberdörster

[Sent via e-mail to Dr. Les Grant, Director of EPA's National Center for Environmental Assessment (NCEA)/RTP, on November 24, 2003]

Subject: Fw: Comments from Dr. Gunter Oberdorster

Dear Les,

Sorry I am late with sending you some comments as we discussed at the CASAC meeting, I was busy getting a major grant out last week.

1. With respect to the results of modeling human COPD lung deposition, I was mistaken when I thought that the results had already been published by Werner Hofmann. He is still writing the paper, only two abstracts have been presented this year at the ISAM Conference in Baltimore in June. Attached are the two abstracts from his modeling efforts, and although you may not be able to cite them - it gives you at least an idea that modeling of particle deposition in human diseased lungs is being done.

2. Table 3-1 of the staff paper summarizes mechanisms of particle effects. The potential mechanisms listed in that table are not really mechanisms but most of them are just effects. For example, in the first category "Direct Pulmonary Effects" you may want to change the so-called "mechanisms" in the first line to something like "activation of alveolar macrophages, epithelial cells"; in the second line, the effect is possibly related to "increased oxidative stress" as a mechanism or decrease in antioxidant defenses"; and in the third line, a potential mechanism may be "stimulation of irritant receptors or sensory nerves in the tracheobronchial region".

In the second category, "systemic effects secondary to lung injury", you could add in the first line "due to pulmonary vasoconstriction, edema". All of these systemic effects listed in the table probably involve acute phase responses with increased acute phase proteins such as IL-6 and others. (I don't see why arrhythmia in the second line of this category is defined as hemodynamic effect?).

In the third category, "direct effects on the heart", you could add as a mechanism for the autonomic control of the heart "via sensory nerves in the tracheobronchial region, connected to vagal ganglia".

I hope this helps.

With kind regards - Gunter

GO/jh

Hofmann_absts.tif

EFFECT OF INHOMOGENEOUS LUNG VENTILATION ON PARTICLE DEPOSITION IN HEALTHY SUBJECTS AND COPD PATIENTS

Elzbieta Pawlak, Robert Sturm and Werner Hofmann

In current particle deposition models, the inhaled air is distributed uniformly throughout the airway system. However, experimental studies have indicated that the human lung is not ventilated homogeneously, but rather in an asymmetric and asynchronous fashion. In this study, a mathematical ventilation model is presented, which describes the asymmetry and asynchrony of lung ventilation by specific time-dependent coefficients for the individual lobes. The effect of inhomogeneous lung ventilation on particle deposition was studied for both healthy subjects and patients suffering from COPD (mean reduction of the airway calibers: 50%) over a wide range of particle sizes. Under standard breathing conditions, particle deposition in COPD patients is significantly increased compared to that in healthy subjects. Depending on particle size, this increase ranges from 2 to 20%. In both healthy and diseased lungs, asymmetric and asynchronous ventilation causes a decrease of particle deposition due to changes of air volumes entering the single lung lobes and a partly significant reduction of particle transport times in the lobes. While asymmetry reduces particle deposition by 5 to 10%, asynchrony leads to a respective drop of up to 20%. Finally, combination of both effects causes a maximum decrease of total deposition by about one third.

Funding: EU Contract FIGD-CT-2000-00053

SIMULATION OF EMPHYSEMA IN THE HUMAN LUNG AND ITS EFFECT ON ALVEOLAR DEPOSITION

Robert Sturm and Werner Hofmann

Emphysema is defined as a slowly progressive, pulmonary disease characterized by the continuous enlargement of air spaces distal to the terminal bronchioles. The abnormal distension of alveolar structures is mainly generated by the permanent destruction of alveolar walls due to a decomposition of their major structural proteins elastin and collagen. In the theoretical approach presented here, different types of lung emphysema (i.e., paraseptal, centriacinar, panacinar, and bullous emphysema) are modelled (a) by defining a distribution of alveolar diameters based on histological sections of diseased lungs and (b) by determining the region within the lung, where a specific type of emphysema preferentially occurs (e.g. paraseptal emphysema in the outermost airway generations). Alveolar deposition of inhaled particles was calculated for patients exclusively suffering from emphysema ('pink puffers') and subjects with a combination of COPD (maximum airway reduction: 50%) and emphysema. For both scenarios, total alveolar deposition is significantly decreased relative to healthy lungs, strongly depending on the considered type of emphysema. Most significant reductions of deposition, as high as 90%, can be observed for bullous emphysema. An additional assumption of volume changes during progress of the disease (increase of FRC, decrease of TV) causes a further decline of the alveolar deposition values.

Funding: EU Contract FIGD-CT-2000-00053

J. Aerosol Medicine 16(2): 224, June, 2003

Dr. Robert D. Rowe

Memorandum

To: Fred Butterfield, Phil Hopke

From: Bob Rowe, Stratus Consulting Inc.

Date: 11/4/2003

Subject: First Draft PM Staff Paper

Modest revisions are provided to my draft comments.

This is a good first draft, and generally well written and presented.

A. Selection of averaging times and levels. Sections 6.3.2, 6.3.3, 6.3.4 for $PM_{2.5}$, and earlier supporting text; and the same for $PM_{10-2.5}$. There are important issues regarding the selection of averaging times and levels. These issues are not sufficiently, or explicitly addressed.

a. Strength and use of evidence for effects at low ambient concentrations and clarity on items factoring into the margin of safety are not made clear.

b. Emphasis on the annual average over the 24 hour measure, for both a $PM_{2.5}$ and a PM_{10} or $PM_{10-2.5}$ is not clear or convincing. Certainly only a few high days will contribute only a small percent to the total risk, but many modest days will contribute a lot to risks. While current evidence suggests using a linear concentration-response function, and little evidence about thresholds, my sense of the panel is that there is much more comfort that a 5 ug/m^3 reduction between 20 and 25 would benefit the public than would the same 5ug/m^3 reduction between 15 and 10, or between 10 and 5. Thus, we should want to emphasize reducing the modest and bad days, rather than further reducing the already good days – but an annual standard does not distinguish between the two (although it perhaps likely that the reduction in the annual average is from reductions on the modest and bad days but it is not necessarily the case). The annual standard deflects control emphasis from episodic controls, which in some locations may provide the most desired reductions in health risks.

c. Basis for selecting a 24 hour standard (pages 6-22 to 6-24), the appropriateness of second statistical approach is unrelated to the health impacts that are being experienced or could be experienced if certain levels were allowed.

d. Future standard setting may benefit from research examining measures such as "dose-days over X ug/m^3" in the epi studies, where X may evaluated for values like 20, 25, 30, 35, 40 ug/m^3, to evaluate critical dose, rather than annual average measures.

e. The interrelated discussions on thresholds, linearity, effects at low levels, and on selecting levels for annual average and 24 hour indicators are scattered in the report

(Section 3, Section 4.6.2.1, Section 6.3) and discuss some of the same, and some different, literature in each place. This should be cohesive in Chapter 3 and/or Chapter 6. The discussion in Section 4.6.2.1 is misplaced as Chapter 4 risk assessment procedures and not presenting the background for the issue.

B. Risk Assessment.

1. Needs more clarity on the intent (goals) and weight to be placed on the risk assessment (see notes below) in the Staff Paper recommendations. This determines the importance to place on many of the detailed issues with the risk assessment (linearity, higher level cumulative uncertainty analysis, etc.).

2. The visual presentation of results that truncates the statistical distributions is a concern and misrepresents the results, even though the argument that negative values (pollution is good) is reasonable. Alternatives were presented in the meeting and one should be selected and implemented.

3. Spatial averaging versus single monitors in standard setting. This issue received little attention in the document and in the meeting, and needs to be more explicitly addressed.

C. Chapter 4: Risk Analysis

1. The use of the risk assessment in the overall process is not as clear as it might be (here or in the introduction). The individual health effect studies identify the potential health effects that clearly are adverse to the affected individuals. The risk assessment helps provide perspective on the level of adversity for the public as a whole. If there are no effect thresholds, not all adverse events may be able to be avoided. Thus, is the population impact potentially important?

2 Page 4-3 lists the goals, which seem to be in inverse order of importance, which is: (1) provide a rough sense of magnitude of risks under current conditions and alternative regulatory strategies (otherwise the other goals are academic). (2) understanding the nature of the risks (is it mortality, morbidity; to what populations, etc.) (3) understanding the importance of various uncertainties and factors in the assessment.

3. The thresholds and linearity discussions found in section 4.2.6.1 (page 4-32 and thereafter) belong in section 3.5 page 3-75 rather than identified on page 3-75 with reference to Chapter 4 and/or in chapter 6.

4. A 1996 finding is used in Section 6.3.2 to discuss averaging times and motivate a focus on an annual average standard: "the few peak 24-hour $PM_{2.5}$ concentrations appeared to contribute a relatively small amount to the total health risks posed by the entire air quality distribution as compared to the aggregated risks associated with the low to mid-range $PM_{2.5}$ concentrations…"(page 4-4, with concept repeated on page 6-13). This is obvious (e.g., a few days won't drive the assessment, especially with linear functions and most of the population not experiencing the peak days. However, the unstated point is that to avoid most

risks would be to set the 24 hour standard at low to mid-levels, reflecting the literature and risk assessment.

D. <u>Chapter 5: Welfare Effects (and related sections in Chapter 6)</u>

Most comments address Section 5.2.5 on the significance of visibility to public welfare.

1. The recent "attitude" studies regarding what is adverse are emphasized. On a scientific basis, the current studies are not strong, but are informative. They are subject to some the same, and some different, issues as in the valuation studies. Attitude and valuation studies should be seen as compliments, not substitutes. For example, some of the valuation studies addressed related issues – Carson et al. (APCA speciality conference proceedings, 1990) and McClelland and Schulze for urban settings, and Chestnut and Rowe (1990) for Class I areas find that visibility impacts on a few days has non-trivial value. One might be able to review these to evaluate the visibility impairment levels for these days to relate these prior economic studies to the more recent attitude studies, although that is not a priority here.

2. Page 5-18 line 7, Page 5-19 line 26: add Chestnut and Rowe (1991), which covers more studies than Chestnut et al. 1994. Chestnut, L.G., and R.D. Rowe. (1991) Economic valuation of changes in visibility: A state of the science assessment. Sector B5 Report 27. In Acidic Depositions: State of Science and Technology Volume IV Control Technologies, Future Emissions and Effects Valuation. P.M. Irving (ed.). The U.S. National Acid Precipitation Assessment Program. GPO, Washington, DC.

3. Pages 5-18 to 5-19 discussion of use and non-use values: (i) Use values include improved aesthetics during daily activities (driving or walking, looking out windows, daily recreation) and for special activities (visiting parks and scenic vistas, hiking, etc.), and viewing scenic photography. Merge in the option value concept. A significant component of value is tied to preserving improved visibility in the event of a visit, even though a visit is not certain. This key component of the measured values is considered by some in use values and by others in non-use values.

4. Page 5-19 mid-line 6 to start of line 9. Remove sentence as it repeats earlier text.

5. Page 5-19 lines 11-16. The lead sentence is distinct from the remainder of the paragraph (which only shows there is a lot of visitation expenditures). The lead is tied to the next paragraph. Perhaps the order of follow-up material should be reversed. I believe there are a few items in the literature for US sites linking visibility to visitation (it would take some time to find them).

6. Page 5-20 Lines 5 – 14. Not all of the valuation studies are CV. Some use hedonic property values, with their own issues of separating visibility from other air pollution impacts (although complimented by surveys can provide indications of the partitioning of hedonic values, which consistently indicates that the visibility component is significant).

7. Page 5-27. It is important to note that each of the studies is in a western environment, and implications regarding eastern (or other) locations cannot be made.

8. Section 5.2.6.2 and similar discussion on pages 6-44 and 6-45. Section 5.2.6.2 should be dropped. The Washington, D.C. work has not be developed sufficiently to warrant this level of discussion in the Staff Paper. The pilot can be cited with the earlier similar public preference attitude studies, but given the limited work on this study, probably no more than as a small note with proper caveat. The similar discussion on pages 6-44 and 6-45 does not appear to belong in the Staff Paper (or at least not at this length) and would be better reserved for a research agenda report.

9. In EPA's cover letter, it requested input on the proposed public attitude studies. See also my comments provided July 30, 2001.

   - The approach has promise, but needs to be conducted in defensible manner (with an advisory panel and peer review), and any new work needs to begin to address a number of issues with the approach. Generally, the approach need not be viewed as an alternative to valuation approaches, but as a complement to these approaches, including both types of questions.

   - More important is that EPA (and/or others) should do more on public losses from visibility impairment, which could lead to setting a secondary standard to reflect welfare impacts, rather than setting the standard to simply match the primary standard. Considerable research identifies that visibility impairment present a substantial public impact. A considerable literature and data base exists for visibility impairment and the public's reaction. A combination of prior research and new research on public preferences would greatly aid to develop a secondary standard. Research along the lines of this project could go a long way toward that objective.

   - Finally, referring to the interviews as "focus groups" may be correct for some past applications, but probably would not be correct for properly conducted future surveys (see edit below for page 6-44).

10. A key conclusion of the limited visibility attitude studies and some of the economic studies is that visibility impairment of modest amounts on a few days is perceived as adverse, which relates to a 24-hour secondary. Other economic studies support reduced annual averages, but the SP does not relate these to any underling annual average concentration levels.

11. Regarding materials impacts (section 5.3.2), there are economic studies that identify values for impairment to cultural resources (such as marble monuments and historic buildings). For example see Morey et al. and the citations therein. [Morey, E.R., K.G. Rossmann, L.G. Chestnut and S. Ragland (2002). Valuing reduced acid deposition injuries to cultural resources: marble monuments in Washington, D.C., in *Valuing Cultural Resources*, S. Navrud and R.C. Ready (eds.), Edward Elgar, Cheltenham, UK, ISBN I-84064-079-0.

E. Minor Text Edits

1. Page 2-26, Figure 2-8 and Page 2-35 Figure 2-13. In the caption, correct "sties" to "sites".
2. Page 2-51, Line 14. Add space in text "Chapter 4and".
3. Page 2-66 Solomon reference. Move date to after authors.

4. Page 3-31 Line 28 “a only” -> “only a”.
5. Page 3-42 regarding life shortening. Need to be clear that the Staff Paper conclusions are by no means definitive and more research is needed.
6. Page 3-76 Line 27. “wiggly” should be replaced with a better description of the concept being alluded to.
7. Page 4-10 Line 3. Should “circle” be “diamond”?
8. Page 4-10 lines 24/26, pg 4-13, line 18 (and related discussions). Should this be US and Canadian. In some places the text refers to reliance on US studies, and in others on US and Canadian. Needs to be consistent.
9. Page 4-13 line 13 – add developmental effects to list of potential but excluded effects.
10. Page 4-27 line 19. Is the y_0 here suppose to be x_0?
11. Page 4-33 footnote 13. More discussion is needed or cross reference to the extended discussion of this matter (now in Chapter 6, but perhaps belongs in part in Chapter 3).
12. Page 4-36 line 25, “term pm exposure” -> “term exposure”
13. Page 5-69 Line 17. “Preservation of values…” -> “Preservation values…”
14. Page 5-71 line 41. “Ben-Davis” -> “Ben-David” and line 41 “Molenar Jr.” -> Molenar J.”.
15. Page 5-31 line 24 is “sorb” to “absorb”?
16. Page 5-64 line 18 “lass” -> “less”.
17. Page 6-23 line 10 “…can provide an appropriate basis for” -> “… can provide useful input to…” to be consistent with similar text elsewhere.
18. Page 6-44 line 6. Remove “focus groups to elicit” (see forthcoming general comments regarding this point).
19. Page 6-49 line 20. Recommend cross reference to materials in SP or CD for the assertion.

Dr. Jonathan M. Samet

Review Comments: OAQPS Paper—First Draft
Jonathan Samet, M.D., M.S.
December 1, 2003

General Comments:

This first draft of the "Staff Paper" is clearly a "work in progress", and consequently these comments are made in that context. As a first issue, in fact, I would urge the staff to consider preparing a far more approachable and "reader friendly" document. At present, the document reads as though segments of the Criteria Document had been juxtaposed with staff interpretation of these critical segments. Of course, this is the intent of the Staff Paper, but the current form of the document precludes gaining a full picture of the new findings and the implications for the NAAQS.

In this regard, it would be useful if each chapter were to highlight, perhaps in a tabular or bulleted form, the state of the evidence at the last Staff Paper and, the relevant incremental gains in knowledge since then. With this summary in hand, any new recommendations with regard to the NAAQS would have a transparent basis.

In my view, the Staff Paper remains bedeviled by the same fuzziness around critical concepts as the Criteria Document, particularly in relation to confounding, effect modification, and causality. There is a sloppiness in the language around these concepts that leads to ambiguity of interpretation. In particular, the document does not carefully separate the quality and extent of the evidence available from the conclusions that might be reached. Some examples will be highlighted in my more specific comments and I have previously commented on this issue in regard to the Criteria Document.

The process of information gathering and synthesis embodied in the Criteria Document and the Staff Paper is very much akin to the conduct of an evidence-based review. As the formalism of evidence-based reviews has evolved over the last decade, emphasis has been placed on elaboration of a clear set of principles for developing the evidence-base, for evaluating the quality of evidence, and for reaching conclusions. Typically, a review provides guidance in its earliest pages as to the approach that was followed.

In this regard, the Criteria Document and the Staff Paper are both deficient. A reader would not be able to judge how conclusions were reached in either document, nor to find any explicit statement as to what principles were followed. In fact, there is substantial variation across the staff paper in the apparent use of criteria for evidence information and the related language for causality. These should be stated explicitly at the start of each document and, in fact, readers and users of these documents should be confident that the stated principles were followed as evidenced by their specific application across the documents.

The proposal for a $PM_{10-2.5}$ (coarse) NAAQS follows the Supreme Court decision and the need to have a regulatory approach for controlling coarse particles. The $PM_{10-2.5}$ size fraction has thus emerged, not primarily on a biological basis, but as a consequence of non-biological happenings. The Staff Paper nonetheless finds support for a $PM_{10-2.5}$ standard, in part on epidemiological evidence.

The agency faces a dilemma in that $PM_{10-2.5}$ is an artificial construct, slicing one segment of the coarse particle mode. Particles in the size range above PM_{10} however, do enter the upper respiratory track and have the potential to cause injury locally as well as more distal/general effects. Epidemiologists commonly study $PM_{10-2.5,}$ because this is the only coarse size fraction for which data are routinely available. There is inherent circularity in justifying a $PM_{10-2.5}$ standard <u>because</u> that is what can be studied.

The Staff Paper reviews the relevant information that would inform having a $PM_{10-2.5}$ standard. Some epidemiological data are presented and a relatively strong interpretation is given emphasizing the "coherence" and finding increased "support for a causal link". The new evidence is not substantial and I have concern about the laxity of evidence interpretation.

Chapter 1: No specific comments.

Chapter 2: Page 2-1, Line 26: Is the distinction between fine particles and coarse particles overstated?

Page 2-2, Line 12: Particle size does not really determine exposure, as stated here. Rather, it determines concentration in relationship to the source of the particles.

Page 2-19, Line 18: I am not certain that the NMMAPS researchers revised these regions, as stated.

Page 2-37, Lines 1-5: Epidemiological implications should be addressed here.

Page 2-41, Lines 11-12: What is the relevance of the statement concerning epidemiological studies?

Page 2-59, Lines 14-17: This statement is not correct as written, as exposure measurement errors do have implications for the magnitude of effect estimated and the precision of estimates.
Pages 2-60, Lines 9-12: This statement should be specified as describing cross-sectional findings.

Page 2-61, Lines 1-2: There are implications for more than "time-series epidemiology". I would also suggest using terminology other than "time-series epidemiology" to describe time-series studies. There should be consistency in reference to various research designs throughout the document.

Chapter 3:

Page 3-2: Here, would be a useful point for setting out the methodologic frame of reference for addressing confound, effect modification, and causation.

Page 3-3, Lines 18-23: These statements are far too sweeping. There are many potential mechanisms by which particles may cause adverse health effects and neither the information available in 1996 nor at present is so conclusively uninformative as stated.

Page 3-5, Lines 24-26: Another similarly sweeping statement.

Page 3-6, Lines 1-3: The same wrong thinking continues here with the proposition that one could "fully define" mechanisms. This is determinism taken to a non-useful extreme. The remainder of the paragraph is similarly off the mark.

Page 3-7, Table 3-1: An arrhythmia is not a "systemic hemodynamic effect". I am also unclear as to what is meant by "PM/lung interactions potentially affecting haematopoiesis".

Page 3-9, Lines 10-12: This sentence is too vague.

Page 3-9, Line 15: Yes, particles are potentially quite rich in their chemical composition. This concept is not well covered in the material to this point.

Page 3-9, Lines 21-25: Again, a continuation of a never-ending search for mechanisms.

Pages 3-10, 3-12: This material needs to be substantially sharpened. I have previously given comments around the Criteria Document that may be useful. Conceptually, the materials are simply too ambiguous.

Mr. Ronald H. White

Comments of Ronald H. White, M.S.T.
EPA First Draft Particulate Matter Staff Paper and Risk Assessment
November 17, 2004

General Comments

Chapter 3

The First Draft PM Staff Paper (SP) provides a generally well written summary of the results of key studies on the health effects of particulate matter. The interpretation and relationship of the key health effects studies to the policy issues regarding the adequacy of the current PM NAAQS and the proposed revisions to the PM NAAQS are also generally appropriately described. The SP does a generally good job of integrating the information on exposure with the results from key epidemiological studies, with appropriate reference to supporting information from controlled human exposure and animal toxicology studies.

Revisions to the discussion of the selected studies drawn from Chapters 7 and 8 of the Fourth Draft PM Criteria Document (CD), as well as from the Integrated Synthesis (Chapter 9), will be needed to reflect the revisions made to those chapters in response to CASAC and public comments on the CD.

A set of criteria or rationale should be provided for the selection of key studies included in the SP.

To provide the reader with a context for the discussion of sensitive groups in Section 3.4, this section should include information on the national magnitude of the sensitive group populations.

Chapter 4

EPA staff is to be commended for the transparent discussion of the selection of health endpoints, locations, and dose-response functions selected for use in the risk assessment.

While noting that negative risk coefficients from epidemiologic studies do not logically represent a beneficial effect from PM exposure, the presentation of negative value lower bound risk estimates in the risk assessment results presented in Figures 4-5 through 4-14 should not be truncated at zero to ensure an accurate presentation of the actual risk coefficient results.

The discussion of mortality and morbidity concentration-response functions and thresholds for health effects should be included primarily in Section 3.3 of Chapter 3 and that information can then be referenced as necessary for the risk assessment discussion in Chapter 4.
Recommendations for additional sensitivity analyses:

1) The use of only a single year of $PM_{2.5}$ data raises a concern regarding the representativeness of the data year selected to typical levels. While relying on a single year of data may be unavoidable for $PM_{10-2.5}$ data due to the limited amount of available monitoring data, EPA's AIRS data base contains a considerable amount of multi-year $PM_{2.5}$ data for several of the cities used in the risk assessment. As a component of the sensitivity analysis, EPA staff should compare the impact of using three years of $PM_{2.5}$ data (e.g. 2000 – 2002) with the single year of data used in the base case risk assessment for a representative sample of cities where multi-year data is available.

2) Add threshold value for $PM_{2.5}$ of 8.0ug/m^3 for short-term mortality analysis. Since the low end of the range of mean city values for the Burnett et al., 2000 study (reanalyzed in Burnett and Goldberg, 2003) is below 10ug/m^3 (9.5 ug/m^3, from Table 3-2), an analysis should be conducted of the impact of a threshold value lower than the lowest city mean $PM_{2.5}$ value associated with an increase risk of mortality.

3) Analyze health outcome risks for alternative forms of the 24-hour $PM_{2.5}$ standard range (e.g. one annual allowable exceedance, 99th percentile, fourth highest value over three years).

4) Compare city-specific health outcome risks associated with meeting the annual $PM_{2.5}$ standard using the maximum vs. average of monitor-specific averages (spatial averaging).

<u>Chapter 6</u>

The proposed emphasis on the annual $PM_{2.5}$ standard as the "controlling standard" is appropriate in the context of shifting the overall annual concentration distribution downward. However, an effective and health protective 24- hour standard is needed to protect against repeated peaks such as those that occur on a seasonal basis (e.g. wood burning, agricultural burning). These acute exposures have been associated with significant morbidity, and repeated episodes may exacerbate the disease condition of sensitive populations such that they are then vulnerable to increased risk of mortality.

While the health evidence for setting a coarse PM standard is substantially less compelling than the evidence for fine PM, the results from studies of noncrustal source coarse PM regarding exacerbation of morbidity, and to a lesser extent increased risk of mortality, as discussed in the CD and SP provide a reasonable basis for setting a coarse PM standard based on a precautionary public health approach. An analysis of key PM_{10} studies where data on the fine fraction (e.g. sulfates) is available and can be removed from the analysis, as suggested by Dr. Lippmann at the November 13, 2003 CASAC meeting, would provide additional confidence in establishing a coarse PM standard.

Specific Comments

Chapter 3

Pg. 3-2, line 28: I would suggest that the results of the intervention studies be described as “avoided excess mortality and cardiopulmonary morbidity” rather than “improvements…in health”.

Pg. 3-24, lines 11-13: The statement that there is little difference in epidemiologic study results when comparing GAM with stringent convergence criteria and GLM approaches is not necessarily true for all studies. See Aphea-2 reanalysis adjusted GAM vs. natural spline results as an example.

Pg. 3-94, line 29: The Dublin intervention study by Clancy et al. should be referenced here as well.

Chapter 4

Pg.4-30, lines 23-24: If the finding of a statistical significant result was not a criteria for the selection of studies for inclusion in the risk assessment, a description of the study selection criteria actually used should be included here.

Pg. 4-31, lines 17-19: Some additional explanation and justification for the statement that the corrected GAM model provides a better effects estimate than the GLM model should be provided here.

Pgs. 4-63, 4-65: Tables 4-10 and 4-11 presenting sensitivity analysis results for different from base case threshold models and historical air quality data indicate that risk estimates less than zero were truncated at zero. However, it is my understanding from the discussion of the risk assessment protocol at the November 12-13 CASAC meeting that negative risk values were truncated only for the purposes of presentation and not for risk calculation. If that information is correct, the statement that negative risk values were truncated at zero should be removed from these tables.

Chapter 6

Pg. 6-2, lines16-20: Beyond establishing “natural background level” estimates for the East and West, it is unclear what additional “risk management implications” that have been considered in the SP. The legislative history of the Clean Air Act and legal precedent are clear that implementation costs and technological feasibility of attainment are not to be considered by the EPA Administrator in establishing the NAAQS level “requisite to protect the public health…with an adequate margin of safety”. The risk management considerations discussed here should be explicitly stated.

Dr. Warren H. White

The roles of background levels and averaging times in the PM RA

Warren H. White, 11/17/03

Ozone provides one risk-assessment model for pollutants with substantial contributions – *backgrounds* – from natural and extra-continental sources. The ozone background varies slowly, and within a limited range, and may thus be usefully approximated as a constant, at least within seasons and geographic regions. Scavenging by NO_x emissions can drive ambient concentrations below background levels, and the resulting ozone deficits must not be miscounted as benefits.

These considerations motivate an analytic framework in which annual risk is appropriately calculated as the sum of the daily risks associated with the excesses of 24h concentrations over a fixed threshold representing the background. It is this framework that the PM RA employs.

Risk assessment for PM requires a different analytic framework, because PM backgrounds relate differently to ambient concentrations. Ambient PM concentrations are just the sum of the background and controllable anthropogenic fractions, and thus are always at or above background levels. It is therefore unnecessary to disaggregate annual risks into their daily increments to avoid including spurious benefits. It is fortunate that annual averages can be used, because 24h PM backgrounds can vary greatly from day to day and are not easily determined.

The excess risk attributable to a 24h concentration A_i of controllable anthropogenic PM is $H[\exp(\beta A_i) - 1] \sim H\beta A_i$, where H is the baseline incidence of the health effect and β is the coefficient of the C-R function. (My 11/10/03 comments, attached below, address the validity of this linearization.) The annual attributable risk is accordingly $H\beta\Sigma A_i = H\beta 365 M_A = H\beta 365(M_{PM}\text{-}M_B)$, where M_A, M_B, and M_{PM} are respectively the annual-average controllable, background and total concentrations. (Actually, this conceals another approximation if seasonal variations in H are acknowledged to correlate with those in A_i.)

Given the C-R model underlying essentially all our epidemiological results, **the risk attributable to controllable PM is thus a function of only the average ambient and background concentrations, not their day-to-day variations. This holds, according to our assumptions, regardless of whether the coefficients H and β refer to acute or chronic effects.**

The annual risk reduction achieved by passing from daily concentrations of PM_i to PM_i^* is, similarly, $H\beta 365(M_{PM}\text{-}M_B) - H\beta 365(M_{PM^*}\text{-}M_B) = H\beta 365(M_{PM}\text{-}M_{PM^*})$. Again within the limits of our epidemiological assumptions, **risk reduction is thus a function of only the decrement in annual-average concentration, independent of how improvements are distributed over individual days. The significance of background is only that it limits the potential for reduction: $H\beta 365(M_{PM}\text{-}M_{PM^*}) \leq H\beta 365(M_{PM}\text{-}M_B)$.** In particular, the "moderate" sensitivity to background of the risk reductions estimated in the draft RA are artifacts of an inappropriate calculation framework.

The (non-)dependence of risk reduction estimates on assumed background.

WHW, 11/10/03

$C_i = B_i + A_i$: ambient concentration on i^{th} day, which is the sum of
B_i: policy-relevant background, and
A_i: controllable anthropogenic.

$M_X = \text{mean}(X_i \mid i = 1,\ldots,365)$, where $X_i = A_i$, B_i, or C_i.
S = annual NAAQS.

Required linear roll-back fraction, p:

$\mathbf{S} = M_B + (1-p)M_A$, so $p = (M_C - \mathbf{S})/M_A$.

Rolled-back concentrations, C^*_i: $C^*_i = B_i + (1-p)A_i$
Concentration reduction: $C_i - C^*_i = pA_i$

Risk reduction: $\mathbf{H}[\exp(\beta pA_i) - 1]$, where **H** is the baseline incidence of the health effect.
Can we employ the linear approximation, $\exp(\beta pA_i) - 1 \sim \beta pA_i$?

i. Consider base-case $PM_{2.5}$ short-term mortality in Detroit, for which $\beta = 0.00074$ (DRA exhibit C.2), $M_C = 15.8$ (DRA exhibit A.2), and max $C_i = 86$ (DRA exhibit A.2). For the current standard **S** = 15 and the base-case average background M_B = 3.5, the required roll-back is p = (15.8-15.0)/(15.8-3.5) = 0.065. The greatest deviation from linearity will occur at the maximum controllable anthropogenic concentration $A_i = 86 - 3.5 = 82.5$. At this extreme daily concentration, $\exp(\beta pA_i) - 1 = 0.003979$, which is 0.2% higher than the linear approximation.
ii. Consider a bounding case, combining the lowest contemplated standard, **S** = 12, the highest background estimate $M_B = 5$, and the upper bound effect estimate, β = 0.0022. Even under these conditions, $\exp(\beta pA_i) - 1 = 0.06516$ is only 3.2% higher than the linear approximation. That's a 3% error on the worst individual day! What else in the entire exercise is known to within 3%?

Conclusion: we can legitimately estimate the risk reduction as $\mathbf{H}\beta pA_i$.

The annual reduction in risk is then just $\mathbf{H}\beta p\Sigma A_i = 365\mathbf{H}\beta pM_A$; substituting $p = (M_C - \mathbf{S})/M_A$ yields $365\mathbf{H}\beta pM_A = 365\mathbf{H}\beta(M_C - \mathbf{S})$, which is wholly independent of any assumptions about the level of the background or its variability from day to day.

The only assumption made above about the background was that $C_i = B_i + A_i$, so that $C_i \geq B_i$ on each individual day. If a constant value $B_i = B$ is used for background, and if observed concentrations C_i sometimes dip below this level, <u>then</u> the risk reduction will depend on the value B assumed. But $C_i < B_i$ is no more physical than $C_i < 0$, a situation that the assessment explicitly rules out by an *ad hoc* computational intervention (RA pages 14 and 15).

The contribution of sulfate aerosol to IMPROVE $PM_{2.5}$ levels in the EUS.

WHW, 11/10/03

Public commenters note that $PM_{2.5}$ annual averages and 98th percentiles at a number of eastern IMPROVE monitors “encroach” on the ranges of annual and 24h standards recommended in the Draft Staff Paper. These comments fail to recognize the large contribution of sulfates to the haze in this region, very little of which can plausibly be attributed to natural or extra-continental sources.

Table 1 summarizes annual average $PM_{2.5}$ mass ($\mu g/m^3$) at a number of the cited monitors during 1988-2002, along with the portion $nsPM_{2.5} = PM_{2.5} - (132/96)[SO_4^{=}]$ of that mass not accounted for by ammonium sulfate, as described at http://vista.cira.colostate.edu/improve/. (Because my $PM_{2.5}$ averages exclude observations with invalid sulfate and sulfur data, some differ from the commenters' values by a few tenths of $\mu g/m^3$.)

Table 1. (annual average)	$PM_{2.5}$	$nsPM_{2.5}$
Acadia, ME	6.5	3.3
Lye Brook, VT	6.8	3.4
Dolly Sods, WV	11.7	5.0
Shenandoah, VA	11.3	4.7
Great Smoky Mtns., TN	12.5	5.8
Sipsy, AL	13.7	6.6

Table 2 provides analogous information for the 98th percentiles. Because the ratio of sulfate to other material varies from observation to observation, the sulfate content of the individual observation supplying the 98th-percentile mass concentration need not be representative. I therefore calculated 98th-percentile $nsPM_{2.5}$ by scaling the 98th-percentile $PM_{2.5}$ by the ratio of non-sulfate and total-mass averages for those observations yielding the top 2% of $PM_{2.5}$.

Table 2. (98th percentile)	$PM_{2.5}$	$nsPM_{2.5}$	$nsPM_{2.5}/PM_{2.5}$
Acadia, ME	22.5	10.4	46%
Lye Brook, VT	26.9	10.8	40%
Dolly Sods, WV	37.6	12.4	33%
Shenandoah, VA	33.1	11.8	36%
Great Smoky Mtns., TN	34.6	16.3	47%
Sipsy, AL	31.9	13.4	42%

I submit that non-sulfate fine mass, $nsPM_{2.5}$, is a much more informative upper bound on policy-relevant background than total $PM_{2.5}$ is. As evidence that EUS sulfate is NOT properly considered PRB, note that regional sulfate concentrations are in fact tracking trends in regional SO_2 emissions! [W.C. Malm, B.A. Schichtel, R.B. Ames, and K.A. Gebhart (2002) A 10-year spatial and temporal trend of sulfate across the United States. *J. Geophys. Res. 107(D22)*, 4627.]

Risk Analysis Appendix B. Linear Trends in Historical $PM_{2.5}$ Data

Warren H. White, 11/17/03

Pages B-1 – B-5 make the necessary point adequately. The statistical discussion on pages B-6 and B-7 adds nothing to the argument. It serves only to impress the statistically credulous, and should be dropped. The fact is that almost ANY two reasonably-shaped distributions containing some small values will generate well-correlated decile averages. How, then, does any particular large R^2 with small intercept "support the hypothesis underlying the proportional rollback method"?

As one example of my claim, regress 1995 Los Angeles deciles on 1992/93 Philadelphia deciles in place of the same-city comparisons. The between-city 'fit' is even better, $R^2 = 0.993$, with a similarly insignificant intercept of -0.15 μg/m^3.

As another example, pull a data set off the web at random – let's consider the numbers of books and serial volumes in each state's public libraries during fiscal 2000 (National Center for Education Statistics, nces.ed.gov/pubs2003/digest02/tables). Calculating the averages in each 5-state decile and regressing them against 1995 Los Angeles concentration deciles again yields $R^2 = 0.993$.

Was that just a lucky shot? Consider instead the numbers of deaths in each state during 1998 among sentenced male prisoners under State or Federal jurisdiction (U.S. Bureau of Justice Statistics, www.ojp.usdoj.gov/bjs/corrections). The decile averages for those data yield $R^2 = 0.973$ with the 1995 Los Angeles $PM_{2.5}$ concentrations. And they correlate at $R^2 = 0.989$ with the library books!

Concluding comments

My core point has been that the effective linearity of the C-R function $H[\exp(\beta A_i) - 1] \sim H\beta A_i$ renders moot much of the argumentation presented in the RA, SP, and public comments (along with some of my own responses such as the above comment on RA Appendix B). I recognize that there may be specific health outcomes and sensitive subpopulations for which the coefficient β is so large that nonlinearity can no longer be neglected. I further accept that it may be prudent for EPA to allow for such a possibility even before the evidence is in. But I strongly encourage the Agency to acknowledge as well, in the interests of lucidity and transparency, the implications of linearity for risk assessment and standard setting.

Dr. George T. Wolff

Comments on the August, 2003 PM Staff Paper

George T. Wolff
11-19-2003

1. Figures 2-6 and 2-7 and 2-9 to 2-12 are impossible to read. You cannot distinguish between the two lowest concentration areas on the maps. I recommend that some sort of a hatched grid be used for the 2nd lowest instead of a shade of gray.

2. p. 3-1, lines 21 – 22 – Change to "Of special importance from the last review were *EPA's* conclusions... I make this distinction because some of the last CASAC PM Panel members did not endorse the third conclusion.

3. p. 3-9, lines 18 – 20 – This is complete speculation and should be deleted.

4. p. 3-15, lines 3 – 8 – This is an overstatement. There are many exceptions. The results are very heterogeneous with respect to strength of an association, whether or not there is a $PM_{2.5}$ association at all, the health endpoint and the pollutant associated the strongest.

5. p. 3 – 17, lines 6 – 10. The basis for this statement is Figure 3-10, which is Figure 5 in the NMMAPs reanalysis. There are two flaws in basing that statement on this figure. The first is that each curve in the figure is based on a different group of cities. They must be based on the same group of cities before any conclusions can be draw. Second, similar graphs do not exist for the gases. An examination of Figures 12 and 14 – 16 in the NMMAPs reanalysis report suggests that for lag 1 (lag 0 for ozone), the graphs for all the gases would look similar to your Figure 3-10. Furthermore, a close examination of these figures indicates that the % changes in mortality/concentration for the single pollutant models are all statistically significant and rival or exceed the PM effect.

6. Figure 3-4 - To look at the impact of possible publication bias, the full 90 individual city results for NMMAPs for total mortality and for cardio and respiratory mortality should be plotted and compared to the individual city PM data in Figure 3-4.

7. p. 3-33, lines 2-4 – This is an understatement. Effects were only seen in persons without a high school education. Also, why does this have to be due to an unidentified socioeconomic effect modifier? It does not modify the effect – it eliminates it. Why can't it be due to an unidentified confounder?

8. p. 3-33, lines 4 and 5 – Doesn't this implausible finding regarding SO_2 cast suspicion on the credibility of the entire study?

9. p. 3-39, lines 6 – 21 – No matter how you stack them there are only 4 long-term cohort studies (ACS, 6cities, AHSMOG, and VA). The first two give positive results and the second two give negative results. However, in both the ACS and 6cities when those with more than a high school education are considered, none of the results are statistically significant suggesting that the studies missed an important confounder. Taking the more than high school education cohorts from each study and the AHSMOG and VA results, the weight of evidence is that there is no long-term mortality effect. Consequently the statement on lines 8 and 9, "lack of consistent findings in the AHSMOG study and the negative results of the VA study, do not negate the finding of the Six Cities and ACS studies," needs to be reconsidered.

10. p. 3-41, lines 2-6 – This mischaracterizes Lipfert's conclusions. He refuted Woodruff's findings by showing that her study was confounded by geographical patterns in infant mortality. Also on line 6, two Chay and Greenstone references are cited, but only one is listed in the back and it is not a peer-reviewed publication.

11. p. 3-62, line 1 to 3-65, line 3 – In comments submitted to EPA at the August 2003 CASAC meeting on the PM CD, Dr. Venditti dismissed much of the material presented here. Since I assume his comments will be addressed in the next CD, this section will need to be revised to reflect Venditti's comments.

12. p. 3-66, lines 4 – 16 – This is a flagrant example of cherry picking. Two studies are cited showing a positive relationship but both used the flawed GAM analysis and were not reanalyzed. A third flawed GAM study reference was also used to support the first two even though when it was reanalyzed the effect became non-significant.

13. p. 3-69, lines 1 – 5 – This conclusion from the 1996 CD is no longer valid given the HEI GAM reanalysis commentary.

14. p. 3-70, lines 23 - 30 – See comment 5 above.

15. p. 3-70, general comment – There is a general problem with the GAM re-analysis. They tended to re-analyze only the PM associations from the original papers and did not address other important aspects of the papers. For example, the original analysis of the 8 Canadian cities by Burnett et al.2000 allocated the mortality risk among different pollutants. However, the reanalysis by Burnett and Goldberg 2003 did not re-analyze these aspects of the paper. The result is that we do not know whether the conclusions of the original paper regarding the role of gases vs. PM are still valid. This is important because the SP includes the new PM results as a multi-city PM study. Without the further reanalysis, it is not clear what credence to give to the single pollutant PM results in the 8 Canadian cities.

16. p. 3-71, line 16 – It is amazing how biological plausibility of the gases can be dismissed but not for PM.

17. p. 3-74, lines 3 – 8 – This is complete speculation. There is not a shred of evidence to support this statement.

18. p. 3 – 74, lines 12 – 13 – See comment 5 above.

19. p. 3 – 75, lines 1 – 4 – See comment 12 above.

20. p. 3 – 76, lines 22 – 24 – This is a serious development and until it is resolved it should be a showstopper for any consideration being given to lowering a $PM_{2.5}$ NAAQS.

21. p. 3 – 77, lines17 – 18 – I cannot find this on p. 8-84 in the CD.

22. p. 3 – 77, lines 21 – 23 – This is too strong a statement given the uncertainties associated with correcting for weather and model selection.

23. p. 3-77, lines 27 –29 – It is unfortunate that EPA did not use the CD to critically evaluate the methods employed, but just blindly accepts them.

24. p. 3 – 82, lines 21 – 23 – HEI reports 96 and 98 do not support this statement. And is contradicted by the statement on p. 3-83, lines 7 – 9.

25. p. 3 – 89, section 3.5.4.1 – How can such heterogeneous results be consistent?

26. p. 3 – 98, section 3.5.4.2 – The long-term studies show no significant response for respiratory disease. This is not coherent with the short-term studies.

27. p. 4-5, lines 15 – 18 – This will need to be revised based on the new CD.

28. p. 4-5, line 33 – Since causality is an assumption, the lower bound on all of the risk estimates has to be zero.

29. Figures 4 –3 to 4 –14 – These figures are misleading. Truncating the lower bound at zero and using the upper bound at the 97.5th percentile creates an artificially high positive picture of effects and obscures the heterogeneity of the data. The 95th should be used throughout and negatives should be shown.

30. Figures 4-4, 4-10 and 4-11 - There are numerous uses of the Klemm et al. 2000 results as re-analyzed by Klemm and Mason 2003 in the Figures in Chapter 4. Since Klemm and Mason showed that the results were sensitive to the degree of smoothing and there is no a priori reason to favor GAM results over GLM results, the range of results from Klemm and Mason 2003 should be shown in the Figures for each PM metric and endpoint shown.

31. Figures 4-10 and 4-11 – When CASAC asked EPA to include the NMMAPS estimates in the risk assessment, we wanted the individual city estimates to be used not the estimated

or pooled estimates. Using the estimated or pooled estimates totally distorts the picture and obscures the heterogeneity.

32. Same Figures - Since the Schwartz 10 city study and NMMAPS both have reported associations for the same cities it would be useful to plot the Schwartz results versus the NMMAPS results to look at the effect of model selection in this subset of NMMAPS cities.

33. Figure 4-11 - To the extent that the regional NMMAPS results are used for comparison in the RA, the variation in regional dose–response in Dominici et al. 2003 should also be discussed. The shapes of the dose-response curves are substantially different among the regions. In addition, given the confidence limits, the shapes in several regions are consistent with a threshold model.

34. p. 4-58 - The SP indicates that sensitivity analyses were carried out for each of the study areas but that the results for Detroit were included in the SP for illustrative purposes. The text indicates that Detroit was selected because it provides an opportunity to examine both mortality and morbidity risk and includes both single and multi-pollutant C-R functions. However, the PM2.5 mortality risk in Detroit as reported in Lippmann et al. 2000 and Ito 2003 is suspect. The strongest positive association for total mortality occurred on lag 3. In Lippmann et al. 2000, the relative risk for 5^{th} to 95^{th} percentile pollutant increment for total mortality was 1.0448 with a t statistic of 1.62, a positive but not statistically significant association. However, the coefficient for circulatory mortality on lag 3 was 1.0042 with a t = 0.1 and the coefficient for respiratory mortality was 1.0005 with t = 0.01. In contrast, the coefficient for the "other' category was 1.0924 with t = 2.28, a relatively strong and statistically significant association. Thus, the positive association on lag 3 was caused by a positive association with "other" mortality and there was little or no association with circulatory or respiratory mortality. When the overall pattern for all 4 days (lags 0, 1, 2, and 3) is considered, the lag 3 association with other mortality was the only statistically significant association of the 12 associations evaluated. In addition, there was evidence of an association with other mortality on three of the four days, no apparent association with respiratory mortality (one positive association and one negative association) and little evidence of an association with circulatory mortality (one positive association). In the re-analysis, all the coefficients were re-calculated but only selected results were presented. This makes it difficult to fully interpret the results. The lag 3 total mortality coefficient was reduced by 40 % with the stringent GAM convergence and 36 % in a GLM model, but it remained the largest daily coefficient although still not statistically significant. Since the results for all the lag/effect category combinations were not presented in Ito 2003, the overall pattern in the re-analysis is not known. However, since the scatter plots in Figures 1 and 2 of Ito 2003 show a wide range of both positive and negative associations in the database, the most likely situation is that the lag 3 total mortality association is the result of random noise in the data and is not a true health effect.

35. Tables 4-10 and 4-11 – The Veterans study needs to be included in this table.

36. General Comment on Chapters 4 and 6 – I do not see the connection between chapters 4 and 6. I thought that the risk assessment was going to be the basis for the selection of the NAAQS. There does not appear to be a connection. What is the purpose of the risk assessment?

37. p. 6-5, lines 6 – 10 – What about the Veterans study?

38. p. 6-6, lines 21 – 26 – This is so important. EPA acknowledges the new awareness of the uncertainties that previously were thought to have been put to rest but have re-emerged, but does not think it through. This is a showstopper and should preclude any efforts to lower the standards.

39. p. 6-10, lines 16 – 17 – There are an equal number of studies (including toxicological) that indicate no effect from sulfates.

40. p. 6-10, line 24 – In the $PM_{2.5}$ discussions, the point that crustal particles are safe is made many times. However, this is in conflict with the discussions of $PM_{10-2.5}$, which are mainly crustal material. How can fine crustal material be ok, but coarse need to be regulated?

41. p. 6-11, lines 1 –2 – Here EPA says there is no basis to conclude that any individual fine particle component cannot be associated with adverse health effects but on the previous page said that there is no association with crustal particles.

42. p. 6-19, line 4 – "precautionary" – The precautionary principle has no place in the standard setting process

43. p. 6-19 – lines 1 – 4 – In the 1996 Staff Paper, EPA proposed a range from 12.5 $\mu g/m^3$ to 20 $\mu g/m^3$, and the Administrator eventually picked 15 $\mu g/m^3$. However, in the final CASAC discussions, the focus of the Panel was in a range of 15 to 30 $\mu g/m^3$, and this debate was never resolved (see Table I). Consequently, the present debate should begin where the last debate left off and consider a range of 12 to 30 $\mu g/m^3$. The overview of the debate is evident in the attached Table where only 2 of 21 Panel members selected a range that went down as low as 15 $\mu g/m^3$.

44. p. 6-23, line 18 to p. 6-25, line 10 – I think there is some merit to the methodology used to select the 24-hr range, but it needs to be expanded upward to accommodate my recommended annual range.

45. section 6.4.3 – There is no long-term study that demonstrates an effect at current $PM_{10-2.5}$ so I support EPA's consideration of not having an annual $PM_{10-2.5}$ NAAQS.

46. section 6.4.4 – While I think the method of selection of the 24-hr range for $PM_{2.5}$ has some merit, it is not appropriate for the 24-hr $PM_{10-2.5}$ range particularly since there is inadequate basis for the selection of an annual range.

47. General Comment - The presentation by Harvey Richmond on the Risk Assessment argued that the use of PM mass was appropriate "given the absence of sufficient information to address either differential toxicity of PM components or differential changes in PM components upon meeting standards." As Harvey indicated, this is indeed a key assumption. However, I am sure that the vast majority of scientists in this field believe (and the Draft CD acknowledges) that there is differential toxicity among PM components. In addition, it is extremely unlikely that any implementation program will control all man-made PM components alike. The implementation will target specific sources (national, regional, and local) and control each to varying degrees based on the availability of technology and cost. And background components will not be controlled. Thus, there are sure to be differential changes in PM components as standards are met. Although there is not sufficient information to fully address this issue, there is enough known to start to address it and do some sensitivity analyses. For example, the major fine and coarse components could be assigned different weights based on available government reference toxicity levels. For implementation, several options could be considered, ranging from focusing on the components of greatest mass, to focusing on the components of greatest toxicity. The point of such an exercise is to demonstrate that differential toxicity and control matter if we want to protect public health.

Summary of CASAC Panel Members Recommendations
(all units µg/m^3)

NAME	Discipline	$PM_{2.5}$	$PM_{2.5}$	PM_{10}	PM_{10}
		24-hr	Annual	24-hr	Annual
EPA		18 -65	12.5 - 20	150[13]	40 - 50
Ayres	M.D.	yes[2]	yes[2]	150	50
Hopke	Atmos. Sci.	20 - 50[3]	20 - 30	no	40 -50[4]
Jacobson	Plant Biologist	yes[2]	yes[2]	150	50
Koutrakis	Atmos. Sci.	yes[2,5,6,12]	yes[2,5,6]	no	yes[4]
Larntz	Statistician	no	25-30[7]	no	yes[2]
Legge	Plant Biologist	≥ 75	no	150	40 - 50
Lippmann	Health Expert	20 - 50[3]	15 - 20	no	40 - 50
Mauderly	Toxicologist	50	20	150	50
McClellan	Toxicologist	no[8]	no[8]	150	50
Menzel	Toxicologist	no	no	150	50
Middleton	Atmos. Sci.	yes[2,3,12]	yes[2,5]	150[3,13]	50
Pierson	Atmos. Sci.	yes[2,9]	yes[2,9]	yes[4]	yes[4]
Price	Atmos. Sci./ State Official	yes[3,10]	yes[10]	no[3,4]	yes[4]
Shy	Epidemiologist	20 - 30	15 - 20	no	50
Samet[1]	Epidemiologist	yes[2,11]	no	150	yes[2]
Seigneur	Atmos. Sci.	yes[3,5]	no	150[13]	50
Speizer[1]	Epidemiologist	20 - 50	no	no	40 - 50
Stolwijk	Epidemiologist	75[7]	25-30[7]	150	50
Utell	M.D.	≥65	no	150	50
White	Atmos. Sci.	no	20	150	50
Wolff	Atmos. Sci.	≥75[3,7]	no	150[3]	50

[1] not present at meeting; recommendations based on written comments
[2] declined to select a value or range
[3] recommends a more robust 24-hr. form
[4] prefers a $PM_{10-2.5}$ standard rather than a PM_{10} standard
[5] concerned upper range is too low based on national $PM_{2.5}/PM_{10}$ ratio
[6] leans towards high end of Staff recommended range
[7] desires equivalent stringency as present PM_{10} standards
[8] if EPA decides a $PM_{2.5}$ NAAQS is required, the 24-hr. and annual standards should be 75 and 25 µg/m^3, respectively with a robust form
[9] yes, but decision not based on epidemiological studies
[10] low end of EPA's proposed range is inappropriate; desires levels selected to include areas for which there is broad public and technical agreement that they have $PM_{2.5}$ pollution problems
[11] only if EPA has confidence that reducing $PM_{2.5}$ will indeed reduce the components of particles responsible for their adverse effects
[12] concerned lower end of range is too close to background
[13] the annual standard may be sufficient; 24-hour level recommended if 24-hour NAAQS is retained

NOTICE

This report has been written as part of the activities of EPA's Clean Air Scientific Advisory Committee (CASAC), a Federal advisory committee administratively located under the EPA Science Advisory Board Staff that is chartered to provide extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. 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.