UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

May 21, 2009

EPA-CASAC-09-009

The Honorable Lisa P. Jackson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Ave., NW
Washington, D.C. 20460

Subject: Consultation on EPA's *Particulate Matter National Ambient Air Quality Standards: Scope and Methods Plan for Health Risk and Exposure Assessment*

Dear Administrator Jackson:

The Clean Air Scientific Advisory Committee (CASAC) and members of the CASAC Particulate Matter Review Panel met on April 2, 2009 to conduct a consultation on EPA's *Particulate Matter National Ambient Air Quality Standards: Scope and Methods Plan for Health Risk and Exposure Assessment.* Typically, CASAC does not provide written comments on planning documents; however for this review, we thought it important enough to provide some overall comments on the plan for health risk and exposure assessment.

Overall, we found the draft plan to be ambiguous in several instances, offering only a number of possibilities that EPA might follow as it develops the ISA. The CASAC's comments might have been sharper if the draft plan had been more closely linked to the completed first draft of the ISA that was reviewed at this same meeting.

The *Scope and Methods Plan for Health Risk and Exposure Assessment* is ambitious, calling for analyses related to several health outcomes and a national-scale estimate for mortality. CASAC suggests that priorities be established quickly in developing the health risk and exposure assessment, giving emphasis to those analyses that may be most informative for establishing particulate matter standards. The Plan should provide a transparent algorithm for selecting endpoints based on the level of certainty and the relative and attributable risks. CASAC suggests that weight be given to the level of classification while still considering the Administrator's obligation to set a standard with a "margin of safety" as described in the Clean Air Act. For example, several CASAC members do not recommend a risk assessment based on birth

outcomes, in part because the level of evidence is still at the *suggestive* level. One panel member proposed setting a higher priority for those health effects shown to have the highest risks in the epidemiological literature. There was support for doing a limited risk assessment for short-term exposure to $PM_{10-2.5}$ for appropriate outcomes such as hospitalization.

With respect to the use of air quality data, the draft plan states that 2005-2007 data have been filtered by application of the Exceptional Events Rule for 24-hour NAAQS designations. The Exceptional Events Rule allows states to exclude air quality data that violate a national air quality standard if an "exceptional event" has caused the violation, e.g. wildfires. It might be informative if the Agency would provide some examples of how the Exceptional Events screening affects the resulting data and metrics like the 98^{th} or 99^{th} percentiles. For the coarse mass data, EPA will be challenged to separate anthropogenic from natural sources, however CASAC generally supports EPA's proposed approach for estimating policy relevant background levels. In general, there was support for the proportional roll-back approach (applying the same percentage adjustment for all concentrations exceeding the policy relevant background to characterize scenarios that just meet specified PM standards) that EPA had followed previously.

With regard to the risk assessment component, CASAC was generally in agreement with the planned approach to identifying concentration-response relationships and using those coming from distributed lag models. We support the Agency's plan to conduct a national scale health impact assessment for long-term exposure mortality related to $PM_{2.5}$. In fact, CASAC believes such a national assessment should play a central role in the overall risk assessment. We also support the general approach to uncertainty analyses, but with a warning to carefully separate sensitivity analyses from uncertainty analyses. With regard to the approach for classifying the degree of uncertainty, we suggest exploring the use of various structured approaches for describing uncertainty. (Recent examples may be found in the work of the World Health Organization and the Intergovernmental Panel on Climate Change.)

CASAC welcomes the inclusion for the first time of a quantitative exposure assessment to focus on 24-hour population exposures to $PM_{2.5}$. The Agency's exposure assessment plan rightly seeks to identify various personal and building-related factors that may account for some of the variability in $PM_{2.5}$-associated health risks. More information is needed on how the results from the exposure assessment will be integrated and used to interpret epidemiological studies. We recognize the Agency's time and resources are not unlimited, but it would be useful to evaluate the APEX model and assess its performance for each of the ten cities for which it will be used.

Since our comments on the Health Risk and Exposure Assessment are very general and brief, we do not expect any formal response from the Agency. Individual committee members' comments are appended here. We thank the Agency for the opportunity to provide advice early in the PM NAAQS review process, and look forward to the review of the First Draft Health Risk and Exposure Assessment in October 2009.

Sincerely,

/Signed/

Dr. Jonathan M. Samet, Chair
Clean Air Scientific Advisory Committee

Enclosures: Enclosure A: CASAC Particulate Matter Review Panel Roster
Enclosure B: Compendium of Individual Comments

NOTICE

This report has been written as part of the activities of the EPA's Clean Air Scientific Advisory Committee (CASAC), a federal advisory committee independently chartered to provide extramural scientific information and advice to the Administrator and other officials of the EPA. CASAC provides balanced, expert assessment of scientific matters related to issues 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 within the Executive Branch of the federal government. In addition, any mention of trade names or commercial products does not constitute a recommendation for use. CASAC reports are posted on the EPA Web site at: http://www.epa.gov/casac.

Enclosure A

Clean Air Scientific Advisory Committee
Particulate Matter Review Panel

CHAIR

Dr. Jonathan M. Samet, Professor and Chair, Department of Preventive Medicine, University of Southern California, Los Angeles, CA

CASAC MEMBERS

Dr. Joseph Brain, Philip Drinker Professor of Environmental Physiology, Department of Environmental Health, Harvard School of Public Health, Harvard University, Boston, MA

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

Dr. James Crapo, Professor of Medicine, Department of Medicine , National Jewish Medical and Research Center, Denver, CO

Dr. H. Christopher Frey, Professor, Department of Civil, Construction and Environmental Engineering, College of Engineering, North Carolina State University, Raleigh, NC

Dr. Donna Kenski, Data Analysis Director, Lake Michigan Air Directors Consortium, Rosemont, IL

Dr. Armistead (Ted) Russell, Professor, Department of Civil and Environmental Engineering , Georgia Institute of Technology, Atlanta, GA

CONSULTANTS

Dr. Lowell Ashbaugh, Associate Research Ecologist, Crocker Nuclear Lab, University of California, Davis, Davis, CA

Prof. Ed Avol, Professor, Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA

Dr. Wayne Cascio, Professor, Medicine, Cardiology, Brody School of Medicine at East Carolina University, Greenville, NC

Dr. Douglas Crawford-Brown, Professor Emeritus, Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC

Dr. David Grantz, Director, Botany and Plant Sciences and Air Pollution Research Center, Riverside Campus and Kearney Agricultural Center, University of California, Parlier, CA

Dr. Joseph Helble, Dean and Professor, Thayer School of Engineering, Dartmouth College, Hanover, NH

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

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

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

Dr. Helen Suh MacIntosh, Associate Professor, Environmental Health, School of Public Health, Harvard University, Boston, MA

Dr. William Malm, Research Physicist, National Park Service Air Resources Division, Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO

Mr. Charles Thomas (Tom) Moore, Jr., Air Quality Program Manager, Western Governors' Association, Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO

Dr. Robert F. Phalen, Professor, Department of Community & Environmental Medicine; Director, Air Pollution Health Effects Laboratory; Professor of Occupational & Environmental Health, Center for Occupation & Environment Health, College of Medicine, University of California Irvine, Irvine, CA

Dr. Kent Pinkerton, Professor, Regents of the University of California, Center for Health and the Environment, University of California, Davis, CA

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. Sverre Vedal, Professor, Department of Environmental and Occupational Health Sciences, School of Public Health and Community Medicine, University of Washington, Seattle, WA

SCIENCE ADVISORY BOARD STAFF

Dr. Holly Stallworth, Designated Federal Officer, EPA Science Advisory Board Staff Office, Washington, DC

Enclosure B

Compendium of Comments

CASAC Particulate Matter Review Panel on

PM NAAQS: Scope and Methods

Plan for Health Risk and Exposure Assessment (Feb. 2009)

Comments from Mr. Ed Avol

General Comments related to Charge Questions for Chapter 4:

1) On overall Structure and Design of the exposure assessment:

It is still not exactly clear what precisely is going to be done. The general layout looks okay, but where are the details? How will this sCope fo Work actually be acomplished? (see several specific comments below)

2) Planned measures of exposure – why only "less than 24hrs $PM_{2.5}$"? Can something be done for longer time frames? If this is truly the only choice, both for pollutant (as compared to coarse, ultra-fine, etc) and averaging time (seasonal? annual?) then (reluctantly) I guess that is the way it is, but I did not find a explicit rationale provided that was convincing.

3) Selection of subset of selected urban study areas – were other options available using another approach? Why were these chosen? They do represent a range across the country, but don't represent the breadth of ambient conditions; rather, they represent where the health data are, or where the monitoring information is available…Have you made the proper selections, and how do you support those choices?

4) Uncertainty and Variability – The plans seems reasonable, so long as the measurements exist to "ground truth" it; otherwise, I defer to statisticians to identify the best method for assessing the errors propagated in this approach.

****

Specific Comments:

P1-10, lines 5-7 – The decision to focus on specific selected health endpoints is a necessary one, but I am not convinced the correct one has been made. The exclusion of other potentially important outcomes (birth effects, for example) seems worthy of examination and discussion.

P1-11, lines 9-13 – If my understanding of the logic behind the REA is correct (that it strictly flows from the discussions and determinations made in the ISA), then this paragraph touches on a previously addressed topic (in the ISA review) and is out of place. However, (since it is raised here), the large and ever-growing amount of information regarding fresh traffic exhaust and exposures (as well as assorted health outcomes associated with living or spending extended amounts of time in close proximity to busy traffic venues) suggests this is a specific environment and source for closer examination. Similarly, the emerging work in ultrafine particles exposures and effects argue that health concerns, exposure reviews, and document evaluations based on particle-size ought to be extended down into the ultra-fine range.

P1-12, lines 1-2 – If the Staff sense is that there is insufficient spatial coverage of coarse particle monitoring data, is something being recommended or undertaken to deal with this perceived

shortcoming? Can't something be done? Would it be useful to do so? Some clarification here would be helpful to the reader…

P2-2, lines 12-13 – I am skeptical that this standard of only using monitoring or health information obtained by FRM or FEM methods is actually applied in the epidemiologic or even clinical arenas.

P2-3, lines 9-13 – Regarding the monitoring data mis-matches for PM2.5 and PM10 for determining PM10-2.5, is there some estimate or idea about the level of uncertainty associated with these mis-matches? Will this be quantified in some manner in the uncertainty and variability analyses?

P3-3, line7-11 – Doesn't this issue of using PM10 as the metric for coarse particles, but being concerned about using the health effects data arising from such use, create an internal inconsistency?

P3-3, lines 12-16 – If the logic is that the REA is to be all mass-based, that ignores what we know about particle composition. If a decision is made to include compositional information, then the decision to include coarse but exclude ultra-fine seems questionable. If existing data was judged to be sufficient to move forward on coarse particles (but use PM10 as the metric and ignore the different chemical composition of coarse particles compared to fines), why vacillate on using data where PM10 was the metric? There would seem to be a significant amount of traffic proximity/fresh combustion exhaust health and exposure research available for review – how much will it take to make this sufficient in the same context as coarse particles?

P3-22, lines 1-6 – The desire to fully differentiate variability and uncertainty (or variability from uncertainty) is a noble one, but practically speaking, what will this accomplish with respect to the REA? If we knew enough to separate uncertainty from variability, we could use this to identify current gaps and needed areas of research; is that part of this document? Do we know enough to assign errors to one or the other of these sub-categories, and does it matter as much as this argument seems to believe?

P4-1, line 22 – If no exposure assessment is planned due to the perceived lack of sufficient spatial coverage for coarse particles, what is planned for coarse PM in the evaluation?

P4-2, line 8 – Why only consider 24hr population exposures? One could be in compliance with 24hr exposure levels and exceed the annual standard. If the limitation is linking it to the selected health studies, perhaps a re-examination of the pool of available health data and health outcomes is needed.

P4-8, lines 3-4 – Shouldn't there be a brief justification (or reference to a justification) for selection of these microenvironments to be modeled?

P4-8, Table 4-1 – Would it help to have some qualitative assignment of likely error for the method to be used for each microenvironmental assignment?

P4-9, line 6 – Presumably, the "shorter" averaging time being considered includes hourly or 8-hr or something,,,but why not consider longer time frames (such as seasonal or annual)?

P4-10, line 11 – is there some significance being attached to employment status here (or is this an assignment of population bins for modeling)?

P4-14, lines 12 and 23 – The decision not to consider the personal cloud seems problematic…not sure in which direction this will drive potential bias, however.

P4-19, lines 1-4 – The assertion that a stochastic model will be developed for a small number of urban areas for near-road concentrations is laudable, but potentially challenging. What is "Plan B" if this doesn't work? Meteorology will be an important issue here, but how available and applicable is the data for the areas to be evaluated?

P4-19, lines 24-28 – What are "size classes of $PM_{2.5}$"? How can there be size classes of a given size? A little explanation is needed here.

P4-20, lines 2-3 – Near-road and in-vehicle activities can disproportionately affect a person's day of exposure, according to several recent studies. Development of a way to capture this is a critical component of any methodology!

P4-20, Exposure Modeling Issues bulleted items – there is a fifth one that is not explicitly listed but probably should be: in-vehicle time, because of the disproportionate exposures, on-road in-vehicle exposure may be worthy of a closer look.

Comments from Dr. Douglas Crawford-Brown

This review considers specific issues related to Chapter 3: Scope and Approach for the Health Risk Assessment.

1. The risk assessment makes use of an appropriate selection of information from the draft ISA, both in regards to effects examined – morbidity and mortality - and exposure pathways that dominate. The decision to focus primarily on the 2.5 fraction is consistent with the ISA as well.

2. It was good to see that the assessment will consider how to place the limited sample size of cities into a broader judgment of national risks. I will need to reserve judgment as to how well that is done, as the draft document provides only a hint as to the general approach, and the details will be significant. But this is an issue for which I have had concern in the past, and so it is heartening to see that it is being addressed directly.

3. I found pages 3-1 and 3-3 to be somewhat inconsistent in regards to the composition of particles. The authors seem to first suggest that composition will be considered, and then on 3-3 state that the risk assessment will use only the mass fraction. I suppose this could be counted as being "considered", in the sense that such a composition approach is rejected, but the discussion as to why it is rejected is so cursory that the two pages end up seeming to be inconsistent.

4. I disagree with the statement on 3-4 that Equation 3-1 does not require more detailed individual-specific exposure data. It certainly is true that 3-1 can be applied to populations even with exposure distributions, but it can equally be applied to studies where detailed individual-specific exposures are used to stratify the population. In fact, it is more accurate in the latter case. So I don't know what the authors mean by this statement.

5. I support the use of the 2005-2007 monitoring data as the basis for the exposure assessment. This is a reasonably complete dataset and relevant to current conditions.

6. The modeling approach outlined on page 3-6 is consistent with REAs conducted for other contaminants, and so is appropriate here.

7. In addition, the proportional roll-back and the modeling approaches have both been used in other REAs, and we have approved these through the CASAC. As a result, I am comfortable with the approach being applied here.

8. The categories of effects noted on pages 3-7 and 3-8 are consistent with those mentioned as prominent in the summaries in the ISA, and so are appropriate ones for this REA. This applies to both the short and long-term exposures. I am less comfortable with the inclusion of birth outcome effects, unless a more compelling case is made for including suggestive effects in ALL REAs as a matter of policy.

9. The criteria for study selection on page 3-9 are appropriate and consistent with past REAs. I don't, however, understand the paragraph beneath this listing of criteria. I have read it several times and can't determine what the authors are trying to say, or why this issue is being raised.

10. The decision on page 3-10 to include both single and multi-pollutant models is a good one, and can be used as a form of sensitivity analysis. Where the two approaches lead to very different results that would have policy implications, however, there will be a need to better specify which approach is preferred. The same comment applies to the other categories of different modeling approaches on page 3-17. Taken as a whole, these various approaches should provide a reasonable exploration of the uncertainty introduced by the availability of alternative approaches.

11. On the issue of cut-points raised on 3-18, the authors should be prepared to offer a scientifically cogent reason for selection of a specific cut-point, and not simply try different cut-points to see what effect this has on the analysis. The draft ISA was clear that there is little evidence for a population threshold in the C-R function.

12. The criteria for city selection on pages 3-18 and 3-19 are appropriate and should lead to a reasonable set of such cities.

13. I am less comfortable with the ways in which the baseline rates are to be determined, or at least how I think they will be determined given what is in the text. It sounds to me as if data on the county containing a city, or even the state, will be used. Given differences in background incidence of many effects in rural and urban areas, I worry if background incidence for an urban area is determined from a geographic area with a large rural population. I will withhold judgment on this until I see what is actually employed.

14. I support the general idea of separating variability and uncertainty in the assessment. The listing of sources of variability on 3-22 and 3-23 is a good one, but I am not sure how the authors plan to separate out these sources, or use the multiple cities in the assessment to get an understanding of this variability. At present, the list strikes me as more aspirational than something that can be built into the methodology.

15. I support the use of a qualitative methodology for uncertainty analysis as a first step. And then I like the idea of a quantitative assessment of uncertainty described in the document. I worry, however, that the fully quantitative approach described may prove infeasible (I hope not, but that possibility remains). If that is the case, I wouldn't want the default position to be the qualitative analysis alone. There have been quite a few studies of uncertainty in PM risk assessment, including subjective probability encoding. I agree that having a high/medium/low categorization of the impact of specific parameters or model choices will be useful, and so support this idea. But it also occurs to me that something more can be said for each of these about the general magnitude of the uncertainty introduced (i.e. factor of 1.5, 2, 10, etc) even if the formal quantitative approach describes does not work in the end. The authors should further

consider whether a more quantitative metric – but not the full sensitivity approach described - can be developed as an adjunct to the qualitative one (not as a replacement). If the full sensitivity approach can be developed, though, this is the approach that will have the most utility.

16. On pages 3-26 and 3-27, I am not sure what is meant by “core” risk assessments. Is this meant to be something like “best estimate”? I think it is. If so, just use that phrase, as “core” is not standard terminology.

17. As I read through subsequent pages, I became less convinced I understood the way in which the authors intend to separate the sensitivity and uncertainty analyses. These are not the same kind of analysis. A sensitivity analysis adjusts each parameter by some fixed fraction and examines the influence on the outcome. An uncertainty analysis employs information on the actual degree of uncertainty in each parameter. But the discussion seems to conflate these two, and so I am left unsure whether an actual uncertainty analysis is being proposed, or only a sensitivity analysis. This must be clarified.

18. I support the approach to presenting the results. It has been effective in other REAs and should continue to serve the purpose here. I liked the mention on page 3-29 of considering the representativeness of a particular geographic area, even if I am not sure how this will be done operationally. But it is certainly an important way of viewing the information.

19. The national approach mentioned is ambitious, and in many ways is another level of complexity to the analyses described in earlier parts of the chapter. As it is so complex, I was struck by the lack of detail provided. I cannot judge the feasibility of the approach given this sketchy description, and so will withhold judgment until the approach is described in detail. It just seems to me a LOT of work to do as an adjunct to the other analyses being conducted, and I don’t yet see the reason for it.

20. I agree both that the PM10-2.5 approach should be similar to that for 2.5 and that the uncertainty will be larger.

Comments from Dr. Chris Frey

Section 1.3.3

Another purpose for doing exposure assessment is to help identify or characterize "vulnerable" subgroups

Chapter 2

Section 2.1. Should explain why the scope excludes PM10 and ultrafines. This is not to imply that they should be included, but just to provide transparency as to why they are not included.

Section 2.2. Recommend insert "epidemiology-based" before "PM risk assessment" in the first line. The typical risk assessment paradigm, as defined by the National Research Council (1983) "redbook" includes an exposure step and a dose-response step. In the "epidemiology-based" approach used here, the exposure step is simplified based on using ambient concentration as a surrogate. Hence, this is a modification of the typical risk assessment paradigm.

Section 2.2.3. The exclusion approach should not be used if the air quality data are being used to conduct or reproduce an epidemiological study or to look for associations between air quality, exposure, and health effects.

Section 2.3.3. The "alternative approach" is a bit unclear. Perhaps a concise step-by-step explanation of the approach would help. For example, would this involve trial-and-error to arrive at emissions reduction scenarios that lead to "just meeting" a regulatory alternative, or would it involve specifying emissions reduction scenarios and accepting arbitrary air quality results?

Section 2.4, page 2-8, bottom of page. Has the use of Bayesian hierarchical modeling been considered in which air quality model results are "updated" with data from those monitors that are believed to represent policy relevant background levels (if any such monitors exist)? Also, as the committee has briefly discussed already, there is likely to be significant relative uncertainty in any estimate of PRB. Has EPA considered attempting to quantify uncertainty in PRB either through assessment of model precisions and accuracy, use of expert elicitation, combinations of these, or other approaches? One could use a mean value of PRB from an estimate of uncertainty in PRB.

Chapter 3

Section 3.3.2. It is not clear as to why it is difficult to incorporate "longer-term" (longer than what?) variability in ambient PM2.5 levels. Such variability can be obtained from monitoring data, air quality model estimates, or combinations of both.

Section 3.3.3. Strongly caution that the use of "high," "medium," and "low" descriptors of uncertainty can be very ***un***informative. See, for example, Granger Morgan's experiment on

interpreting qualitative descriptors of uncertainty using the EPA SAB as an example. See also IPCC efforts to carefully define descriptive terms. See also World Health Organization (2008) recent guidance on uncertainty in exposure assessment, including a structured qualitative approach as well as quantitative approaches. Any terms used qualitatively must be clearly defined and consistently applied.

Last bullet on page 3-24 seems to be referring to lag effects – this could be stated more clearly.

Section 3.3.4. 2nd paragraph. It is not the case that "any existing correlations between those input parameters" must be "clearly defined." Such definitions are needed only when there are correlations between two inputs under the following conditions: the output is sensitive to both inputs, and the strength of the correlation is strong enough to affect the output. Furthermore, if two inputs are correlated, one needs to consider whether a model is properly specified. This is true whether one conducts uncertainty analysis. In general, it is better to explicitly model the dependence, rather than treat dependencies using statistical correlations, wherever possible. Other techniques for dealing with dependence include stratification or "bootstrapping" from paired databases. See Cullen and Frey (1999) for more details.

One must be careful not to "sell" sensitivity analysis as a replacement for uncertainty analysis. Sensitivity analysis merely evaluates the sensitivity of a model response (output) to changes in one or more inputs. Unless one is defining and appropriately dealing with simultaneous variation in ranges of inputs that represent uncertainty, the answer cannot represent uncertainty.

The text mentions "multi-factor" sensitivity analysis but does not explain what specific technique will be used. If the idea is to use nominal range sensitivity analysis but to try to vary two or more inputs, there is no guarantee that the answer will have much relationship to a range of uncertainty that would have been obtained from a probabilistic analysis in which uncertainties in multiple inputs were simultaneously sampled. One suspects that the so-called multi-factor sensitivity analysis method might be merely a repetitive application of a local sensitivity analysis method, which cannot produce results that would be obtained from a global sensitivity analysis method. For more on sensitivity analysis methods, see books by Andrea Saltelli and a series of reports and journal papers by Frey and colleagues such as Mokhtari and Patil.

p. 3-27. There is typically a conceptual disconnect between the notion that one can identify ranges of inputs for sensitivity analysis but somehow cannot develop a distribution of uncertainty for the same input. If the state of knowledge is sufficient to develop a range of values, why is it not sufficient to develop a best estimate of the distribution of such values?

If the goal of a sensitivity analysis is to identify and rank key sources of uncertainty, then it is less controversial to assign distributions to inputs for use in a global sensitivity analysis, which will produce more robust results than a local sensitivity analysis. Hence, why not consider doing a "one-dimensional" Monte Carlo simulation of uncertainties, using in combination with

sensitivity analysis methods such as Pearson or Spearman correlation coefficients, ANOVA, regression, or categorical and regression trees (also known as hierarchical-based tree regression)?

p. 3-28. Top of page, first line. It is not correct to imply that a local sensitivity analysis is an "uncertainty simulation." Uncertainty typically refers to range and relative likelihood, not just an arbitrary sample of point estimates in the modeling domain.

End of section 3.3. the last sentence is dubious. Developing a set of arbitrary point estimates of unknown confidence with regard to coverage of the uncertainty in the model response could lead to arbitrary artifacts – while it could be reasonable to describe each point estimate as an example of a scenario, it would not be appropriate to claim that a set of local sensitivity analyses could be interpreted as a probability sample or as covering a range of uncertainty.

Chapter 4.

The use of APEX for conducting exposure assessment for PM2.5 as described is generally appropriate. EPA has identified several key issues that are works in progress. Given that results are not available or that some of the methodology is not yet established, we will have to wait for more information in order to provide more detailed feedback.

Is there prior knowledge upon which the choice of microenvironments as given in Table 4-1 is based? If EPA might decide to focus risk or exposure analysis on susceptible or "vulnerable" subgroups, might this affect the relative importance among the microenvironments? For example, elderly susceptible groups might spend a larger portion of time at home and less time at "offices" than groups comprised of healthy working adults. In terms of developing the parameters for the mass balance or factor approaches for each microenvironment, is there a sense of priorities among these in terms of data quality needs?

Section 4.5.3. Why not use hierarchical Bayesian methods for combining model and monitoring data?

Section 4.5.5. If EPA has decided not to include indoor sources of PM2.5, does this affect a preference for a "mass balance" versus "factor" approach?

For clarity, it should be explained as to what metric(s) APEX uses for exposure. Does it use time-weighted concentration? Or does it also estimate potential dose based on breathing rates.

Are there any plans to deal with PM components? Or to apportion exposures to emission sources?

The capability for a "2-D" method for quantifying variability and uncertainty is good.

Comments from Dr. David Grantz

1. Include, or better justify exclusion of, "exposure assessment of simulated air quality that meets current or alternative PM standards".

The Plan repeatedly declines to conduct such an analysis, but the reasons are not apparent. They should be made explicit, as it seems that such an analysis would be a useful and even obvious goal. The authors appear to have serious misgivings about excluding it, since it is restated several times (e.g. at page 1-9, lines 15-16; page 1-12, lines 12-14; and page 4-2, lines 12-15). This seems to contrast with the explicitly stated goal (Figure 3-1) of making quantitative risk estimates at these simulated levels of air quality.

2. Better justify exclusion of air quality data from health endpoint analyses (exceptional events, indoor, personal cloud), that are reasonably excluded form regulatory analysis.

As exceptional events subject populations to exceptional levels of PM exposure, these events may have health endpoint consequences. Their exclusion from exposure modeling (page 2-3, lines 22-27) may inappropriately underestimate exposure, and bias selection of a health protective standard. It may also introduce avoidable variability between C-R relationships obtained at contrasting locations.

The proposed exposure analysis explicitly ignores indoor PM, while giving careful consideration to time spent indoors and therefore partially shielded from outdoor PM (page 4-14, lines 6-12). Assuming that health impacts of indoor and outdoor PM are similar, exclusion of indoor sources will inappropriately underestimate exposure. Similarly, exclusion of the PM contribution of the personal cloud, particularly in the unstirred air indoors, requires justification. At a minimum, the suggested offset of 2-4 $\mu g\ m^{-3}$ (page 4-14, lines 13-23) could be incorporated in indoor exposure estimation, while suggesting that further research is required.

3. Evaluate but do not rely on the proposed model (CMAQ)-based-rollback to simulate air quality just meeting current or alternative NAAQS.

The application of CMAQ to this problem has considerable merit. However, there is little evidence presented that serious error will be inserted into the analysis by using the historical data and a proportional rollback approach. At this time it seems prudent to continue the conceptually straightforward practice of assuming historical continuity in the patterns of PM reduction, while using the current review to demonstrate the power of the new technique.

4. **The Plan can be condensed for clarity and ease of reading.**

Many sections of the text can be condensed. Others are partially redundant. For example, the description of the APEX model can be made much more straightforward (Page 4-4, line 16-23, can be combined with page 4-3, lines 5-16). Consideration of alternative models can be much reduced with inclusion of appropriate references.

Comments from Dr. Joseph Helble

The general plan presented in section 2 is reasonable. The following two comments are offered.

1. The text refers to “air quality distributions” and “air quality concentrations” that will be simulated or otherwise used for assessment. It is not clear what is meant by an AQ distribution or concentration. PM concentrations, particle size distributions, and chemical composition distributions as a function of particle size can be measured. Presumably this is what EPA means in this section, but additional detail to describe plans and avoid the use of terms such as “AQ distributions” would be helpful.

2. In the PRB discussion, CMAQ modeling of a baseline considering only natural emissions from the US, Mexico, and Canada is described. It is also worth modeling background considering only natural emissions from the US with natural plus anthropogenic emissions from the US and Canada to provide a clearer indication of what falls directly under US oversight.

Comments from Dr. Rogene Henderson

Response to charge questions

1. Selection of health effects

a. I agree with the endpoints selected.

b. I see the quantitative risk assessment for birth outcome as potentially quite problematic. The timing of the exposure would be important and this type of exposure information would be difficult to obtain. I am not enthusiastic about this approach.

c. I agree with the approach presented on page 3-3, lines 1-4, in which a limited quantitative risk assessment would be conducted for short-term exposures to $PM_{10-2.5}$

2. Specification of concentration-response functions

a. Agree

b. I am not qualified to comment here. I am not familiar with Empirical Bayes.

c. Agree

d. Agree, especially because various cut-points will be considered (Page3-18, lines 5-7.

3. Selection of urban study areas

a. I think the selection criteria are complete and appropriate. I was curious as to what urban areas would be representative of vulnerable populations.

b. I think the agency has done as well as they can considering the limited data. I wondered why Table 3-3 did not include the studies on coarse particles conducted in the Coachella Valley of California by Lipsett and Ostro.

4. Addressing uncertainty and variability

a. I agree with it.

b. I think a major uncertainty is how to deal with co-pollutants and how they contribute to the health effects attributed to PM. The document discusses this well.

5. National patterns of risk

a. I think it is good to attempt this. There will be many uncertainties, but this is discussed in the document.

b. I like this idea.

Comments from Dr. Mort Lippmann

Chapter 1:

Overall, the chapter is well constructed, concise, and on target.

Specific Comments:

Page line Comment

1-6 8-12 I can't tell, on line 12, what was similar for the two cities. Was it incidence or something else?

1-6 21-25 The Administrator, by all precedents and logic, should have used a larger margin of uncertainty in such a situation!

1-8 11 "considering" is too uncertain a commitment. For the health effect of greatly increasing certainty and overall impact, it must be done!

1-10 4 Does this apply only to $PM_{2.5}$?

1-10 12 Does birth "outcomes" include mortality?

1-11 12 Correct "$PM_{10-2.5}$".

Chapter 2:

No comment other than well done.

Chapter 3:

Overall, this chapter is well constructed, covers many complex issues in a logical manner, and has selected studies on health effects appropriately.

Specific Comments:

3-3 10 It may be true that is "difficult", but that is not an adequate reason.

Note: For my next two comments, I understand that OAQPS has to use the judgments made in the draft ISA, and I will do my best to see them modified before the final ISA is completed.

3-7 26 For mortality effects associated with short-term exposure to $PM_{2.5}$, "likely" should be deleted.

3-8 2 For mortality effects associated with long-term exposure to $PM_{2.5}$, "likely" should be deleted, at least for cardiovascular mortality.

Chapter 4:

This chapter is very well constructed, covers many complex issues in a logical manner, and has selected exposure models appropriately. Its success in the face of a great deal of complexity is attributable to the authors' now long experience, dedication, and persistence. I have no nits to pick on this chapter.

Comments from Dr. Kent Pinkerton

Chapter 3 – Scope and Approach for the Health Risk Assessment

1. Regarding **selection of health effects endpoints** to model in the risk assessment:

    a. To what extent are the Panel members supportive of EPA's planned approach to **focus on selected health effects endpoints** (e.g., emergency department visits and hospitalizations for ischemic heart disease) within broader health effect categories (e.g., cardiovascular morbidity) initially classified in the first draft ISA as having a causal or likely-causal association with ambient $PM_{2.5}$?

Reply: The stated goals for the design of the human health risk assessment are 1) to provide estimates of the potential magnitude of premature mortality or selected health effects with recent ambient exposures to PM just meeting the PM standards, 2) to develop an understanding of the influence of inputs and assumptions on risk estimates, and 3) to gain insights on the distribution of risks and patterns of risk reduction and uncertainties in risk estimates made.

Based on these stated goals, I agree with the approach to focus on selected health effects endpoints such as emergency department visits and hospitalizations for ischemic heart disease for health effect categories such as cardiovascular morbidity. However, I also feel other endpoints which will include respiratory endpoints should also be considered. As listed, short-term PM effects include cardiovascular morbidity (causal), respiratory morbidity (likely causal) and mortality (likely causal). However, it is not clear how will you take into consideration susceptible individuals with pre-existing health conditions and/or susceptibility to other factors such as infection. Never-the-less, selection of health endpoints which have an established causal or likely-causal association is highly appropriate.

    b. What are the Panel members' views regarding EPA's plans to consider for inclusion **additional health effects endpoints (e.g., birth outcomes)** for $PM_{2.5}$ that are within broader health effect categories (e.g., reproductive, developmental, prenatal and neonatal outcomes) that have been initially classified in the first draft ISA as having suggestive evidence of a causal association? What are the Panel members' views with respect to addressing the challenges in designing a quantitative risk assessment to appropriately consider birth outcome endpoints?

Reply: To expand further the focus of the risk assessment which include additional health effects categories is reasonable. Inclusion of birth outcomes is important and timely, based on

established casual or likely causal association with PM2.5 which will also encompass reproductive, developmental, prenatal and neonatal outcomes. However, under the current rationale provided in Chapter 3, it remains difficult to determine how highly divergent endpoints such as ischemic heart disease and birth outcomes will be handled. Each of these health endpoints will require extensive efforts to obtain and interpret findings for establishing a PM standard. It would appear these health endpoints for the consideration of a health risk assessment plan will examine short-term effects of PM, while others will be dependent on long-term PM effects. Finally, the authors need to clearly delineate how the health risk plan will be implemented that extends beyond what is already available in the published literature regarding these health endpoints.

c. Are the Panel members generally supportive of EPA's planned approach to conduct a risk assessment for $PM_{10-2.5}$ considering health effect endpoints within broader health effect categories that have been initially classified in the first draft ISA as having suggestive evidence of a causal association?

Reply: This planned approach appears to be based on limited data from what is presented in the ISA first draft which yields no strong health endpoint categories for thoracic coarse particles. A broader range of health effect categories would be logical for the analysis of risk assessment for thoracic coarse particles. There should be available published studies to provide just justification. However, the EPA planning team will need to render a clear rationale of how they will be able to elucidate the effects of thoracic coarse particles that will distinguish the effects from concomitant or simultaneous exposure to PM 2.5.

2. Regarding specification of **concentration-response functions** for use in the assessment:

a. In modeling health impacts associated with short-term ambient PM, to what extent are Panel members supportive of EPA's planned approach to place emphasis on distributed lags, where they are available, with additional lags (e.g., 0, 1 day lags) being included as part of sensitivity analyses, based on consideration of the degree of biological support for these lags?

Reply: Concentration-response relationships are of great relevance to clearly establish. The components of the model to include 1) concentration-response, 2) air quality parameters and 3) baseline health effect incidence rates and demographics. Each of these components of the model is highly relevant. However, each also present highly complex and labor intensive efforts to obtain. An emphasis on distributed lags where available and additional lags (e.g., 0, 1 day lags)

seems very logical. Lag times continues to be an important issue in order to establish the plausibility and relationship to adverse PM health effects. Epidemiologic studies represent our best choice to establish such relationships.

b. What are the Panel members' views on EPA's planned approach to place emphasis on **multi-city studies which provide city-specific effect estimates**, particularly Empirical Bayes adjusted effect estimates which consider both the regional signal as well as the local (city-specific) signal in deriving adjusted city-specific effect estimates?

Reply: Multi-city studies with city-specific effects estimates sound like a powerful way to best determine and tease out specific PM effects. The question is will the selection of cities and the robustness of the data collected (or that is already available for analysis) be sufficient to use empirical Bayes-adjusted effect estimates? I have very little knowledge of this estimate technique approach to have an informed opinion. I would ask that the staff team simply justify why this approach is being proposed. It is also critical to explain why existing peer-reviewed studies already published are not sufficient from which these estimates could be made. The authors have listed 50 possible sites, all with one or more publications for short-term and 17 studies for long-term epidemiological studies for C-R functions for the PM 2.5 risk assessment.

c. What are the Panel members' views regarding EPA's planned approach place **equal weight on single and multi-pollutant models** in recognition of the competing advantages and disadvantages provided by both types of models?

Reply: To use single and multi-pollutant models seems highly logical in view of the fact these models have been used previously. It is clear that every study in a multi-city analysis will have to deal with multiple pollutants. Therefore, to be able determine the extent to which co-pollutants contribute to the observed health effects will be essential.

d. Based on information provided in the first draft ISA regarding potential population thresholds, EPA is planning to place primary emphasis on modeling risk down to policy-relevant background or the lowest reported measured level in the epidemiological studies. In contrast to the prior review, EPA is planning to place less emphasis on consideration of hypothetical population thresholds. What are the Panel members' views on this approach or alternative approaches that could be considered?

Reply: To model risk down to policy-relevant background levels and/or the lowest reported measured levels from past epidemiological studies seems very reasonable. Please move forward with your ideas!

3. Regarding **selection of urban study areas**:

    a. EPA plans to include 15 to 20 urban study areas in the $PM_{2.5}$ health risk assessment, with areas selected based on application of the criteria presented in the plan. What are the Panel members' views on the planned criteria for selecting urban study areas?

Reply: The selection of urban study areas that may provide risk estimates with a higher degree of overall confidence due to location-specific data sounds like good science. The use of well-documented 1) air quality data, 2) location-specific C-R functions and 3) baseline incidence rates and demographic data add confidence to making the most appropriate judgments for selection of urban study areas. The additional factors of selecting for geographic heterogeneity, areas with large vulnerable populations and opting for epidemiological studies with more refined exposure metrics adds further confidence to your planned approach. However, again please clearly define what or who represent vulnerable populations. Also be sure the selection of diverse geographic heterogeneity as well as vulnerable populations does not dilute your power to interpret your findings.

    b. The scope of the planned assessment for $PM_{10-2.5}$ is much more limited. What are the Panel members' views regarding the overall approach and, specifically, on the criteria for selecting study areas for evaluating the health impacts associated with ambient thoracic coarse particles?

Reply: Again, consider the power and certainty of any analysis for the assessment of thoracic coarse particles. Also consider how best to use the resources you have available for conducting the health risk assessment for the PM document.

4. Regarding the approach for addressing uncertainty and variability:

    a. In the plan, EPA describes an uncertainty analysis approach based on the application of single and multi-element sensitivity analysis. What are the Panel

members' views on this planned approach for addressing uncertainty in the risk assessment?

Reply: I assume the multi-element sensitivity analysis refers to the use of high, medium and low designations? The authors define uncertainty as the lack of knowledge for both the actual values of the model input variables and the physical systems or relationships. This definition is not very helpful to the uninformed, naive reader. Although the definition continues to remain unclear to me, the uncertainty analysis approach as outlined in Figure 3-2 seems to incorporate the essential elements used in risk assessment modeling to provide some sense of the users know what they are doing.

b. Do Panel members generally agree that the planned approach sufficiently captures key sources of variability related to PM-related risk? Are there any important sources of variability which are not captured by the proposed risk assessment approach and, if so, what are the Panel members' views regarding how these sources of variability could be incorporated into the analyses?

Reply: Based on the listed key sources of variability: 1) PM2.5 composition, 2) spatial gradients in PM2.5, 3) demographics, 4) behavior related to PM2.5 exposure, 5) susceptibility of the population, 6) differences in baseline incidence of disease and 7) longer-term temporal variability in ambient PM2.5 levels, I think virtually all bases have been covered!

5. Regarding analyses being considered to place the urban study area risk results in a broader context with regard to national patterns of risk and risk-related indices:

a. EPA is considering conducting a national-scale health impact assessment of the mortality impacts in the U.S. population associated with long-term exposure to ambient $PM_{2.5}$ under recent air quality conditions to support interpretation of the risk estimates generated for the urban study areas. What are the Panel members' views related to including such a national-scale analysis in the risk assessment? What are the Panel members' views regarding the general structure and overall design of this analysis?

Reply: The approach of the risk assessment team seems to be soundly based. The analysis of studies of longer-term PM2.5 exposures with mortality as an endpoint and strong available C-R information would be major strengths. The Benefits Mapping Analysis Program (BenMAP) sounds like an excellent resource.

b. EPA's planned approach also includes analyses to compare the information for the selected urban study areas with national statistics for a set of key PM risk-related indices (e.g., baseline incidence rates for health effects modeled in the risk assessment, rates of air conditioner use, housing stock). What are the Panel members' views on this planned comparison?

Reply: Sounds good, but how straight-forward is this type of information for interpretation? If it is reasonable, easily accessible, this would be a nice option to pursue.

Comments from Mr. Rich Poirot

1. **Do Panel members generally agree with the planned approach for obtaining and analyzing the air quality data that will be used in the risk and assessments?**

Yes, I think the proposed approach for obtaining and analyzing data for the risk assessments is reasonable.

You indicate that the 2005-2007 data have been filtered by application of the Exceptional Events (EE) Rule for 24-hour NAAQS designations. Does this mean that no data were found to qualify for exceptional event status relative to the annual standard (or is there no policy for this)? It might be informative for the panel if you provided some examples of how the EE screening affects the resulting data and metrics like the 98^{th} or 99^{th} percentiles at selected sites. I wonder how EE-type events were handled in the epidemiological studies?

You indicate that continuous PM data from non FRM/FEM samplers may be used for risk assessments at locations where epidemiological studies were based on such monitors. If that's the case, what procedures, if any, will be applied to convert the non-FRM/FEM data and response functions to "FRM-like" units?

Given that the availability of $PM_{10-2.5}$ data are already quite limited, what approaches, if any, will be considered to account for what may be assumed is the large spatial variability in coarse particle concentrations and exposures?

Given the typically different sources for fine and coarse particles, I'm not sure its logical to assume EEs for $PM_{2.5}$ would also be EEs for $PM_{10-2.5}$. Forest fires, for example, may result in relatively small contributions to coarse mass.

2. **With regard to approaches for simulating air quality that just meets the current or alternative standards under consideration:**

a. What are the Panel members' views on the planned use of a proportional (i.e., linear) approach to adjusting air quality (proportional rollback)?

I think the proportional rollback approach has been reasonable in the past given the absence of viable alternatives for estimating a "more realistic" shift in the distribution of concentrations below the standard(s). It would be useful to include this approach in the current assessment, for comparison with both historical results as well as those from other methods of reducing concentrations – such as those based on historical trends or modeling future emissions changes.

b. What are the Panel member's views on also considering the alternative rollback approach being considered for $PM_{2.5}$ (model-based rollback)?

I like the proposed alternative rollback approach, based in part on logical assumptions about emissions controls pending and/or likely in the relatively near future. It should be cautioned though that sometimes the best laid plans…don't always work so well. For example, a year before the 1996 PM NAAQS revisions, EPA had already issued the final CAIR Rule (Clean Air Interstate Rule - for which the first phase reductions were to commence in 2009 for NOx and 2010 for SO_2), but the current status of those reductions is uncertain. Still, I think this (modeled) approach could be a useful current and future tool. It might allow, for example, consideration of differential effects of different PM species, sources, or pollutant mixtures, comparisons of alternative control strategies, and assessments of benefits gained or lost by speeding or delaying the implementation

For example, EPA had estimated that CAIR would reduce acidification, improve visibility, and result in $85-100 billion in health benefits each year, preventing 17,000 premature deaths, 22,000 non-fatal heart attacks, 12,300 hospital admissions, 1.7 million lost work days and 500,000 lost school days. If all this is true, then why was it prudent to phase in the program so gradually over a 10 to 15 year period and to limit it to the Eastern US only (and what are the health costs of further delaying its implementation)? Maybe a modeled rollback approach could be used to evaluate benefits of more rapid or larger-scale emissions reductions.

It isn't clear to me how estimated reductions of specific future controls could be linked to the concept of just attaining the annual or 24-hour standards, as it seems likely such programs would tend to undershoot or overshoot the 24-hr or annual NAAQS. The approach seems promising, but more detail would be helpful.

3. What are the Panel members' views on the planned approach for estimating and using policy-relevant background concentrations?

I think the proposed approach (GEOS-Chem + CMAQ) for estimating policy relevant background seems reasonable. I assume this approach can be demonstrated to work significantly better than use of GEOS-Chem alone? Is it also assumed that this approach is superior to the use of data from selected IMPROVE sites and species (excluding sulfate or sulfate & nitrate) that was considered in the last review cycle?

A relatively poor model performance (underestimates) in the West – if such results are due, as suggested, to failure of the relatively coarse 36 Km grid structure to capture influence of local emissions in mountain valleys - doesn't necessary reflect poorly on the model's estimates of Western PRB. To evaluate this you might stick to comparisons with only higher (relative) elevation ridge-top monitors. It would also be interesting to compare these modeled PRB calculations (or the natural source component of PRB) with the estimates of natural background (mean and deviations) that have been made for the IMPROVE sites.

Presumably, the PBR calculations include many site-days that might have qualified as exceptional events, had they been measured and caused violations. How do these influences relate to use of measurement data with EEs removed? Are there PRB calculations for $PM_{10-2.5}$?

As indicated by CASAC comments in the last PM review, PRB is un-measureable and therefore fundamentally unknowable, and might most efficiently be handled by using health assessment metrics that minimize its importance.

Comments from Dr. Frank Speizer

Chapter 1:

Page 1-6, Last sentence beginning line 21. I believe this is somewhat revisionist history. The Administrator not only "more heavily weighed…" than did almost everyone else, he essentially bowed to pressure from the White House Offices to redefine any scientific considerations of the meaning of margin of safety to make the definition of uncertainty fit the political needs of his bosses. I suggest a more thorough chronology of the facts be presented, perhaps including the recent court ruling that rejected his logic.

Page 1-8, Section 1.2: As part of the goals of this REA I believe it would be appropriate to re-introduce the concept of "margin of safety". As this section is written it appears the thinking is continuing along the line of "uncertainty" as the objective itself (and thus a reason set less stringent standards) rather than uncertainty being used to inform the margin of safety.

Page 1-10, Line 10. What about total mortality for PM2.5? More generically should these risks be assessed for all those specific categories in which the data summarized in the ISA identified as "suggestive of causal" rather than just those in which causal is likely or firm? Although this might broaden the work load considerably it would provide a more thorough picture of certainty and uncertainty of risks. (It probably makes more sense biologically and scientifically than focusing on birth outcomes).

Page 1-11, line 7: What does "sufficiently suggestive" mean. The ISA uses suggestive without the adverb.

Page 1-11, Paragraph beginning line 9. Although I would tend to agree I THINK CASAC AS A GROUP NEEDS TO SIGN OFF ON THIS. IN PARTICULAR I AM CONCERN THAT MORE MIGHT BE SAID ABOUT ULTRAFINES

Page 1-12, Figure 1-2: Because of the way the chapter 8 in the ISA is organized should there be a separate line in the APEX model for Vulnerable Groups to go along with Sensitive Populations? For example, urban/downtown centers or near stationary power sources may have significantly greater impact on a sub segment with closer proximity than the entire metropolitan district, which might include suburbs as well. In fact the rest of the page essentially says that is what will be done!

Chapter 2—Useful summary of plan.

Page 2-9, paragraph 2.5. Not mentioned here but probably to be considered is the variation in seasonal effects across regions. If data exist suggest it also should be considered.

Chapter 3

General Comment: As indicated above I am biased toward accepting "suggestive of causal" for evaluation. This will need to be discussed more fully among the members of CASAC with staff as to what it will mean to add this category to the risk assessment. The summary tables suggest that there might be additional endpoints, which have not made it into the categories of causality (e.g. cancer). The details of the methods seem reasonable but I leave to others with more expertise to comment.

Specific Comments;

Page 3-3, last two sentences beginning line 7: This presents an interesting problem. It seems that this says we must throw out the bulk of the data we have to make estimates with a very weak set of data that can only lead to substantial uncertainty in what gets done. Would it not be better to explore more fully ways in which the PM10 data (in conjunction with the PM2.5 data) could be more effectively used to make estimates for PM10-2.5? I admit I do not know how to do this but have we really explored all possibilities?

Section 3.2.2 Needs to be expanded to include suggestive causal categories.

Page 3-21, Section 3.3 This is an important and appropriate section. Either up front in this introductory section or certainly in the section later that deals with uncertainty, some discussion need to deal further with what happens with the uncertainty, particularly to the degree that it can be quantified. Surely an alternative, or as part of the discussion of what to do with the uncertainty must relate to how it is used in estimating the appropriate margin of safety. Can we begin to decide on how much uncertainty needs to be considered in constructing an adequate margin of safety? Or how much uncertainty results in setting the standard at a level that is not too stringent, given the law states with an adequate margin of safety?

Page 3-33 Section 3.6: Agree with plan to use, if possible suggestive causal data for long term PM10-2.5 as outlined in this section. Reasonable set of studies as outline in Table 3.3. However, as indicated above suggest explore other possible approaches to use PM2.5 and PM10 data to make estimates of PM10-2.5 to expand potential health data base for these analyses. (Maybe this will come up as option in Chapter 4).

Chapter 4. I applaud the effort to do this, and the chapter provides sufficient information to suggest that data do exist. The problem will be both having the time and expertise to carry out the appropriate analyses. I lack the technical expertise to know how many of the procedures proposed are "off the shelf" and already in the appropriate literature. If this is the case than I encourage staff to go ahead. The issue will be to the degree that the analyses are novel they should be peer reviewed and at least in press by the time they are used.

Comments from Dr. Helen Suh MacIntosh

Chapter 4: Population Exposure Analysis

1) *What are panel members' views on the general structure and overall design of the exposure assessment to provide insight on population exposures with respect to informing the interpretation of available epidemiological studies?*

The inclusion of a quantitative exposure assessment is an important component of the health risk assessment. The proposed exposure assessment rightly focuses on $PM_{2.5}$ and its intent is appropriate – to provide insight on population exposures with respect to informing the interpretation of available epidemiological studies. Specifically, the exposure assessment is intended to help identify various personal and building-related factors that may account for some of the variability in $PM_{2.5}$-associated health risks. To do so, the exposure assessment will predict 24-h population exposures for each of ten cities using the APEX model. The focus on 24-h $PM_{2.5}$ concentrations is appropriate given the importance of community time-series study findings; however, given the likely causal relations between $PM_{2.5}$ and mortality, it is also important to characterize annual exposures for the interpretation of chronic $PM_{2.5}$ studies.

Population exposure assessment will be useful in characterizing the relation between the ambient concentration and mean population exposure (or the ambient exposure factor) and the variability in population exposures – both of which are important to the interpretation of community time series studies. Since APEX is able to estimate both, it has the potential to achieve the goals proposed for this exposure assessment. Before APEX can be used, however, it should be validated and its performance assessed for each city, beginning with the pilot study in Detroit. In addition, the added value of APEX over other more simple exposure assessment methods should be assessed and described. This added value should be balanced with the model complexity and its many data inputs and assumptions and also should be balanced with the fact that exposures given future or projected scenarios will not modeled. Simple methods to be considered may include (but are not limited to) GIS-based linear regression or spatial models to estimate impacts of location on outdoor spatial variability, statistical models examining the relation between measured exposures, indoor, outdoor, and ambient concentrations for distinct population sub-groups, and/or analytical models accounting for the influence of different exposure-related variables on pollution-effect associations. While not perfect, some of the simpler methods may be equally or more effective in identifying factors affecting population $PM_{2.5}$ exposures and their relation to corresponding ambient concentrations and health risks. If so, these simpler methods may be preferable to the more complicated model. .

In addition, although the document provides a good and thorough description of the APEX model structure and data inputs, the document does not provide a framework for how the model and its results will be integrated and used to interpret epidemiological studies. As currently presented, the APEX model and its results are disconnected from the basic intent of the work. To maximize the effectiveness and usefulness of the population exposure assessment, a data analysis plan should be developed that describes how the model results and calculated personal exposure factors will be used to (1) explain observed health risk variability and (2) identify and quantify the influence of important personal and building factors on this variability. In addition, the plan should propose an approach to integrate results within and across cities.

In this regard, the pilot exposure assessment to be conducted for Detroit presents a valuable opportunity to refine aspects of the proposed exposure assessment framework and approach. Results from this pilot study should be connected to a time-series health study to examine whether and how health risks obtained using ambient concentrations as the exposure measure differ from those using the estimated population exposures for the entire city, by season, for different susceptible age groups, and for different SES groups.

2) *What are panel members' views regarding the planned measures of exposure?*

As above, the exposure assessment is focused on 24-h population exposures, which is appropriate given the importance of time-series study findings in the causal determinations. Since long-term PM2.5 exposures were also found to be "likely causal" of adverse health impacts, it would also be important to develop a plan to characterize annual exposures to help in the interpretation of chronic $PM_{2.5}$ studies.

While interesting, the consideration of additional indicators of exposure (as indicated magnitude and duration of exposures, frequency of repeated high exposures, and ventilation rate) will not likely inform the interpretation of time-series epidemiological studies. These efforts, however, will be useful in explaining variability in autonomic function and other intermediate marker studies that show sub-daily exposures to be important exposure windows.

3) *EPA is planning to focus the exposure assessment on a subset of the urban study area evaluated in the risk assessment. What are panel members' views regarding the selection of these study areas and the planned time periods to be modeled?*

EPA proposes to select ten cities and time periods to correspond to the cities used in the health risk assessment. These cities will also be selected to be diverse, as assessed by geographic location, $PM_{2.5}$ composition, air conditioning use, demographics included SES, and/or baseline health rates. This strategy is appropriate and coordinates well with the planned health risk assessment.

4) *Regarding the approach for addressing uncertainty and variability, are Panel members generally supportive of the planned approach?*

The approach used to address uncertainties and variabilities in the model and its inputs is well thought, thorough, and builds upon previous work and findings.

Comments from Dr. Ted Russell

I am generally pleased with the PM NAAQS Scope and Methods Plan for Health Risk and Exposure Assessment (hereafter, SM). It lays out a reasonable path that will provide desirable information on the potential risks from and exposures to PM, how those risks may respond to revised PM NAAQS (both $PM_{2.5}$ and $PM_{10-2.5}$).

I do note a few deficiencies, both in the document as well as the plan.

First, it would have been very nice if the document had a section summarizing criticisms by CASAC and others on the prior risk and exposure assessments, and how they have responded. This could be done by grouping the types of comments made, and how they plan to address them, and where in the current document the planning takes on those criticisms, very much like a typical response to review document. This should become standard in the process.

Chapter 1

On page 1-8, they are considering a nationwide assessment of the potential magnitude of premature mortality. This should definitely be more than a consideration, and should be done. This analysis should, likewise, provide comparisons of the potential risks of meeting the current and alternative standards. I am actually less keen on the exposure analysis next discussed (though still believe it should be done), unless there is a clearer linkage identified between the results of that analysis and how consideration of those results would be reflected in the decision process of revising the NAAQS. We have used those results in the past, though particularly when there was clinical data to suggest exposure levels of concern. Besides additional understanding of the epidemiologic results, what is to be gained? This analysis may prove to be very resource intensive, so additional thought needs to be given as to what aspects of the product would be of use. I personally find the distribution of exposures, both between individuals and by location and source, to provide insight, and to assess the potential of certain subpopulations or individuals to be particularly exposed. I would actually like to have this extended to simulating meeting the standard/alternatives.

Chapter 2:

The air quality considerations chapter adequately lays out the data needed, though the treatment of compositional data should be strengthened. There is growing evidence of the differences in health impacts between PM components, and those components vary spatially and respond differently to controls. Yes, the data to address compositional differences is less extensive, but I think still adequate to consider a more thorough treatment.

A major concern is the potential treatment of PM10-2.5. This issue is made more difficult given the relatively weak foundation in the ISA. A major concern is the relative impact of PM10-2.5 from natural and anthropogenic sources, and the correlation between those two.

In this SM, they will use estimates of PRB in their adjustment of PM levels. One question this brings up is what is the correlation between estimated PRB and the observed concentrations? There is reason to suggest that there will be some correlation (possibly negative). For example, natural dust is likely to be higher on windy days when anthropogenic PM levels are low. This has potentially important implications in terms of rolling back 24-hour levels. It is true, however, that given that the PRB levels are typically rather low, so the concern might be minor. This should be assessed.

On page 2-7, they consider using a data melding process using CMAQ and observations to estimate how PM levels should be adjusted. I can agree that this might be "better", but they need to identify how they will identify if the results are, indeed, better.

On page 2-8, they ascribe a potential issue with the poor performance of CMAQ in the West to model resolution. This should not be said without some further foundation. There are a variety of other possible reasons as well. One question this brings up is does the poorer performance impact the analysis significantly.

Chapter 3.

As noted above, I think there is sufficient information to begin addressing how composition may play a role in the PM risks, and that they can start to address this. It may be that, in the end, the decision is that there is such uncertainty as to not place much emphasis on the results using composition, but the information is still informative, and lays the foundation for the next review. Compositional information is going to be used in the Visibility SM, and it can be used here as well.

On page 3-29, they note that they have greater confidence in the risk estimates for the base case than for the other cases. However, the estimate of difference in risk may not be. This should be assessed.

As noted previously, I strongly encourage EPA to move the national-level assessment from being considered to being a central part of the REA. Also, they suggest they will use a CMAQ model run for one year, and then add parenthetically, e.g., 2005. Given that they are going to use 2004 in their PRB calculation, I would consider doing the same for consistency, all else equal.

Specific: Page 3-19: Do you mean vulnerable and/or susceptible populations?

Chapter 4.

As noted previously, I would think long and hard about how the results from the detailed exposure modeling will be used in the decision process to possibly revise the NAAQS. This might impact how the modeling is done and which results are highlighted. I was hoping for a more extensive plan to evaluate the model results.

Comments from Dr. Sverre Vedal

Chapter 1. Introduction.

This chapter works well as an Introduction/Summary.

1. I support the decision to:

   i. also carry out a nationwide risk assessment for mortality endpoints(1-8, line 11);
   ii. not carry out risk assessments for PM components or sources, or for the ultrafine size range of PM;
   iii. not to carry out an exposure assessment for the purpose of performing quantitative risk assessments based on clinical studies (1-11, line25).

2. I do not support the decision to:

   i. consider a PM2.5 risk assessment for birth outcomes or other outcomes for which the evidence is merely suggestive.

3. Based on details provided in Ch.3, I presume that the example endpoints provided for the risk assessment (1-10, lines 4-5) are just a small subset of the potential endpoints being considered.

4. Do EPA staff have confidence that the risk assessment this time will play a more influential role in the Administrator's decision-making process than it did in the most recent PM NAAQS deliberations (1-6)?

Chapter 3. Health risk assessment.

1. The list of contributors to uncertainty for the qualitative uncertainty characterization seems incomplete (3-24). For example, what about measurement error and its effects on the slope of the C-R function and on the ability to identify a threshold concertration?

2. I agree with the plan to limit the nationwide risk assessment to mortality outcomes associated only with long-term exposure (3-30, line 20). However, there is evidence that there are substantial regional differences in the long-term exposure estimates of PM effect, just as there are for short-term exposure effect estimates (see note 17, 3-30).