Compendium of Individual Preliminary Panelists Comments
Quantative Health Risk Assessment for Particulate Matter (February 2010)

Comments from Mr. Ed Avol

Comments on 2nd draft RA PM document Feb2010

General:
The second external review draft of the Quantitative Health Risk Assessment for Particulate Matter is a dramatic and substantial improvement over previous versions, and sets a new thoughtful foundation and prototype for subsequent risk assessments to generally follow. Improved presentations, superior figures, and useful tables and footnotes are present to guide the reader through the process by which EPA staff and consultants have reviewed available data, performed additional analyses, and interpreted the results of those analyses. The NAAQS health risk assessment process is becoming increasingly transparent and trackable (to the benefit of the reader and the public), but some convolutions and meanderings still persist to prevent complete clarity of presentation and focused presentations of conclusions, decisions, and positions. Chapter 4 provided a detailed step-by-step examination of the approach, the analyses, and the results in appropriate depth. Chapter 5 seemed minimal and uncertain as to whether the authors wanted to present it in the main body of the risk assessment or in an appendix. The addition of Chapter 6 as an opportunity to present an integrated view of the work performed is much-appreciated and was generally well-presented.

Charge Questions to CASAC:
1a) Utility of the peak-shaving approach:
The peak shaving approach seems useful and warranted, and aids in understanding the impact and effects of various reduction approaches for a range of possible 24hr/annual standards.

1b) Use of the composite monitor approach:
This may well be the appropriate approach to take, but it would be more convincing if some supporting data for this approach were provided in the appendix or main body of the document. This could readily be done using sites in several of the urban study areas with complete or near-complete data, and dropping out one or more data points, creating the "need" for fill-in strategies and a ready comparison with the actual (observed) data.

2) Selection of model inputs:
This seemed generally clear and supported. The justified emphasis on $PM_{2.5}$ and cardiovascular endpoints may be appropriate at this point in the NAAQS review cycle, but ongoing concern about other particle sizes and constituents, as well other health endpoints, makes the current specific approach time-limited and in need of likely review in successive cycles.

3) Addressing uncertainty and variability:
The discussion and clarifications regarding uncertainty and variability are much-improved, much-appreciated, and markedly enhance the credibility and stature of the entire document. The presentations are definitely appropriate and mostly clear.

4) In general, the approach seemed appropriate and well-documented. Some assumptions appear to have been made (or at least, not justified in the main body of the presentation), such as the

consideration of Philadelphia and Los Angeles as "representative" East and West Coast urban areas, respectively, for sensitivity analyses.

5a) The approach to include both 24hr and annual design values in assessing core risk estimates is a reasonable one that helps to clarify differences between reduction strategies that may not otherwise have been apparent. Expansion of the discussion to 15 urban study areas is a worthwhile endeavor that yields a number of insights and benefits (not the least of which is added credibility to the calim of national representativeness). The approach was generally well-done, added to the discussion, and enhanced the validity of the presentation. Several specific questions remain (see detailed queries below) but in general, I am wholly supportive of the presented approach and applaud the staff and consultants for their work.

5b) (No additional insights or recommendations at this time)

6a) The key policy-relevant questions have been captured in the presentation.

6b) The observations in Section 6.2 are well-supported but not always clearly articulated.

6c) The characterization of confidence in the core risk estimates is a marked improvement over previous versions of the document, and helps to establish the credibility of the overall approach. Given the available information, the presentation appears appropriate.

6d) The presentation of the placement of the risk assessment results in to a national context appeared a bit tentative and unsure. A more forceful presentation and assertion about representativeness could and should be made.

6e) The key observations listed at the end of Chapter 6 are generally present, but not presented in as clear or concise a manner as could be done. I believe it is useful to present estimated risks for a range of standards and levels, so that the Administrator has a more complete perspective of the possible implications of the various reduction strategies. However, I don't think it is worth a great deal of effort or time to estimate risk reduction at 6 or 8 or 10 ug/m^3 annual average. Practically speaking, there are US urban areas unlikely to ever get those low levels, so with a uniform national standard, it seems more of an academic exercise.

Questions:
P3-11, lines 30 on (the composite monitor approach) – The procedure is described for filling in missing data, but there is not presentation or reference to any validation of how precise, accurate, or appropriate the fill-in approach is, based on more complete data at sites within the respective urban study areas. Was any assessment of the efficacy of this approach done? The comment in the text (p3-13, lines 7-11 that in New York, for example, 2/3 of the data was interpolated may raise some cause for concern that could be allayed by providing some sample analyses with existing and retrievable data.

P4-28, line 18, sensitivity analysis modeling using Philadelphia and Los Angeles as "representative" East and West Coast urban areas…how was this decision made? On what basis

or comparison criteria? How ”representative” are either of these urban areas of their respective coastal counterparts? Some explanation or justification should be provided for these selections, even if it is only to refer to an Appendix or footnote.

P4-49, line2 – the explanation behind this discussion of determinants of $PM_{2.5}$ health estimates is confusing. On p4-48, lines 21-24, the estimates of risk in the risk assessment framework are claimed to be based on four elements: “…population, baseline incidence rates, air quality, and the coefficient rating air quality and the health outcome (i.e, the $PM_{2.5}$ effect estimates)”. On the next page, the determinants of the $PM_{2.5}$ effect estimates are claimed to”… be grouped into three areas: demographics, baseline health conditions, and climate and air quality” (which are actually *four* areas). However, it would appear that these statements argue that the fourth element of the basis for the risk estimates is composed of the first three elements (if population = demographics, baseline incidence rates = baseline health conditions, and air quality = air quality), which seems a bit circular in description and confusing.

P4-49, line 14 bullet – It is not clear why climate and air quality are included in the same bullet, since the references and variables listed seem to cover each of the climate and air quality categories separately and without overlap.

P4-51, Table 4-5, Co-Pollutant Levels row – in the text discussion of $PM_{2.5}$ and relevant risk estimates, SO_2 is specifically identified and included in the figures and tables, yet in the table presented national-level data for use in developing risk determinants, ozone is the only co-pollutant presented. It would therefore seem appropriate to note why Ozone is listed (cardiovascular mortality reference?) and why SO_2 is not.

P4-54, line27 – It is a bit confusing for the reader to follow the discussion from text to table to figures, since the number of urban case study areas plotted vary in the table (4-6) by category, are not presented in any key with the appropriate figures, and don’t appear to be consistent between the table listing and the related figure.

P4-57, Figure 4-9 through 4-16: (a) It would make it easier for the reader to follow the discussion and the illustrated material in the figures if some additional information were provided in a key or caption with the figures, such as the number (n) for each of the three data categories plotted (all counties CDF, Case study counties CDF, case study counties) and some comment about the splines or plots and inflection points displayed; (b) there is an inconsistency in the presentation of these figures, since some specifically comment on the number of urban case study areas above or below a given percentile (Figures 4-9, 4-11,4-14, 4-15, 4-16), while others have no comment (Figure 4-10, 4-12, 4-13); A summary comment for these three figures would help guide the reader as to the key implications of the provided figure; (c) Would there be additional clarity or power of argument gained by somehow highlighting or identifying the plotted values of the 15 urban study areas from amongst the larger number shown?

P4-66, lines 10-12 and P4-67 (Figure 4-17) – The text and figure are inconsistent, since the text says (with respect to the 35/15 plot in Figure 4-17) “…Atlanta, Birmingham, and Houston fall into this zone…” (Zone C), but Figure 4-17 shows Birmingham in Zone B2, Atlanta arguably on

the cusp between C and D, and Houston well into Zone D…and in fact, the next bullet in the text (P4-66, lines 13-16) discusses Houston's presence in Zone D.

Chapter 5 (National-Scale Assessment of Long-Term Mortality) – This chapter seems a bit tentative. The data presented is critical to the overall discussion, but the manner in which it is presented doesn't fit with the previous chapter or document flow. The introductory sections are fine and lay the ground for presentation of the results, but the methods section seemed spotty and uncertain in tone as to whether this should be better placed entirely in Appendix G with more details, or presented more conceptually here, with reference to the appropriate appendix for details. The two figures presented are the key information for the chapter. The text provides a description of what is literally presented in the figures, but provides little in the way of interpretation (see text related to Figure 5-3, for example). What is the intended message associated with Figure 5-3? A sentence or two would help to direct readers' interpretation or consideration of this.

P5-2, lines14 thru 21 (interplay between regional-scale reductions in adjacent urban areas) – the issue under discussion and the key problem being identified still is not clear…

P5-5, line 15 - The reference to "…7 Health Effects Institute PM regions…" seems unnecessary and distracting; this discussion of health impacts relates to geographic regions of the country (southwest, southeast, etc, as presented in Table 3.2, p3-14), and how the Health Effects Institute identifies these areas does not seem germane to the central argument.

Chapter 6 (Integrative Discussion) offers the promise of pulling together the previous chapters' presentations into one focused summary for reader review, but up through Section 6.3, too often self-destructs by using meandering phrasing and including extraneous comments in conclusionary statements. The net result of the writing approach employed is to obscure the message and frustrate the reader. There are some excellent, insightful, and important points to be made in this presentation, but they are often hidden in multiple qualifying commentaries when the main point is being presented. Use of bullets in the commentary can be an effective way of focusing the reader on the main persuasive points of the presentation, but in this chapter, their use suffers from a lack of clarity. With each bullet, make the statement, and then provide a terse justification in the ensuing paragraph. For example, the bullet might be "Dramatically different rates of mortality risk reduction are observed in areas with peaky $PM_{2.5}$ distributions, when peak-shaving, compared to proportional rollback , approaches are used." A brief supporting paragraph could use Salt Lake City or Los Angeles data as an example to make the point.

P6-7, line 21-22 – This comment is self-obvious and not necessary; of course the variability in incidence estimates is driven by differences in the study population (if the analyses are correctly performed and identifiable confounders are adjusted for). Remove this opening sentence and begin this paragraph with something like, "Substantial variability in incidence estimates was observed for the 15 urban study cases. Substantially less variability would be expected in estimates…"

P6-17, lines 26 thru 29 – This seems like a very odd disclosure comment to make as a closing comment. Under what set of circumstances does EPA believe someone reading this document

would interpret the 15 urban areas under discussion being the ONLY 15 areas in the country having health risks associated with PM2.5? This last sentence in the paragraph seems entirely unnecessary, but if there is a sense that this point is worthy of a comment, I would suggest that the last two sentences in this section of the text (lines 24-29) be deleted and replaced with something like the following: "The results of the national-scale mortality analysis suggests that the 15 urban study areas selected for presentation in this document are representative of a continuum of urban area results towards the upper end of the cumulative mortality distribution."

P6-19, lines 16 thru 23 (the final bullet on the page) – This comment, which establishes the credibility for evaluating the 15 urban study areas chosen, does not directly address the policy – relevant question posed under which it is listed, but does go to a central issue in the considering the relevance of the risk assessments undertaken. Accordingly, this bullet should be moved up to become the first bullet in Section 6.4, establishing the credibility of the study areas used in the analyses, *before* the results of those analyses and their relevance to policy are presented.

Editing Details:
P2-6, line 28 – delete one of the periods at the end of the sentence
P2-8, line 27 – something is missing where the "(chapter 5)" placeholder is in the sentence; perhaps "long-term mortality"?
P3-7, line 34 – need hyphen between "empirically estimated"
P3-13, line 6 – "synced" is mis-spelled; shouldn't this be "synched" or "synchronized"?
P4-17, line 22 – remove underline for "of total incidence"
P4-17, line 28 – change "these head negative point estimates" to "these had negative…"
P.4-23, line 13 – change "…2-4hour average risk…" to "24-hour average risk"
P4-27, line 26 – capitalize "eleven" to become "Eleven"
P4-38, line 2 – "…morality…" should be "….mortality…"
P4-40, line 26 - "…at the same urban study are…" should be "…at the same urban study area…"
P4-41, line 25 - This first sentence anticipates a listing of several factors that are then never presented, so the sentence should be re-structured, perhaps as, "There are several observations regarding key sources of … that can be made."
P4-41, line 28 – change "factor" to "factors"
P4-42, line 1 – what does "…the *highest* sensitivity analysis results…" mean? Does it mean the biggest changes, or the best outcome? A little re-phrasing would help.
P4-42, line 25 – for consistency with previous line, probably better to say "…log-linear model with fixed effects…"
P4-42, line 28 – remove hyphen from local-sources
P4-42, line 34 – should be re-phrased to read "…in the area of the curve between the LML and the PRB" rather than "…as you move below the LML"
P4-42, lines 36 to 38 – this is an important comment that, as written, is difficult to understand. I would offer the following: " Due to large uncertainties associated with C-R functions in the range between the LML and PRB, such estimates should be excluded when considering *reasonable* alternative risks associated with core risk estimates."
P4-45, footnote refers to open circles in Figures 4-6 and 4-7, but neither has open circles nor displays SO_2 risk estimates; I believe Figure 4-8 is the correct reference.
P4-46, Figure 4-8 key – add a "g" to "modelin"

P4-47, line 13 – change phrasing from "…attributable to secondarily formed $PM_{2.5}$" to "…partially attributable to secondary formation of $PM_{2.5}$"
P4-47, line 18 – this statement can be more definitive, based on Figure 4-8; Instead of "Most if not all", change to read "Virtually all of the alternative model specifications…"
P4-47, lines 29, 31, and 32 – Figure references should be to Figures 4-7 and 4-8, not 4-6 and 4-7.
P4-49, line2 – shouldn't this be four areas, not three (demographics, baseline health, climate, and air quality)?
P4-49, line3 – change "…, and climate and air quality…" to "…, climate, and air quality…"
P4-49, line31 – Incorrect table referenced; should be Table 4-5 in current draft version.
P4-54, line 5 –Kolmogorov- Smirnov test results appear in Table 4-7, not in Table 4-4.
P4-54, line21 – Figure numbers are incorrect in text.
P4-70, Table 4-8 – the entry for Urban study area Pittsburgh PA has a superscript "5" following it, but no footnote or key to what this refers to.
P4-71, line 5 – should be Figure 4-19, not F-19.
P4-75, footnote 61 – second line should read "…related mortality – see section…"
P4-75, lines 22 thru 27 – this sentence can and should be divided into several shorter sentences.
P5-5, lines 27 thru 29 – this looks like the title of a table or figure, not part of the text…???
P5-6, footnote, second-to-last line – should read "…are subject to greater uncertainty…"
P6-1, line 31 – remove one of the two periods at the end of the sentence.
P6-4, line 23 – "INTERPRETATION" is mis-spelled (missing a T).
P6-6, line 2 – it would be clearer to re-phrase this to read: "…relatively high 24-hr design values compared to the respective area's corresponding annual average design values…"
P6-6, line 6 – the term "composite monitor" should be defined or explained in a parenthetical comment or footnote in this integrative summary chapter, to avoid having readers backtrack to previous chapters to try and find the intended meaning.
P6-6, line 18 – the correct figure reference (for the current standard values) is Figure 4-17.
P6-7, lines 17 & 18 – Actual IHD mortality values in Table 4-1 are 15-19 for Salt Lake City and 1755-2222 for New York City.
P6-7, line 25 – should read "…current suite of standards…", not "…if standards…"
P6-9, line 27 – change "that" to "than"
P6-11, line 1 – change "…just meeting the a given…" to "…just meeting a given…"
P6-14, footnotes 78 and 79 – references to Figure 4-20 should be 4-19
P6-15, lines 7-9 – This sentence is garbled and confusing, and needs re-writing and clarification.
P6-15, footnote 81 – reference Figure 4-18, not 4-20.
P6-18, lines 11-12 – delete the rest of the sentence and end after "…2,000 cases per year." (This is repetitive and not needed in this summary statement).
P6-18, lines 14-15 – delete the rest of the sentence after "…in a given year for the urban study areas."
P6-18, line 17 – Delete "Generally comparable" and begin the sentence "Estimates of CV-related mortality…"
P6-18, lines 24-25 – Correct first sentence to read, "A broader array of health effects has also been associated with $PM_{2.5}$ exposures, including reproductive effects and developmental effects in children."
P6-18, lines 29-32 – Change first sentence to read:"Given the quantitative and qualitative assessments of uncertainty and variability performed as part of the risk assessment, it is unlikely

that the degree of risk remaining upon the simulated meeting of the current (15/35) suite of standards has been over-stated."
P6-18, line 41- Re-write to read: "This variability in annual-average $PM_{2.5}$ concentrations is most prominent in study areas where the 24-hour standard is "controlling"."
P6-19, line 14 – Delete the word "well"; not clear what the definition of "well below a value" is, if some of the values are said to be approximately equal to it). In this context, it seems sufficient to state that the annual average concentrations under discussion are below the level of the current annual standard.
P6-19, line 19 – delete the phrase "in most such areas"; it is redundant, given the beginning of the sentence and the sentence structure.
P6-19, line 21 – replace "reflective of" with "as do"
P6-19, line 36 – add "%' after "11"
P6-21, line 2 – delete comma after "… with a"
P6-21, line 3 – insert "U.S." before "…counties"
P6-21, line 5 – replace the phrase "…likely capture well…" with "do represent"
P6-21, lines 7-9 – why is this comment here? Isn't it self-evident in selecting a representative sub-sample? As I have commented previously, this is unnecessary and can be deleted.

Comments from Dr. Joe Brain

Charge Question #6a-e
Responses from Joe Brain
March 4, 2010

6) We have developed an integrated discussion of the $PM_{2.5}$-related risk estimates which considers the results of the qualitative and quantitative treatment of uncertainty and variability together with the various national-scale assessments completed for the analysis to support interpretation of the core risk estimates. As part of the integrative discussion, we also provide key observations that bear on policy-relevant risk-based questions.

a) To what extent does the Panel believe that we have captured the key policy-relevant questions that can be addressed by this risk assessment?

Response: Chapter 6 is the culmination a rationale developed throughout the document. It needs to pull all the pieces together and clearly integrate them in a form that provides the basis for decisions. The panel responds positively to Question 6a. We believe that this second draft does effectively capture the key policy-relevant questions that can be addressed by the preceeding risk assessment.

b) We provide a set of key observations related to estimates of risk associated with simulations of just meeting the current and alternative suites of standards. These observations are based not only on consideration of trends in risk reduction across alternative suites of standards and residual risk remaining after simulation of just meeting specific suites of standards, but also on additional factors that can impact risk (e.g., the role of annual and 24-hour design values, the peakiness of $PM_{2.5}$ distributions within a study area, and application of different rollback approaches). To what extent do the Panel members believe that the observations presented in section 6.2 are well supported by the results of the analyses? Are there other observations that might be made that would help to address the policy-relevant questions identified at the beginning of the chapter?

Response: The panel believes that Section 6.2, Interpretation of Urban Study Area Results, is well supported by the results of the previous analyses. We are not aware of other observations that might affect the answers to the policy-relevant questions identified at the beginning of Chapter 6.

c) Part of our interpretation of the core risk estimates presented in section 6.2 is our characterization of confidence in the core risk estimates and in observations made based on those estimates. These assessments of confidence are based on consideration of the results of the sensitivity analysis as well as on the qualitative assessment of uncertainty and variability. To what extent does the Panel believe that the characterizations of confidence in the core risk estimates and associated policy-related observations are reasonable given available information?

Response: The authors have done an appropriate sensitivity analysis and have adequately characterized both uncertainly and variability. The panel believes that the degree of confidence expressed in the core risk estimates is appropriate. We believe that the adjacent policy-related observations are reasonable given the information available to the EPA and summarized in the final version of the PM ISA.

d) As part of the integrative discussion, we use the results of several national-scale analyses (i.e., the national scale $PM_{2.5}$ mortality analysis, the representativeness analysis, and the new exploration of design values and patterns of $PM_{2.5}$ monitoring data presented in section 4.5) to place the results of the risk assessment in a broader national-context. What are the Panel members' views on appropriateness of this effort to place results of the analysis in a national context?

Response: One could always suggest additional cities or parts of the country where PM data could be applied. The panel believes that the national-scale analyses which were carried out are appropriate and are reasonably varied and thus they do provide the context for placing the results of the policy analysis in a reasonably representative national context.

e) We conclude chapter 6 with a list of key observations. Does the Panel believe that we have appropriately highlighted key findings of the risk assessment in these observations? Of particular note is the observation that, while alternative 24-hour standard levels can be used to reduce annual-average $PM_{2.5}$ concentrations and thus to reduce estimated risk, the results are likely to be highly variable across urban areas. More consistent lowering of annual-average $PM_{2.5}$ concentrations across study areas, and thus more consistent reductions in estimated risk, may result from application of alternative annual standard levels. We also note that simulation of the alternative 24-hour standard level of 25 $\mu g/m^3$ resulted in reductions in annual-average $PM_{2.5}$ levels for some study areas that were well below the lowest annual standard level assessed (i.e., below 12 $\mu g/m^3$). As a consequence, we observed risk reductions reflecting these changes in annual-average $PM_{2.5}$ levels below 12 $\mu g/m^3$. Given these results, does the Panel believe that there is utility in estimating risks for alternative annual standard levels below 12 $\mu g/m^3$?

Response: Section 6.4, Key Observations, is a valuable part of this chapter and of the document as a whole. The three questions asked are important and the answers provided are well founded and useful. They get to the heart of the matter. What are the consequences of maintaining the current standard? What are the likely benefits of lowering the standards? To what extent are policies about PM2.5 applicable throughout the country? The bullets following the three questions succinctly address these questions, and will be useful to the administrator in forming the basis for her decision.

Consistent with discussions at the last meeting of the PM advisory committee, we are comfortable with not estimating risks for alternative annual standard levels below 12 $\mu g/meter^3$. At these concentrations, data is increasingly less common and less reliable. Therefore, estimates of risk will be subject to greater and greater uncertainty. We believe that the risk estimates at levels higher than this provide an adequate basis for decision making.

Comments from Dr. Wayne Cascio

General Comments:
With the second draft assessment document the staff of the EPA's Office of Air Quality Planning and Standards (OAQPS) has produces an excellent document that largely addresses concerns raised by members of CASAC and the public. Specifically, the second external review draft now contains a more comprehensive discussion of the core risk estimate through the integration of uncertainty and variability analyses from 15 urban centers, and national-scale long-term exposure PM2.5 mortality assessment. The rationale for not including a quantitative assessment of risk associated with PM10-2.5 presented in section 2.3 and appendix H is justified based on the limitations in the data available for characterizing PM10-2.5 exposure and risk. The overall presentation and discussion of the consideration of variability and uncertainty is much improved.

Charge Question 2. Selection of model inputs (section 3.3): We have expanded and clarified the discussion of our rationale for identifying modeling choices comprising the core risk model, focusing in particular on selection of C-R functions (section 3.3.3). To what extent does the Panel consider this discussion to be clear and the model selections appropriate?

Response. The expanded discussion provided in Section 3.3.3 Selection of Epidemiological Studies and Concentration-response (C-R) Function within those Studies conveys the rationale for identifying modeling choices and adequately justifies the selection of specific epidemiological studies utilized to establish the C-R functions. The rules employed for study selection as described on 3.27 and 3.28 provided a strong foundation to identify studies that provide the most accurate data to derive the C-R relationship. The expansion of the discussion and integration of the ISA was also very useful in that it provides an opportunity to reinforce the gaps in knowledge, for example on 3-30 where it is stated that, "There were no multi-city studies for this [short-term exposure to PM2.5 and emergency department visits for cardiovascular and/or respiratory illnesses] category of health endpoint."

Comments from Dr. Rogene Henderson

Comments on Assigned Charge Questions for 2nd Draft of PM-RA
Rogene Henderson, March 1, 2010
Charge Question 5: Consideration of design values and patterns of PM2.5 monitoring data in interpreting core risk estimates (section 4.5): To enhance our interpretation of the patterns of core risk estimates generated for both the current and alternative suites of standards, we have included analyses of 24-hour and annual design values together with patterns of PM2.5 monitoring data for the 15 urban study areas. This reflects the fact that these two factors play a key role in determining the degree of risk reduction estimated upon just meeting the current and alternative suites of standards under alternative rollback approaches. As part of the consideration of design values, we have also contrasted the 15 urban study areas with patterns of design values seen for the broader set of urban areas in the U.S. in order to help place the urban study area in a broader national context

a) To what extent is the Panel supportive of these additional assessments?
I found Figures 17-21 quite helpful in putting the 15 urban areas into context with the monitoring values from other cities in the US. The figures also helped me to visualize what the overall picture looked like and what the controlling values are.

I would like to have more discussion about how the peak shaving rollback method was used. When the short-term values are controlling, it does not make sense to me to use the same percentage rollback for the annual values as is required by the short-term values. This results in non-feasible annual values (e.g., 7 ug/m3) for some cities such as SLC. Do we have actual evidence to indicate that if the short-term value were reduced 55% that the annual value would also be reduced 55%? I do not see the logic to that assumption.
Why not consider the risk of mortality from short-term exposures separately from the risk of mortality from the average annual exposures. It is stated in the beginning of Chapter 6 (page 6-1) that the primary focus will be based on risk associated with long-term exposure to PM, because long-term exposure to PM2.5 has been shown to produce substantially larger mortality risk compared to short-term PM2.5 exposure. Based on that, one might expect the calculated risks from the annual exposures to be the most critical consideration to protect public health. I do not think it is appropriate to mix the two design values in the rollback procedures.

But perhaps I do not understand how the peak shaving rollback method was used. It states on page 3-15, lines 20-23, that the proportional rollback method was the only one used to generate core risk estimates and the other two rollback methods were only used for sensitivity analysis. I would appreciate more discussion of this at our meeting.

I suggest adding the term "design value" to the list of terms on page viii, with a reference to page 3-16 for a definition.

b) Does the Panel have any recommendations for additional insights based on consideration of patterns in design values and PM2.5 monitoring data across the 15 urban study areas and at the national level?
It would be nice to have Figures 17-19 in three dimensions with some measure of health effects for the 15 cities on the third axis. Can that be done? It stated that the C-R functions are fairly linear and it would be nice to see that.

Chapter 6 – Integrative Discussion of PM2.5-related Risks

6) We have developed an integrated discussion of the PM2.5-related risk estimates which considers the results of the qualitative and quantitative treatment of uncertainty and variability together with the various national-scale assessments completed for the analysis to support interpretation of the core risk estimates. As part of the integrative discussion, we also provide key observations that bear on policy-relevant risk-based questions.

a) To what extent does the Panel believe that we have captured the key policy-relevant questions that can be addressed by this risk assessment?
I thought the policy-relevant questions were well stated.

b) We provide a set of key observations related to estimates of risk associated with simulations of just meeting the current and alternative suites of standards. These observations are based not only on consideration of trends in risk reduction across alternative suites of standards and residual risk remaining after simulation of just meeting
specific suites of standards, but also on additional factors that can impact risk (e.g., the role of annual and 24-hour design values, the peakiness of PM2.5 distributions within a study area, and application of different rollback approaches). To what extent do the Panel members believe that the observations presented in section 6.2 are well supported by the results of the analyses? Are there other observations that might be made that would help to address the policy-relevant questions identified at the beginning of the chapter?

I thought the list of observations was appropriate.

c) Part of our interpretation of the core risk estimates presented in section 6.2 is our characterization of confidence in the core risk estimates and in observations made based on those estimates. These assessments of confidence are based on consideration of the results of the sensitivity analysis as well as on the qualitative assessment of uncertainty and variability. To what extent does the Panel believe that the characterizations of confidence in the core risk estimates and associated policy-related observations are reasonable given available information?

The confidence statements were well presented.

d) As part of the integrative discussion, we use the results of several national-scale analyses (i.e., the national scale PM2.5 mortality analysis, the representativeness analysis, and the new exploration of design values and patterns of PM2.5 monitoring data presented in section 4.5) to place the results of the risk assessment in a broader national-context. What are the Panel members' views on appropriateness of this effort to place results of the analysis in a national context?

I thought it was a good idea.

e) We conclude chapter 6 with a list of key observations. Does the Panel believe that we have appropriately highlighted key findings of the risk assessment in these observations? Of particular note is the observation that, while alternative 24-hour standard levels can be

used to reduce annual-average PM2.5 concentrations and thus to reduce estimated risk, the results are likely to be highly variable across urban areas. More consistent lowering of annual-average PM2.5 concentrations across study areas, and thus more consistent reductions in estimated risk, may result from application of alternative annual standard levels. We also note that simulation of the alternative 24-hour standard level of 25 μg/m3 resulted in reductions in annual-average PM2.5 levels for some study areas that were well below the lowest annual standard level assessed (i.e., below 12 μg/m3). As a consequence, we observed risk reductions reflecting these changes in annual-average PM2.5 levels below 12 μg/m3. Given these results, does the Panel believe that there is utility in estimating risks for alternative annual standard levels below 12 μg/m3?

Definitely not. I think there is too much uncertainty to do that with much confidence.

Comments from Dr. Philip Hopke

In general, there are no major problems with this assessment. It is generally following the approaches that we have now seen in multiple rounds of review.

1) Air quality inputs (section 3.2): We have expanded the consideration of alternative approaches to simulating just meeting the current and alternative suites of $PM_{2.5}$ standards (i.e., rollback approaches) to include a peak shaving approach, in addition to the hybrid and proportional approaches considered in the first draft assessment. This peak shaving approach is intended to represent more localized, rather than regional, patterns of $PM_{2.5}$ reductions (discussed in section 3.2.3.3).
a) To what extent does the Panel believe that the use of the peak shaving approach provides useful additional exploration of variability associated with how ambient $PM_{2.5}$ concentrations are simulated to change upon just meeting the current and alternative suites of standards?

It does not seem useful to me since it is unlikely that there are obvious sources that could be controlled that would let one shave peaks in practice. Thus, it provides unrealistic scenarios that really do not contribute useful information to the assessment. The other rollback approaches seem fine and it is hard to see a role for the peak shaving approach.

b) We have used comparisons of composite monitor annual averages generated using the different rollback approaches as a surrogate for differences in long-term exposure-related mortality in looking across all three rollback approaches. To what extent does the Panel believe that this is a reasonable approach for assessing the impact of variability associated with simulating changes in air quality patterns on estimates of long-term exposure-related mortality?

It would make the compositing easier to follow if equations were provided. If you do not wish to put them into the main chapter, then an appendix can be added to guide the interested reader through exactly how all of the calculations were performed.

I agree with the approach to removing quarters in which too few values are reported to provide a reliable estimate of the average concentration during that period. However, it makes little sense to replace missing values with a mean value. If you are going to attempt to impute missing values, then a much more sophisticated approach should be employed using other monitors in the area as well as historic data for similar meteorological conditions. I would suggest that for those quarters where there are missing values, but a sufficient number of values to provide a valid mean value, then that should be the value used in the health analyses.

2) Selection of model inputs (section 3.3): We have expanded and clarified the discussion of our rationale for identifying modeling choices comprising the core risk model, focusing in particular on selection of C-R functions (section 3.3.3). To what extent does the Panel consider this discussion to be clear and the model selections appropriate?

3) Addressing uncertainty and variability (section 3.5): We have clarified the process used to evaluate sources of variability and added coverage for specific sources of variability (section 3.5.2); expanded our discussion of the qualitative analysis of uncertainty (section 3.5.3); and included analyses of pair-wise interactions of sources of uncertainty (section 3.5.4). To what extent does the Panel consider these discussions to be clear and appropriate?

It is time to move toward stochastic uncertainty analysis. Every time we get a risk assessment, we continue to get the qualitative review of uncertainties and some hand waving. If there are insufficient resources available to do the full stochastic risk assessment, then the Agency should state this clearly so that it can be made clear why they have chosen not to move to a more complete analysis. Otherwise, we should not see them continue to duck this approach.

Comments from Dr. Morton Lippmann
CASAC PM RA Review, 2/16/2010

Overall Comments:

The authors have been very responsive to the prior CASAC PM Panel comments and suggestions, and the 2nd draft provides a thorough and well-prepared presentation of realistic quantitative estimates of the effects of long-term ambient air $PM_{2.5}$ exposures on premature mortality in 15 US urban areas, their variability and uncertainties, and their representativeness of the US population as a whole. The alternate risk assessments (RAs) for recent years' concentrations, meeting the current suite of $PM_{2.5}$ NAAQS, and alternate NAAQS (14, 13, and 12 ug/m^3 annual and 30 and 25 ug/m^3 daily, and combinations thereof) is an appropriate way to tee up the selection options for the Policy Assessment (PA).

Having offered a strong endorsement of the overall work of the OAQPS in preparing the RA document, I need to raise two issues that I have raised before that have not been properly addressed. These are:

1) The fact that the ACS cohort is not a representative US population, but rather is of higher SES. Within this cohort, the mortality risk is lower for those of higher SES. Thus, the risk coefficient for a representative population would be higher, as it is for the 6-cities cohort that was selected to be more representative (at least for cities in the eastern half of the US).
2) The description and use of the data from the Ito et al. (2007) study has not been corrected. As noted below, this was a study covering all of New York City (Kings, Queens, New York, Bronx, and Richmond Counties), not just New York County (Borough of Manhattan).

Specific Comments on Text Entries:

Page Line Comment

3-13 Table 3.1. The entry for New York City is incorrect. Change New York City (Manhattan)" to "New York County (Manhattan)".

3-13 6 Change "New York (Manhattan)" to "New York City ".

3-30 11 Delete "(Manhattan)".

3-31 7 The number of members of the ACS cohort in the 156 MSAs with air quality data is much smaller than 1.2 million.

3-40 Counties Column for "New York". Change New York City (Manhattan)" to "New York County (Manhattan)", and delete "New York City (Manhattan)" where it precedes "Ito et al. (2007)". The Ito et al. (2007) study covered all of NYC, not just Manhattan!

3-50 Table 3-10. There should not be two entries for New York, NY. As noted above, the Ito et al. (2007) study covered all of NYC, not just Manhattan! The second New York entry, which shows only the population of New York County (Manhattan), should be deleted, and the calculations for the Ito et al. study need to be revised to reflect the whole NYC population and all of the NYC $PM_{2.5}$ monitors.

3-77 Section 3.54.3. This section refers to single- and multiple-elements, as well as to single- and multiple-factors, in terms of sensitivity analyses. There is no definition provided as to what constitutes an element or a factor. I suggest that element is a poor choice in terms of terminology, since it implies, at least to me, a chemical element. This will become important in the next round of PM NAAQS review. [On page 4-42, there is a clarification indicating that elements refers to "modeling elements"]

4-19 10 Change "fig" to "fit".

4-19 11 Change "greater" to "great".

4-36 12 Change "CPD" to "CVD".

4-45 Note at bottom of the page: Change "Figures 4-6 and 4-7" to "Figure 4-8".

4-47 29 Change "Figures 4-6 and 4-7" to "Figures 4-7 and 4-8".

4-48 para. 2 : Once again, there are elements and factors with no definitions. Here the elements are not modeling elements, but something else. [On page 4-54, line 21, I learned that these were "critical risk function elements"]. Please reword to avoid confusion!

4-55 23 Change "population" to "populated".

6-1 31 Delete extra period at the end.

6-14 25 Change "Pitts" to "Pittsburgh".

6-18 15,16 Where can we find these estimates?

6-18 21-23 Where can we find these estimates?

6-21 8 Change "PM2.5" to "$PM_{2.5}$".

Charge Questions to the CASAC PM Review Panel - focus on the charge questions listed below in review of the *Quantitative Health Risk Assessment for Particulate Matter*

Chapter 3 – Urban Case Study Analysis Methods

Charge Question #2: Selection of model inputs (section 3.3):

We have expanded and clarified the discussion of our rationale for identifying modeling choices comprising the core risk model, focusing in particular on selection of C-R functions (section 3.3.3). To what extent does the Panel consider this discussion to be clear and the model selections appropriate?

Response:
The Panel commends the authors for expanding and clarifying their rationale for identifying modeling choices comprising the core risk model in a logical and satisfactory manner. Their model selections were appropriate

Chapter 6 – Integrative Discussion of $PM_{2.5}$-related Risks

Charge Questions #6a-e:
We have developed an integrated discussion of the $PM_{2.5}$-related risk estimates which considers the results of the qualitative and quantitative treatment of uncertainty and variability together with the various national-scale assessments completed for the analysis to support interpretation of the core risk estimates. As part of the integrative discussion, we also provide key observations that bear on policy-relevant risk-based questions.

Charge Questions #6a: To what extent does the Panel believe that we have captured the key policy-relevant questions that can be addressed by this risk assessment?

Response:
The Panel considers that the authors have captured the key policy-relevant questions that can be addressed by this risk assessment that is focused solely on $PM_{2.5}$.

Charge Questions #6b:
We provide a set of key observations related to estimates of risk associated with simulations of just meeting the current and alternative suites of standards. These observations are based not only on consideration of trends in risk reduction across alternative suites of standards and residual risk remaining after simulation of just meeting specific suites of standards, but also on additional factors that can impact risk (e.g., the role of annual and 24-hour design values, the peakiness of $PM_{2.5}$ distributions within a study area, and application of different rollback approaches).

To what extent do the Panel members believe that the observations presented in section 6.2 are well supported by the results of the analyses? Are there other observations that might be made that would help to address the policy-relevant questions identified at the beginning of the chapter?

Response:
The Panel considers that the observations presented in section 6.2 are consistent with the analytical results, and provide all of the information needed in the development of the PA

document.

Charge Questions #6c: Part of our interpretation of the core risk estimates presented in section 6.2 is our characterization of confidence in the core risk estimates and in observations made based on those estimates. These assessments of confidence are based on consideration of the results of the sensitivity analysis as well as on the qualitative assessment of uncertainty and variability.

To what extent does the Panel believe that the characterizations of confidence in the core risk estimates and associated policy-related observations are reasonable given available information?

Response:
The Panel considers that the characterizations of confidence in the core risk estimates and associated policy-related observations are reasonable.

Charge Questions #6d: As part of the integrative discussion, we use the results of several national-scale analyses (i.e., the national scale $PM_{2.5}$ mortality analysis, the representativeness analysis, and the new exploration of design values and patterns of $PM_{2.5}$ monitoring data presented in section 4.5) to place the results of the risk assessment in a broader national-context.

What are the Panel members' views on appropriateness of this effort to place results of the analysis in a national context?

Response:
The Panel considers that the effort to place results of the analysis in a national context was appropriate.

Charge Questions #6e: We conclude chapter 6 with a list of key observations.
Does the Panel believe that we have appropriately highlighted key findings of the risk assessment in these observations?

Response:
The Panel considers that Staff has appropriately highlighted key findings of the risk assessment in these observations

Of particular note is the observation that, while alternative 24-hour standard levels can be used to reduce annual-average $PM_{2.5}$ concentrations and thus to reduce estimated risk, the results are likely to be highly variable across urban areas. More consistent lowering of annual-average $PM_{2.5}$ concentrations across study areas, and thus more consistent reductions in estimated risk, may result from application of alternative annual standard levels. We also note that simulation of the alternative 24-hour standard level of 25 $\mu g/m^3$ resulted in reductions in annual-average $PM_{2.5}$ levels for some study areas that were well below the lowest annual standard level assessed (i.e., below 12 $\mu g/m^3$). As a consequence, we observed risk reductions reflecting these changes in annual-average $PM_{2.5}$ levels below 12 $\mu g/m^3$.

Given these results, does the Panel believe that there is utility in estimating risks for alternative annual standard levels below 12 $\mu g/m^3$?

Response:
The Panel does not consider that estimating risks for alternative annual standard levels below 12 $\mu g/m^3$ is a worthwhile endeavor on several grounds. One, the extrapolation of the risk coefficients at such concentrations becomes increasingly uncertain; and two, the prospects of achieving compliance with such levels in the foreseeable future is vanishingly small.

Comments from Dr. Robert Phalen

General Comments:

The second draft of "Quantitative Health Risk Assessment for Particulate Matter" is clearly-written and the logical development is well-described. The staff has done an excellent job. My enthusiasm for the entire document is diminished by factors that are apparently outside of policy-driven limitations imposed on EPA staff. These limitations include:

- PM2.5 is a mass-based metric (indicator) for assessing health-effects. It is likely that specific components, such as vanadium, nickel, and elemental carbon, are actually driving the health effects. The use of a mass-based indicator can lead to air-quality standards that do not permit appropriate abatements.

- Secondary health effects, e.g. as generated by abatements that adversely affect the economy (and the many associated health effects), are not considered. Such indirect adverse health effects are real, and the affected populations must face them along with the direct effects.

- The current risk assessment does not conform to the recommendations made by the National Research Council of the National Academies (Science and Decisions: Advancing Risk Assessment, the National Academies Press, Washington D.C., 2008). Specifically, ... "that risk assessment should be viewed as a method for evaluating the relative merits of various options for managing risk rather than as an end in itself." The concept of managing risk must include all of the significant risks associated with a particular decision, not just some selected direct effects associated with the decision.

- National air quality standards have the disadvantage of forcing abatements on regions of the U.S. for which the abatements may harm health to a greater extent than they improve health. As an example consider restrictions on diesel-emissions to try to meet standards in areas where a significant portion of PM2.5 can be found in soil fine-particle mass. Control of soil aerosols is not practical, so the diesels must be targeted.

Specific Comments on Section 3.3

This reviewer found only minor errors.

Table 3-4, pg. 3-24: Replace "LA" with "Los Angeles", as LA is also the designator for Louisiana.

Line 34, pg. 3-31: Drop either "the" or "our" for clarity.

Comments from Dr. Ted Russell

Overall, I continue to be pleased with the amount of analysis conducted in this PM Risk Assessment (RA), though I am still disappointed that a more comprehensive exposure modeling effort was not made. The document provides the type of information needed to inform the review of the primary NAAQS, and does s good job of providing a quantitative assessment of the potential risks, and has done a commendable sensitivity analysis.

While I appreciate the addition of Chapter 6, it was not as effective as it might be. First, I was looking for Chapter 6 to be more integrative of the results from Chapter 4 and 5 as it is now more focused on Chapter 4. Second, I was looking for it to also integrate uncertainties in to the discussion to a greater degree. This is the point that they could bring in the issue of using the LML as the zero risk level, and how that might influence the overall interpretation of the results. Third, it was a bit repetitive, bringing up the issue of "peaky nature" and explaining it more often than necessary. It is an important concept, but it was overdone in Chapter 6. Still, Chapter 6 is a good addition, but it could be made stronger in response to the above comments.

Chapter 3 Charge Questions:

1) Air quality inputs (section 3.2): We have expanded the consideration of alternative approaches to simulating just meeting the current and alternative suites of $PM_{2.5}$ standards 5 (i.e., rollback approaches) to include a peak shaving approach, in addition to the hybrid and proportional approaches considered in the first draft assessment. This peak shaving approach is intended to represent more localized, rather than regional, patterns of $PM_{2.5}$ reductions (discussed in section 3.2.3.3).

a) To what extent does the Panel believe that the use of the peak shaving approach provides useful additional exploration of variability associated with how ambient $PM_{2.5}$ concentrations are simulated to change upon just meeting the current and alternative suites of standards?

Response: This approach, along with the hybrid approach, provide a reasonable method to get an estimate of the lower bound impact of what would happen when an urban area attains a specific standard. This approach is specifically of interest in areas where the 24-hour standard will be the driving standard, and the proportional roll-back would lead to reductions in the annual level beyond that which might be viewed as likely. The method, either in the report or in the appendices, should be specified mathematically, as well as in words. They might present a specific set of example applications of the approaches, e.g., for three different cities, at this point.

b) We have used comparisons of composite monitor annual averages generated using the different rollback approaches as a surrogate for differences in long-term exposure-related mortality in looking across all three rollback approaches. To what extent does the Panel believe that this is a reasonable approach for assessing the impact of variability associated with simulating changes in air quality patterns on estimates of long-term exposure-related mortality?

Response: First, the approach to compositing should be better explained. A critical question is how well it aligns with what is actually done in reporting for the calculation of design values. Indeed, the approach they used should be motivated by first explaining how the annual average PM is found (provide specific equations as needed), and then showing that the approach taken is in line with how the annual average PM is now found. Also, they need to explain how and when compositing is used in calculating design values. This will explain why quarterly averages are first calculated, and may impact my thoughts on how the current approach to compositing should be altered.

I do not like how they replace missing values, i.e., using the average value to replace missing values when a certain number of samples are missing. It seems to be a bit arbitrary, and could lead to a bias. Is this what is done in practice (i.e., specified)? In terms of compositing, the values could each be adjusted using a centering approach. In this case, the annual average from each monitor being used in the composite is subtracted from the daily value from that monitor, leading to a string of values that have a mean of zero. These annual averages are also used to calculate the composite annual average of the stations being used in the composite. The daily values for each monitor (after subtracting the mean of that station) are then averaged as available. This leads to the average variation from the mean for that day. The annual average composite value is then added back to get the daily composite value. This is relatively insensitive to stations dropping out. Further, it should exactly give the observed annual mean at each station, and well as the composite mean. The current approach for imputing missing days can lead to an average that would not agree with the reported value. The same approach for calculating quarterly averages, that are then used to calculate the annual average and design value, can be used.

Like my response to part (a), the mathematical equations should be provided here or in the appendix.

5) Consideration of design values and patterns of $PM_{2.5}$ monitoring data in interpreting core risk estimates (section 4.5): To enhance our interpretation of the patterns of core risk estimates generated for both the current and alternative suites of standards, we have included analyses of 24-hour and annual design values together with patterns of $PM_{2.5}$ monitoring data for the 15 urban study areas. This reflects the fact that these two factors play a key role in determining the degree of risk reduction estimated upon just meeting the current and alternative suites of standards under alternative rollback approaches. As part of the consideration of design values, we have also contrasted the 15 urban study areas with patterns of design values seen for the broader set of urban areas in the U.S. in order to help place the urban study area in a broader national context

a) To what extent is the Panel supportive of these additional assessments?

Response: These are beneficial. Figures 4-17 to 4-19, and Table 4-8 do a good job of identifying which are the controlling standards for each of the 15 cities, and also show that the cities examined do a good job of spanning the space of conditions for cities around the US. It might have been good to also identify some of the major outliers on Fig. 4-17 (there really is only one to identify). Also, you could consider color coding the dots on the figure, such that

each portion of the US is assigned a color, so it would be apparent if there were specific areas where an issue might identified. This would also help address why no Upper Midwest city is targeted. Without this, I don't think this section really does as good job of contrasting the 15 urban study areas with other areas, that they are supposed to represent, particularly at a regional level. Section 4.4 presents additional information putting other areas in to perspective, but what would be valuable is to show that all regions and sizes of cities are addressed adequately.

b) Does the Panel have any recommendations for additional insights based on consideration of patterns in design values and $PM_{2.5}$ monitoring data across the 15 urban study areas and at the national level?

Response: See above. I think this section could be strengthened by being able to have a concluding statement somewhat along the lines of "We have captured the range of conditions found for cities that would be in non-attainment of the various combinations of standards in each of the six of the seven (I would have preferred all seven) regions of the country, and the 15 cities capture X% of the population. The analysis has identified cities that span the range of non-attainment levels, and which standard would require the greatest level of control based on this analysis." It is close to being able to say this.

Other details:

Page 3-8, Footnote: This footnote is not needed and I find it confusing. Also, while it need not be added at this point, unless $\beta\Delta x$ is above about 0.2, the error in linearizing (3) and (4) is rather small. This might simplify interpretation of later analyses since it makes the response to concentration changes linear.

In Chapter 3, no city in the Upper Midwest was identified. Might this be alleviated?

Comments from Dr. Frank E. Speizer

CASAC PM Risk Assessment Draft 2 dated February 2010

Pre-meeting Comments: 3/5/2010

General Comment:
Staff has done an excellent job of using the available literature and assessments from the ISA to generate a series of outcome assessments over an appropriate range of alternative levels of both 24 hour and annual averages of PM2.5 that are justified by the available data. They appear to have been response to our previous comments on the first draft, and where they have not followed our suggestions have presented evidence that they considered our requests and accepted or rejected or modified with appropriate justification. The one area where I would have like to have seen more analyses is with regard to PM10-2.5, simply to have demonstrated the ":inappropriateness" of trying to rely on such an analysis would have emphasized the need for more data related to the course fraction effects.

Specific Comments and Charge Questions
Chapter 2
Page2.4, line 14. Typo: PM2.5

Page 2.6, line 2-4 Query the continuing developmental work on population exposure analysis methodology. It would have been useful to read or hear further what this plan would be. One would have thought that over the years much of this would have already been worked out.

Page 2.6, line 22-24. Appendix H provides a well reasoned argument for Staff concluding that they could not do a quantitative risk assessment for PM10-2.5. From my perspective it would seem to me that some estimate of risk should be made if for no other reason than to document wide confidence intervals and thus reason for not including it in the quantitative assessment.

Page 2.7, line 10-11. Perhaps at the end of this process a note could be sent to CASAC to indicate what the plan might be for the future to prepare for the next round by carrying out the methods development necessary to use "specialized analysis of risk…" Is this important enough for CASAC to include such a request in our letter to the administrator?

Page 2.12, line 13-20. The full set of model choices offered seems appropriate, given the plan is to maintain an annual as well as a 24 hour standard..

Chapter 3
Page 3.19-3.20 and Charge Question 1a and 1b. Use of peak shaving rollback approach as discussed is confusing. Staff indicates (at top of page 3.20) that because of time constraints they did not calculate health risks using this method. They go on to indicate what the method does. The appendix that discusses the method more fully refers to Tables F-49 and F-50. These tables present the curious finding that except for Fresno (where there is virtually no change) and Tacoma, all of the other cities show a progressive increase in maximum values going from the "proportional" to "hybrid" to "shaving". This certainly does not seem to indicate that the

shaving analyses were not done. I t also raises for me concern that more effort is needed to understand which is a better technique or more representative of "truth". Thus I am worried that "time constraints may be once again getting in the way of what the proper analysis should be.

Section 3.3., Charge Question 2 asks about the discussion of rationale for modeling choices and selection of CR function. The section follows a logical rationale and provides appropriate documentation of both model and site selection.

Page 3.49, Section 3.4.1.3. It might be worth adding a sentence that approximately 52 x 10^6 people are represented in the 15 cities or about 20% of the total population of US (even if not a representative sample).

Page 3.53-54, Table 3.11. there is far too much missingness in this table to make it at all useful. It raises more questions about the quality of the health input than would be justified. The lack of COPD for all sites but LA just doesn't compute for the 4th major cause of death in the US. Why are there not values for All Ages for All Causes? The data are presented for some but not all categories of disease. Someone needs to take the time to get all these cells filled in. These data all exist at the Federal level in one book o(or on one web site)!

Charge Question 3, Uncertainty and Variability. The discussion is quite complete dealing with a number of issues. What is not fully considered is the role of the potential for non-random missingness in both site selection within the 15 cities and thus the selection of the cities themselves.

Chapter 4
Charge Question 4-Sensitivity Analysis
Section 4.3, Table 4.1, page 4.6. This table needs to define denominators for Incidence (in title). Assuming both are the same then in general hospitalization considerably more frequent than mortality. (My concern is mortality may be x/100,000 and morbidity might be y/10,000 and if this is the case there are strikingly different numbers that need to be discussed more fully).

Page 4.17, line 6. Need to indicate denominators for these incidence rates.

Page4.46 Figures 4.7 &4.8. I have difficulty in interpreting these figures. For HD total incidence for LA go from 5-10% to as high as 19%. For Philadelphia from10-15% to14-15% with outlier at 23%. For total mortality the variation are separately 2-6% and 3-4 to 8%. For both these cases the core analysis seems too high and the delta seems too big, as tests of sensitivity. Is this the wrong interpretation?

Page 4.47, line 29. Should this be figures 4.7 and 4.8 rather than 4.6 and 4.7?
Page 4.54, line 5. This should be changed to Table 4.7 (page 4.56).
Line 21. Figure numbers seem not to match up with text being discussed. Please check.

Page 4.67-4.69, Figures 4.17-4.19. These are very effective presentations of design values. I suggest either change titles or text numbers so that 24 hour/annual are presented consistently in both.

Charge Question 5a. Design values analysis and assessments. These analyses are quite useful, as they are described. They provide visual interpretations directly of the range of effects related to the alternative standards. Notably in Chapter 3 the suggestion is made tht the use of peak rollback shaving is not to be considered; however, it looks like it is being used here.

Charge Question 5b. Additional suggestions. I think the ranges of outcomes are well demonstrated in figures 4.20 and 4.21. Rather than simply ending with the description of how they were constructed it might be worth adding a paragraph or two that summarize the degree to which the various alternative would provide changes in some fraction or all of the cities, rather than just ending with selected examples (maybe this will happen in Chapter 6).

Chapter 5

This chapter leaves me uneasy and I would like to suggest we spend some time discussing it. It look to me that it is a straight forward extrapolation of results from the two large cohort studies (ACS and 6 Cities). The level of exposure are extracted from the 15 urban areas and then scaled up to the US. This leads to ~88,000 (4-8% of total mortality—a figure repeated in Chapter 6). However, this seems too high. At one point there is a suggestion that the figures come from upper end of urban mortality risk, but isn't the calculation for the whole nation where exposure levels must be a lot lower? Lets discuss!

Chapter 6
Page 6;.7, line 25. Typo change "if "to "of"

General Comment: The integrated discussion pulls together and highlights some of the specific details presented in Chapter 4 and the Appendices. It reflects the obvious and not so obvious conclusions that results from manipulating the various alternative scenarios between current existing exposure, estimates to current NAAQS levels and the various alternatives proposed. In spite of the evidence that there does not appear to be a threshold the lowest level assessed 12/25 seems appropriate in that the evidence of an effect below those levels is simply too uncertain to evaluate. On the other hand the question of margin of safety remains and it will need to be argued that any level chosen (above that level) will need to be defended as to whether there is adequate margin of safety.

With regard to the Charge Questions:

a. Staff has done an excellent job in presenting and capturing the key-policy relevant questions. However, as indicated early on in this document there was to be a qualitative discussion on PM10-2.5 and on those effects that were deemed only "suggestive" but might have important public health implications (e.g. lung cancer, reproductive effects), but for which quantitative risk assessment was not thought warranted but that would appear in the PA (page 2.6). I would have thought that some remarks in this chapter would be necessary to assure that the PA would discuss the issues.

b. Key observations are presented and discussed with adequate discussion of the relevant contribution of the role of annual and 24 hour design values and the role of "peakiness" of distributions. One observation that appears to be focused upon and may be a driving

force is the uneven distribution among the 15 urban sites on the impact of the various scenarios and whether this fact is sufficiently taken into account in scaling up for the national estimates. More discussion and or analyses on this point may be warranted. For example what role does the actual estimates from these 15 sites play is coming to the estimates of 3-9% excess mortality? It may be too much to expect (in spite of the statistics) that 63,000-88,000 premature deaths would be prevented. Part of the country is already well below the proposed alternative levels and thus would not contribute to lives saved. Are there additional alternative sensitivity analyses that would provide either alternative estimates or put more confidence in these estimates by taking into account better population weighted C-R analyses?

c. See above. In spite of the last comment, the uncertainties and variability of the core assessments seems to be as good as it can be.

d. Evaluation of the several national scale analyses, as indicated above is of some concern. If I read the Tables in Appendix E correctly, the effect of moving to the lowest alternative (25/12) in some cases within the 15 urban sites produces a range of 32-67% (with one outlier at 11% and one at 100%) reduction in the IHD compared to the current standard. The question is, is this the best baseline for the comparison or should it be the current recent measurements, which would drop the percent changes considerably (and perhaps provide a more realistic estimated of the potential benefits from implementing changes). Obviously, the proportional ranking and changes would be the same, but the impact on "lives saved" on a national scale might be considerably less and more realistic.

e. Key observations seem to be presented in a balanced and fair way. Although the national assessment suggests a range of 63,000-88,000 premature deaths per year attributable to PM2.5 does not jive with a fairly often quoted figure from 2006 that moving the annual standard from 15 to 14 ppm would result is "more lives than perished in 9/11". (That figure translated into about 3000 lives.) Staff acknowledges that the rage of effects are in two categories: 3-9% and 0-3% in two halves of the country. This may be less precise than what the data indicate in that it would appear from their own estimates that the bulk of the effect comes fro the upper end of the exposure in the counties (pg 5.8, line 12-15).

Comments from Dr. Helen Suh:

Comments to Charge Questions 2 and 3
***Quantitative Health Risk Assessment for Particulate Matter*: Second External Review Draft**

Helen H. Suh
Associate Professor
Harvard School of Public Health
Boston, MA 02215

March 2010

The Second Draft of the Risk Assessment is a clear and comprehensive presentation of the rationale, methods and results for the assessment the acute and chronic PM-mediated health risks. In this draft, the authors have thoughtfully considered and incorporated the comments and suggestions from the CASAC panel. As an overall comment, the rationale that was used in the decision to forego a risk assessment for $PM_{10-2.5}$ made logical sense. Further, inclusion of a discussion of $PM_{10-2.5}$ exposure and risk in the upcoming draft Policy Assessment document is welcomed. However, it still seems that a qualitative and indirect assessment of $PM_{10-2.5}$ risks could be made based on PM_{10} monitoring and health data for metropolitan areas where $PM_{10-2.5}$ concentrations comprise a large fraction of PM_{10}. While certainly not ideal, this qualitative assessment may provide valuable information about $PM_{10-2.5}$ risks.

***Charge Question 2**: Selection of model inputs (section 3.3): We have expanded and clarified the discussion of our rationale for identifying modeling choices comprising the core risk model, focusing in particular on selection of C-R functions (section 3.3.3). To what extent does the Panel consider this discussion to be clear and the model selections appropriate?*

The revisions to section 3.3 represent a significant improvement over the previous draft, with the selection rationale for the core risk model clearly and cogently presented and previous concerns addressed. The rationale provided is a thoughtful and sensible approach to assess particle-mediated health risks. Further, the summary tables (Table 3.5 - 3.8) provide a useful and nice synopsis of the model inputs for the core risk models and sensitivity analyses. As a very minor comment, it might be possible to condense Table 3.7 somewhat by replacing certain columns with check boxes instead of text (for example to indicate short-term or long-term).

***Charge Question 3:** Addressing uncertainty and variability (section 3.5): We have clarified the process used to evaluate sources of variability and added coverage for specific sources of variability (section 3.5.2); expanded our discussion of the qualitative analysis of uncertainty (section 3.5.3); and included analyses of pair-wise interactions of sources of uncertainty (section 3.5.4). To what extent does the Panel consider these discussions to be clear and appropriate?*

The discussions of uncertainty and variability were clear and thoughtful, representing a substantial improvement over the previous draft. The sources of variability and uncertainty were well represented, although some explicit discussion of the impact of seasonality on risks should be included given its effects on each of the other discussed sources.

Comments from Dr. Sverre Vedal:

PM Risk Assessment, Charge Question 6.

General:

1. There is a too much formulaic and detailed presentation of results, for example when presenting findings of just meeting the current standards and just meeting alternative standards. This seems out of place in an integrative discussion.
2. While there is some opinion that effects of long-term exposure on IHD mortality (as opposed to all-cause cardiovascular mortality) are especially strong, this is based on relatively weak evidence from the ACS cohort (Pope 2004), in my opinion. It is therefore not clear why it was elected to present risk assessment results largely for IHD in this discussion (p. 6-6) and in chapter 4 of the RA. On second thought, it is clear because all-cause CVD mortality effects were not presented in Krewski 2009, unfortunately, and Krewski 2009 was chosen to provide the most defensible effect estimates.
3. The remaining percent PM-attributable effect of long-term exposure on <u>total</u> mortality is presented (p. 6-7, line 18; p. 6-8, line 8 and lines 14 & 15). These are ridiculously high. Presumably this should be on <u>IHD</u> mortality. The first bullet under the first Key Observation (p. 6-18) gets it right, I believe.
4. I question whether short-term exposure-related risk (p. 6-5, line 3) is also driven by changes in long-term average PM concentrations. Short-term effects are observed independent of long-term PM concentration, ie, down to the lowest baseline concentrations. Therefore the argument in this paragraph for motivating attention to the annual average is not sound.

6.a. Key questions captured?

Yes

6.b. Role of additional factors?

Observations in 6.2 are well-supported. No other observations are apparent.

6.c. Characterization of confidence.

Again (see above), the emphasis on IHD mortality as opposed to all-cause cardiovascular mortality reduces our confidence in effect estimates; unfortunately, all-cause cardiovascular effect estimates are not provided in Krewski 2009. IHD effect estimates are highest in the ACS cohort. Otherwise, estimates are conservative.

6.d. National context.

I'm not sure I agree that the 15 cities capture "the overall distribution of risk for the nation," (p.6-16, line 36) given the fact that these are the more polluted cities in the nation. More correct is the contention that these cities reflect the experience of cities with relatively elevated levels of PM-related risk and attributable mortality. Both can't be true.

6.e. Key observations.

These observations are fine. I see little utility in estimating risk below an annual standard of 12 mcg/m^3.