UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

December 16, 2008

EPA-CASAC-09-003

The Honorable Stephen L. Johnson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: Clean Air Scientific Advisory Committee's (CASAC) Review Comments on EPA's *Risk and Exposure Assessment to Support the Review of the* NO_2 *Primary National Ambient Air Quality Standard*

Dear Administrator Johnson:

The CASAC Oxides of Nitrogen Primary National Ambient Air Quality Standards (NAAQS) Review Panel (see Enclosure A for Panel Roster) is providing review comments on the Environmental Protection Agency's (EPA) final *Risk and Exposure Assessment (REA) to Support the Review of the* NO_2 *Primary National Ambient Air Quality Standard.* We focus on Chapter 10, which summarizes the findings in the document and discusses options for revisions of the NAAQS. This is the fourth advisory letter that CASAC has provided on the NO_2 REA.[1]

CASAC decided to provide review comments on this REA, even though EPA has finalized the document to meet the court-ordered schedule for the NAAQS review, because this draft represents the first instance of completing an REA under the revised approach for NAAQS review. We understand that CASAC's comments will be given consideration as EPA moves forward to develop the NO_2 primary NAAQS Advance Notice of Proposed Rulemaking (ANPR) for publication in January 2009.[2] Because this review was initiated by CASAC, EPA did not provide specific charge questions. The CASAC panel held a public teleconference on December 5, 2008 to review the completed REA.

[1] CASAC provided a letter on June 25, 2008 (EPA-CASAC-08-014) providing peer review of the first draft REA; a letter on September 24, 2008 (EPA-CASAC-08-021) providing peer review of a second draft of the REA that did not include Chapter 8 on exposure or the summary Chapter 10; and a letter on October 28, 2008 (EPA-CASAC-09-001) providing peer review of the REA draft Chapter 8.
[2] CASAC also notes that it is providing advice and review on the secondary NAAQS for oxides of nitrogen and that the setting of the primary NAAQS has implications for the related secondary standard, to be published later in 2009.

Overall, CASAC found this version of the REA satisfactory in its approach to moving from the scientific foundation developed in the Integrated Science Assessment (ISA) to setting out evidence-based options for the NAAQS. The REA provides the needed bridge from the evidence presented in the ISA to a characterization of the exposures and the associated risks with different profiles of exposure. It draws on toxicological and epidemiological evidence and addresses risk to an identified susceptible population, people with asthmatic conditions. EPA has also systematically described uncertainties associated with the risk assessments. We commend EPA for developing a succinct and thoughtfully developed synthesis in Chapter 10. This summary chapter represents a long-needed and transparent model for linking a substantial body of scientific evidence to the four elements of the NAAQS.

This letter provides the panel's comments concerning the most critical aspect of Chapter 10, its recommendations concerning the four elements of the NAAQS for NO_2. We also provide general comments on other major issues identified by the panel in its review. CASAC also prepared a set of more specific panel comments, included as Enclosure B with this letter. Enclosure C provides comments of individual panel members.

We turn first to the discussion of the NAAQS in Chapter 10. CASAC concurs with EPA's judgment that the current NAAQS does not protect the public's health and that it should be revised. With regard to the options for revisions to the four elements of the NAAQS, we recommend, as follows:

- **Indicator:** The panel concurs with retention of NO_2 as the indicator.

- **Averaging time:** The panel concurs with having a short-term NAAQS primary standard for oxides of nitrogen and using the one-hour maximum NO_2 value. CASAC also recommends retaining the current standard based on the annual average. It bases this recommendation on the limited evidence related to potential long-term effects of NO_2 exposure and the lack of strong evidence of no effect. Additionally, the findings of the REA do not provide assurance that a short-term standard based on the one-hour maximum will necessarily protect the population from long-term exposures at levels potentially leading to adverse health effects.

- **Ambient air concentrations:** For the short-term standard, EPA presented two sets of recommendations: from 50 to 100 ppb and from 50 to 200 ppb NO_2. The evidence reviewed in the REA indicates that adverse health effects have been documented in clinical studies of persons with asthma at 100 ppb and the REA finds "…strong support for a level at or below 100 ppb" (page 309). CASAC firmly recommends that the upper end of the range not exceed 100 ppb, given the findings of the REA. For the annual standard, the panel recommends retaining the current level, as evidence has not been cited that would lead to either an increase or decrease.

- **Statistical form:** The selection of the percentile depends on the ambient air concentration chosen for the NAAQS. CASAC advises that EPA choose a health protective percentile appropriate for the level chosen for the one-hour standard. For example, if the lower bound of 50 ppb were recommended as the level, the 98th percentile would be appropriate. If EPA were to decide that the lower bound should be a higher number, then the Agency would need to consider a percentile higher than 98 as the form.

The REA makes some general comments about the role of scientific evidence and policy judgments in revising the NAAQS. CASAC notes that the NAAQS are based on the scientific evidence, and that policy judgments addressing gaps and uncertainties in the evidence base are within the purview of the REA as it offers options for the NAAQS. With regard to characterization of uncertainty, the REA provides a relatively comprehensive but qualitative description of the major uncertainties. We encourage continued evolution of the characterization of uncertainty to include some quantitative analyses as well, particularly for the sources of uncertainty that are found to be major. Any biases arising from the available data or assumptions made to bridge uncertainties also need explicit consideration. The REA leaves open how the uncertainties should be considered in promulgating the NAAQS. CASAC offers the reminder that the intent of the Clean Air Act is to protect public health with an adequate margin of safety and consequently uncertainty should be considered as a reason to move towards the lower end of the range of levels and not to the upper.

Second, we turn to the NAAQS review process in general, which will also be the focus of a forthcoming CASAC advisory letter. CASAC is troubled by the implications of the greatly accelerated NAAQS schedule in the NO_2 review for CASAC's fulfilling its statutory advisory and review role. It is especially important that, in the future, sufficient time be included in the document development schedule so that CASAC can receive the completed REA for review and comment before it is finalized by the EPA. CASAC also notes that some issues critical to public health protection relating to monitoring and implementation will be addressed in the January 2009 ANPR prior to CASAC review. One important issue concerns the siting of monitors in a manner that recognizes the spatial variability of NO_2 levels.

In closing, the panel noted that the final REA has satisfied the need for a document that can link the scientific foundation for the NAAQS review, provided in the ISA, to future presentation of evidence-based options for the NAAQS, which will be provided in EPA's ANPR, now scheduled for publication in January 2009. CASAC recommends that the remaining REA technical issues identified in the panel comments in the enclosure to this letter be addressed in the ANPR, so that the scientific evidence in the supporting the options can be clearly and accurately outlined in the ANPR. We look forward to reviewing the ANPR early next year.

Sincerely,

/Signed/

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

Enclosures

cc: Marcus Peacock, Deputy Administrator

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

U.S. Environmental Protection Agency
Clean Air Scientific Advisory Committee
Oxides of Nitrogen Primary NAAQS Review Panel

CHAIR

Dr. Jonathan M. Samet, Professor and Chair, Department of Preventive Medicine, Keck School of 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, USA

Dr. Donna Kenski, Data Analyst, Lake Michigan Air Directors Consortium, Des Plaines, IL

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

CONSULTANTS

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

Dr. John R. Balmes, Professor, Department of Medicine, Division of Occupational and Environmental Medicine, University of California, San Francisco, CA

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

Dr. Terry Gordon, Professor, Environmental Medicine, NYU School of Medicine, Tuxedo, NY

Dr. Dale Hattis, Research Professor, Center for Technology, Environment, and Development, George Perkins Marsh Institute, Clark University, Worcester, MA

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

Dr. Patrick Kinney, Associate Professor, Department of Environmental Health Sciences, Mailman School of Public Health , Columbia University, New York, NY

Dr. Steven Kleeberger, Professor, Lab Chief, Laboratory of Respiratory Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC

Dr. Timothy V. Larson, Professor, Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA

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

Dr. Edward Postlethwait, Professor and Chair, Department of Environmental Health Sciences, School of Public Health, University of Alabama at Birmingham, Birmingham, AL

Dr. Richard Schlesinger, Associate Dean, Department of Biology, Dyson College, Pace University, New York, NY

Dr. Christian Seigneur, Director, Atmospheric Environment Center, Université Paris-Est, Champs-sur-Marne, France

Dr. Elizabeth A. (Lianne) Sheppard, Research Professor, Biostatistics and Environmental & Occupational Health Sciences, Public Health and Community Medicine, University of Washington, Seattle, WA

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

Dr. George Thurston, Professor, Environmental Medicine, NYU School of Medicine, New York University, Tuxedo, NY

Dr. James Ultman, Professor, Chemical Engineering, Bioengineering Program, Pennsylvania State University, University Park, PA

Dr. Ronald Wyzga, Technical Executive, Air Quality Health and Risk, Electric Power Research Institute, Palo Alto, CA

SCIENCE ADVISORY BOARD STAFF
Dr. Angela Nugent, Designated Federal Officer

Enclosure B:
Additional Panel Comments

This enclosure provides more detailed comments on aspects of the REA, particularly those chapters that were modified since the prior peer review along with Chapter 10. In evaluating EPA's *Risk and Exposure Assessment to Support the Review of the NO_2 Primary National Ambient Air Quality Standard* (REA),the CASAC panel considered the extent to which chapters covered in previous reviews had been modified. A review of Chapters 1 to 6 in the final draft indicates that they have not changed appreciably in any way that would affect the conclusions in Chapter 10. As a result, the current review does not provide comments on Chapters 1 through 6, except to the extent they influence our assessment of conclusions drawn in later chapters.

As to Chapters 7 and 8, our review focused initially on whether the final draft contains any significant differences in numerical results that affect applications in Chapter 10. We find that the final draft does contain substantial revisions to the material presented originally in Chapters 7 and 8, at least in the quantitative magnitude of the size of the population affected by exceedances if not necessarily in the qualitative significance of the results as reflected in Chapter 10. These differences are due in part to methodological changes, and in part to the fact that the final draft report contains estimates of daily maximum exceedances while the previous draft reported total number of one-hr exceedances (so various numbers of exceedances in a day counted an equivalent number of times in calculating the average).

With regard to the analysis of exceedances in Atlanta alone (the second approach), the methodology used is close to that of the earlier draft. As a result, the estimates of key figures such as number of asthmatic-days exceeding a health benchmark are similar in the two documents. This similarity can be seen, for example, by comparing Figure 8-16 in the final draft against Figure 8-15 in an earlier draft, where the figures of the latter are within 20% of those in the former (the final draft values being lower than those of the earlier draft, resulting from corrections to the methodology suggested in previous CASAC reviews of the REA). These differences do not substantively affect the findings of neither the analysis nor the scientific foundation for any of the conclusions in Chapter 10.

We note first that the recommendations in Chapter 10 arise from three different analyses performed in Chapters 7, 8 and 9. These are specifically:

- an approach based on air monitoring results and health benchmark exceedances in 18 cities throughout the U.S. (Chapter 7);
- (ii) an approach based on more detailed exposure and health benchmark exceedance results in Atlanta (Chapter 8); and
- (iii) an epidemiologically and exposure-response approach, based again in Atlanta (Chapter 9).

We focus our attention on the methods and results of Chapters 9 and 10; the fidelity with which the results of each approach are summarized for Chapter 10; and how well the three approaches are integrated in drawing conclusions in Chapter 10, including the impact of uncertainty on the answers produced through the three approaches and the degree to which final recommendations for the NAAQS are bounded by these three sets of analyses.

The analysis of incidence of emergency room visits in Chapter 9 employs a methodology identical in both reports and consequently the incidence results are the same. The EPA has chosen in the final draft to summarize the results as percent of the incidence attributable to NO_2 exposures, rather than as total incidence, which provides a more easily understandable summary conclusion as to the magnitude of health effects. However, it would be useful to see both total incidence and percent incidence in adjacent tables of the final draft, as both provide information of value in decision-making. This could be accomplished by bringing a few of the relevant tables, such as Table 4-1, from the Appendices back up into the body of the final draft, especially since the text in the final draft mentions the incidence numbers but the reader must go to the Appendices to find them.

In addition, the results on incidence are summarized as average daily daily number of incidents, averaged over a year. This approach masks the seasonality of incidence, which can vary significantly throughout the year. There may be policy implications of such seasonal variations due to the capacity of the health care system to cope with periods of high incidence, and there may also be implications for the uncertainty in risk estimates. At the least, the REA should mention these issues and their potential implications for policy decisions.

Uncertainty analyses have improved across the sequential revisions of the REA. While it still is not possible to place rigorous and quantitative uncertainty distributions around the key conclusions, the results of Chapters 7, 8 and 9, and therefore of Chapter 10 which relies on these chapters, now include at least substantial qualitative analyses of uncertainty. The major sources of uncertainty have been identified in all three of these chapters, and some assessments made of their sources and implications. In Chapter 9 (the epidemiological approach), the EPA has also developed a reasonable system of assigning measures of the degree of uncertainty (low/medium/high); and the direction of bias (if any) introduced by a specific source of uncertainty has been noted. This characterization of each of the sources of uncertainty is useful and can be used for an overall and full assessment of uncertainty. However, there insufficient discussion of the basis for the specific classifications of the subjective degree of uncertainty. CASAC consequently cannot comment on whether these specific judgments are reasonable. We support the characterization of uncertainty but recommend that a fully transparent approach be used.

There is a difference in the approaches of the uncertainty analyses performed in Chapters 7, 8 and 9. The useful format employed in Chapter 9 is more qualitative, but summarizes the results in a way that is most relevant for subsequent decisions. By contrast, Chapters 7 and 8 contain more quantitative analyses of uncertainty, although the degree of quantification varies significantly from factor-to-factor. The result is a somewhat unbalanced and disjointed exploration of the "space" of uncertainty, for example with much discussion of uncertainty due to fitting a distribution to data but more cursory exploration of the effects of uncertainty in developing health benchmarks. Chapters 7 and 8 should have ended with tables similar to the one on uncertainty in Chapter 9.

It is unlikely, however, that further development of these procedures is needed to serve the purpose of the REA, or that a fully quantitative, Monte-Carlo-style, uncertainty analysis is feasible given the time available for the assessment process. In any event, some of the results, such as those in Chapter 9, have had quantitative confidence intervals placed around the estimates of incidence. These are admittedly conditional intervals (i.e., conditional on the choice of exposure-response model form and conditional on the air monitoring results used), but they are useful nonetheless and should be informative in decisions. In any event, we are pleased that the EPA authors have addressed the previous comments of CASAC in regards to taking uncertainty into account in developing this most recent REA.

The choice of the Tolbert/Peel epidemiological study is reasonable, given EPA's decisions to give increased weight to North American studies and to use Atlanta for a model exposure analysis. In general, the REA uses and reports the results of the Tolbert at al. analyses appropriately. There is some discussion about the use of average baseline rates and some potential bias due to the potential influence of NO_2 exposures on actual incidence; this discussion could be extended to consider the seasonal patterns of both ED visits and NO_2 in Atlanta. The REA correctly discusses the pros and cons of using single and multi-pollutant models in several places; however, on p. 273, this discussion is not as well-balanced as elsewhere. In Chapter 10 (p. 289-90; 308), emphasis is given to "central estimates" without mention of the uncertainty of these "central estimates". This description needs to be modified.

EPA has applied the concentration-response function from the Tolbert study appropriately to Atlanta ambient data. The analysis in Chapter 9 is, however, based on an implicit assumption that changes in ambient air concentrations associated with different NAAQS values will be approximately uniform in magnitude across the 5 counties of the Atlanta metropolitan area, or at least across the areas of most significant exposure. This assumption, commonly made in virtually all time series studies, is a natural consequence of using a monitoring station result as the measure of exposure for what is in reality a complex spatial pattern, and should be acknowledged at some point in the chapter or perhaps in Chapter 10. We understand that the analysis was limited by the way in which the epidemiological study characterized exposure through this monitoring station. If the spatial field of ambient air concentration is simply shifted up or down at all spatial points equally when a different NAAQS is introduced, this assumption will be valid. But where this uniform shift is not the case – and we are convinced it is not – the assumption will be less valid. The contribution of this problem to uncertainty in the results of Chapter 9, and to the implications for conclusions in Chapter 10, is neither clear nor discussed. Nonetheless, the epidemiology-based approach provides additional information that supports EPA's evaluation that a more protective short-term standard is needed to protect public health. And again, as in uncertainty analyses, the EPA authors have been responsive to previous reviews of this REA, using a study that is of acceptable quality.

In considering further the choice of the study by Tolbert et al. (2007) – which is an update of Peel et al. (2005) - we note that several epidemiological studies have examined the associations between ambient air concentrations of NO_2 and respiratory health endpoints; the results of these studies have been mixed with a large majority of them finding a positive association. These results are buttressed by findings of several human clinical studies, which under controlled exposure conditions, find an association between NO_2 exposure and respiratory symptoms and lung function responses. The authors chose to consider studies conducted in the US and to consider a health endpoint that was unambiguously adverse, ED admissions. Six key studies were identified in the ISA; of the six, four considered multi-pollutant models, an important consideration given the relatively high correlations between ambient levels of NO_2 and other pollutants and the concern that NO_2 could be serving as a surrogate for other pollutants. Two of these four studies were undertaken in Atlanta, and two in New York City. The two Atlanta studies are based in the same study with the second updating the first with additional years of data. The REA analysis could have used any combination of these four studies; indeed it would be useful to consider some additional analyses based upon all four studies as a sensitivity analysis. Given the variation of their results, this would likely demonstrate a range of responses, at least for asthma, the only endpoint that the studies have in common.

A key issue underlying the conclusions in Chapter 10 is whether Atlanta is to be taken as representative of exposures and effects in other moderately large urban areas. The EPA justification for Atlanta is found in Chapter 8, which reads "…Atlanta is roughly in the middle of the distribution with respect to estimated population and total roadway miles when compared with the other locations examined. Given that there are factors that would lead to either under- or over-estimation with respect to the influence of on-road and other mobile sources on the representativeness of Atlanta relative to the other 17 urban areas examined in the air quality analysis, EPA judges that the Atlanta exposure estimates are likely representative of other moderate to large urban areas included in this comparison. EPA does recognize that the Atlanta exposure results are likely lower on a population-weighted basis compared to the largest urban areas such as Los Angeles, New York, and Chicago given the greater proximity of the population to mobile sources in these large urban areas."

Chapter 10 contains a direct response to a key policy question that was the basis for the existing NAAQS. The existing NAAQS was established based on the claim that "Retaining the existing standard would also provide protection against short-term peak NO_2 concentrations at the levels associated with mild changes in pulmonary function and airway responsiveness observed in controlled human studies" (60 FR 52874, 52880 (Oct. 11, 1995)). CASAC supports the conclusion of the REA that health effects studies conducted since implementation of the existing NAAQS, and the results of the current REA under review, suggest that the existing standard, both in terms of level (0.053 ppm) and averaging period (annual) is not protective of human health in regards to short-term peak NO_2 concentrations. This conclusion is supported uniformly by the three approaches employed in the REA. In this regard, we do not fully agree with the statement on Page 284 that "(t)he way in which exposure and risk results will inform ultimate decisions regarding the NO_2 standard will depend upon the weight placed on each of the analyses when uncertainties associated with those analyses are taken into consideration."

On Page 286, the authors note competing uncertainties, or biases. The first is that AERMOD may over-predict upper percentile exposures, and the second is that the

Atlanta case study may under-predict near-road exposures, both in Atlanta as well as in cities where near-road populations are more common. It would be helpful if the authors were to give a summary statement as to the degree to which this introduces significant uncertainty into the final conclusions of the REA. Members of CASAC have differing opinions on the degree to which one or the other of these biases dominate and some members believe that these two sources of bias are likely to roughly cancel each other. Given the robustness of conclusions across the three approaches, these biases may not be the primary sources of uncertainty in drawing policy conclusions.

The authors also note that "(h)owever, to the extent that a decision regarding standard level emphasizes the general uncertainties associated with quantifying the contributions of NO_2 to respiratory effects in epidemiologic studies and uncertainties regarding the public health significance of NO_2-associated airway hyperresponsiveness (particularly at 0.1 ppm), a level as high as 0.2 could be supported." We note that the treatment of uncertainty in the causal connections between NO_2 exposures and effects is not quantified, and so it is not possible to use uncertainty as the basis for a specific magnitude of adjustment to the upper limit on levels to be considered (here, by a factor of 2). The lack of guidance provided in the document on how the various sources of uncertainty and the associated magnitude of uncertainty influence any specific conclusions as to exposure and risk leaves open as to how the uncertainty analyses should figure in their use in policy decisions. We encourage the REA authors to provide more guidance as to conclusions in Chapter 10 that are meaningfully uncertain, and the direction of this uncertainty (whether it is likely to be for or against health protectiveness, or neutral).

A further key issue is whether the Atlanta study results (Chapters 8 and 9) can be extrapolated elsewhere, or at least be considered as representative for a larger U.S. population. The ISA shows that results can vary across studies. Indeed a consideration of the above-cited studies for asthma ED visits and NO_2 shows variable results. For this reason it would have been preferable for the risk assessment to consider additional studies as well; unfortunately the number of studies that consider NO_2 in both a single-pollutant and a multi-pollutant context is limited; nevertheless some consideration of the New York studies could be helpful and considered as a sensitivity/uncertainty analysis. Secondly the results of a specific study are tied to the characteristics of the study and the location in which the study is conducted. The REA points out several of these characteristics, but it should also show that the use of EDs varies geographically. For some populations, EDs are a major source of healthcare, and it is unclear how Atlanta may differ from other metropolitan areas in this regard.

There is a fundamental issue, however, that is not discussed sufficiently in this chapter; namely, the issue of differences between exposures and monitored concentrations of NO_2. The REA (and the ISA) notes that exposures are likely to be significantly higher near roadways, and monitors may not capture these high exposures. The limited discussion of this issue on p. 300 is based upon comparisons of personal exposures and ambient monitoring results for extended periods of time (>24 hours). For the much shorter period of one-hour, these comparisons are of uncertain relevance. The highest exposures likely occur when individuals are near roadways, but personal-activity patterns would not lead to extended stays at such locations. Unfortunately there are no measurements of short-term personal NO_2 exposures that might better inform this issue. In the REA, adjustment is made to estimate peak NO_2 air concentrations on roadways from monitored levels. The need for this adjustment contradicts the statement made at the bottom of p. 300. There should be an expanded discussion of exposure issues in interpreting the results of the epidemiological analysis in the context of standard setting. The siting of regulatory monitors is a related issue and one that does not receive attention in Chapter 10, but needs to be considered in the ANPR.

A remaining problem with both Chapters 9 and 10 is inconsistency inconsideration of alternative 1-hour concentrations. EPA shifts between reporting them in ppm and ppb. Eliminating this inconsistency would improve the document and allow the reader to focus on conclusions rather than trying to understanding the basis for these differences. This inconsistency should addressed if possible.

Enclosure C: Compilation of Individual Panel Member Comments on EPA's completed *Risk and Exposure Assessment (REA) to Support the Review of the NO_2 Primary National Ambient Air Quality Standard*

This enclosure contains final written comments of individual members of the Clean Air Scientific Advisory Committee (CASAC) Oxides of Nitrogen Primary National Ambient Air Quality Standards (NAAQS) Review Panel. The comments are included here to provide both a full perspective and a range of individual views expressed by panel members during the review process. These comments do not represent the views of CASAC or the CASAC Panel.

<u>Comments Received</u>:

Comments from Prof. Ed Avol

Comments on Draft NO2 REA Document
(EPA-452/R-08-008a, November 2008)
Ed Avol

(Samet-proposed focus of CASAC Dec 5 teleconference)
Chapter 10 of the REA and the adequacy of the NOx ISA and NO2 REA for asking the question below:

What scientific evidence and/or scientific insights have been developed since the last NOx review to indicate if the current public-health based NAAQS need to be revised or if alternative levels, indicators, statistical forms, or averaging times of these standards are needed to protect public health with an adequate margin of safety ?

Chapter 10 should be a NO2 document "grand unification" of sorts, in which the salient conclusions and judgments of the several hundred pages of thoughtful presentation, appendices, tables, and analyses are presented in an accessible and succinct presentation. In general, this has been achieved, and staff should be commended for the accomplishment. The chapter is carefully organized, framing the discussion in the context of indicator, averaging time, form, and level of the standard. Clarity and logic of argument is provided across most of the four main categories of discussion.

The scientific insights and research evidence that have become available since the 1996 NOx review have been substantial. Several informative epidemiologic, and controlled-human-exposure studies, as well as a wealth of ambient monitoring exposure information, has become available in the decade since the previous review. The staff has carefully reviewed this data and made a rational and substantive case for the Administrator to consider revision of the current standard to protect public health.

Chapter Section 10.4 ("Potential Alternative Standards") is generally well-constructed, logically described, and clearly presented. The explanation for the choice of "Indicator" is clear, logical, and compelling. The presentation of "Averaging Time" considerations is informative and convincing. The discussion of "Form" of the standard helpfully provides several examples using several metropolitan areas and seems justified.

The recommendations of the chapter section concerning "Level" of the standard (Section 10.4.4), however, seemed inconsistent with the presented evidence. The staff's conclusion that the upper end of the range of a one-hour daily maximum standard reasonably supported by the evidence is 0.2ppm seems inconsistent with the prior conclusions that "…the scientific evidence provides strong support for a standard at or below 0.1ppm (100ppb)…" Where is the adequate margin of safety? Where is the acknowledgement and adjustment for sensitive sub-populations (such as more severe asthmatics) who were under-represented in the scientific studies? It appears that uncertainty concerns, with respect to measurement errors, health, and monitoring outcomes have been used to move the discussion in only one direction (towards a less-protective level). This is arguably the wrong direction, if the stated goal is to protect the public's health.

The final concluding section of the document (Section 10.4.4.3) appropriately lays out the range of NO2 exposure levels, which based on the current available evidence, would be protective, but this makes the conclusions of the evidence-based considerations section (10.4.4.1) seem out-of-step with the collective view of the chapter.

Comments from Dr. James Crapo

The revised REA document, including Chapter 10, is well done and provides an excellent review and assessment of the scientific data that underlie health risks of exposures to NO2 at ambient or near ambient levels in the United States. The agency is to be complemented for the scope, design and implementation of this document. It is a substantial advance over the equivalent documents in prior reviews in which I have participated over the past 5 years.

The critical data on risk for adverse human health effects at ambient/near ambient levels of NO2 are appropriately reviewed in Chapter 10 and I would concur with the primary recommendations for the NAAQS for NO_2.

The indicator is appropriately recommended to remain as NO_2. The scientific evidence now provides a sound basis for a new short-term standard for NO_2. I would concur with the staff recommendation for a 1-hour standard. There is little rigorous data to critique the choice of the form for the standard and I would concur with the selection of either the 98th or 99th percentile.

With regard to the level of a new 1-hour NO_2 standard, I find the data summarized in Chapter 10 convincing that the level should be no higher than 0.1 ppm. Controlled human exposure studies suggest that levels of NO_2 from 0.1 to 0.2 ppm would result in adverse symptoms in some asthmatics. I would thus strongly concur with the staff conclusion that the optimum level for the 1-hour NO2 standard would be in the range of 0.1 to 0.05 ppm.

I concur with the recommendations for modifications to Chapter 10 as summarized in Enclosure B with the CASAC letter which can be addressed in the upcoming ANPR. With these modifications for clarity and rigor, I compliment the staff on an excellent REA document and on recommendations for a new NO2 standard that will reflect the scientific data and be appropriately supportive of the public health.

Comments from Dr. Terry Gordon

Chapter 10 – NOx REA

Charge question: "What scientific evidence and/or scientific insights have been developed since the last NOx review to indicate if the current public-health based NAAQS need to be revised or if alternative levels, indicators, statistical forms, or averaging times of these standards are needed to protect public health with an adequate margin of safety?"

Overall, the chapter is an excellent concluding chapter for the NOx REA. The indicator, form, averaging time, and alternative levels sections are very clear and well organized. Based upon the logical pattern of the assessment of the data, it is not clear how the range is extended up to 0.2 ppm (or even warranted). Despite the chapter's logical flow and clarity for the suggested alternative standard, the Staff does not directly address what to do with the current standard. This is a key issue that the NAAQS development process is required to address.

Minor comments:

Page 277, para 3 – There seems to be some issue with plural/singular (suggests vs suggest and its vs. their support)
Page 277, last bullet – 'areas' is unclear (although I know it is meant for geographic region).
Page 279, para 1, line 6 – Were the human exposure and animals studies only qualitative?
Page 288, para 2, line 2 – Unclear sentence: 'effects that we cannot currently…' implies that you did assessments on effects that you couldn't quantify.
Page 290, para 1 – 'inclusion' in the 'inclusion' could be reworded.
Page 290, para 2, line 7 – The sentence is unclear: 'To the extent that these considerations are emphasized,'?
Page 292, para 1, line 5 – It is unclear what the 'air quality correlations' are correlating with.
Page 293, Table 10-1 –The table legend needs a little more information so that it can stand on its own (e.g., Max (99^{th}): Annual – what is Annual?)
Page 296, Table 10-3 – It would be more clear if '1-h' were added after 0.05, 0.1, etc. Also, 'alternative' was used in the title but 'different' was used in the table heading.
Page 297, para 3, 1^{st} sentence – It could be me, but the second half of the sentence isn't clear: 'could exceed the standard level and still attain the standard'?
Page 299, para 1 – The font size appears to be different.
Page 299, para 2, line 1 - The choice of 'greater protection' is odd. Are we trying for appropiate protection or greater protection than the current standard?
Page 309, para 1 – I realize that this is the conclusion paragraph for 'level', but it is odd reading 'level' with no mention of an averaging time (of 1 hour or anything else).

Comments from Dr. Dale Hattis

Premeeting Comments for Discussion of the Draft Final REA for NO2

Dale Hattis, Clark University

I am happy to report that there have been substantial efforts to address the issues I raised in my previous comments on the probabilistic methodology for APEX exposure modeling and the critical on road/off road concentration adjustment factor. In particular

- EPA staff in the revised Chapter 8 has done some sensitivity analysis for their distributional assumptions related to the indoor removal rate distributions. While I am not completely happy with the analysis, I think it reasonably suffices to reduce concerns that there is some large hidden uncertainty that could materially alter the results of the exposure analysis.

- EPA have now reasonably fully documented their database of empirical observations of Cv/Cb (or "m") that led to their assumption of the empirical distributions on-road/off-road sources of emissions for modeling of the NO2 concentrations on and near roadways (Section A-8.2 beginning on A-111 of Appendix A). Based on this documentation I have been able to critically evaluate EPA's conclusion that no parametric distribution can be found that reasonably describes the available data. Briefly, I think this is incorrect. However, using the documentation provided, I have developed a relatively simple alternative distributional suggestion that could at least be used for sensitivity analysis in the calculations for the eventual ANPR. This analysis is given briefly below:

The reason why it is desirable to avoid simple empirical distributions in the kind of exposure/risk analysis being done in the REA is that empirical distributions are necessarily limited to the range of values that have been directly observed in the available data. When there are only 41 available data points there is reason to suspect that this could effectively cut off the high tail of the distribution of actual concentrations; and this is unfortunate because the high end tail of the on road and near road concentrations determines the estimates of the frequency with which high end concentration benchmarks are exceeded. This concern is magnified in the current case because the 41 observations used for the analysis already reflect three exclusions of extreme values (see p. A-113)—two from the "not-summer" category and one from the "summer" category.

Figure 1 shows probability plots of the logarithms of the Cv/Cb ratios in the database that EPA has assembled and classified into its "summer" vs "not-summer" categories.* It can

* In this type of plot, the straight line represents a hypothesis that the data are lognormally distributed, and the correspondence of the points to the line is a quick qualitative indicator of the correspondence of the data to the theoretical distribution. The intercept and slope are estimates of the mean and standard deviations of the log-transformed values, respectively. Lognormal distributions are expected when there are many factors

be seen that the datapoints—particularly those from the "summer" distribution—are not perfectly described by the straight lines representing the lognormal hypothesis—but there is not a radical systematic departure either. Subdividing the "summer" points in to those that are listed as "downwind" measurements vs those classified as "both upwind and downwind" (Figure 2) locates the departure from lognormality mostly in the second category.

The two distributions shown in figure 1 have relatively similar slopes (standard deviations of the log-transformed values). Therefore there is reason to suspect that a simple regression analysis could capture both distributions by deriving a simple multiplicative factor to relate the distributions for the two seasonal categories. The regression I derived for this purpose is:

Response:	log(m)
Summary of Fit	
RSquare	0.152
RSquare Adj	0.130
Root Mean Square Error	0.227
Mean of Response	-0.091
Observations (or Sum Wgts)	41

Parameter Estimates

Term	Antilog(Est)	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.643	-0.192	0.052	-3.69	0.0007
Season2--Summer	1.539	0.187	0.071	2.64	0.012

It can be seen that the "summer" parameter is statistically significant at well under $P<.05$. To check on whether the regression with its root mean square error of .227 (analogous to the standard deviation of the data points about the regression line) can be used to describe the dispersion of the observations, I did a probability plot of the residuals** from the regression (Figure 3). It can be seen that the regression residuals are in fact well described as lognormal--having a mean very close to zero and a standard deviation within rounding error of the expected 0.227. Therefore I recommend that in place of its empirical distribution, EPA should substitute lognormal distributions for their "m" values with the following parameters:

	Gmean	GSD
Not Summer	0.643	1.685
Summer	0.989	1.685

This should be done at least for sensitivity analysis for the calculations in the final analysis for the ANPR.

---

that cause differences among individual observations, and those factors tend to act multiplicatively on the measured parameter.

** Each "residual" is the regression model predicted-value minus the observed value for an individual data point.

Figure 1

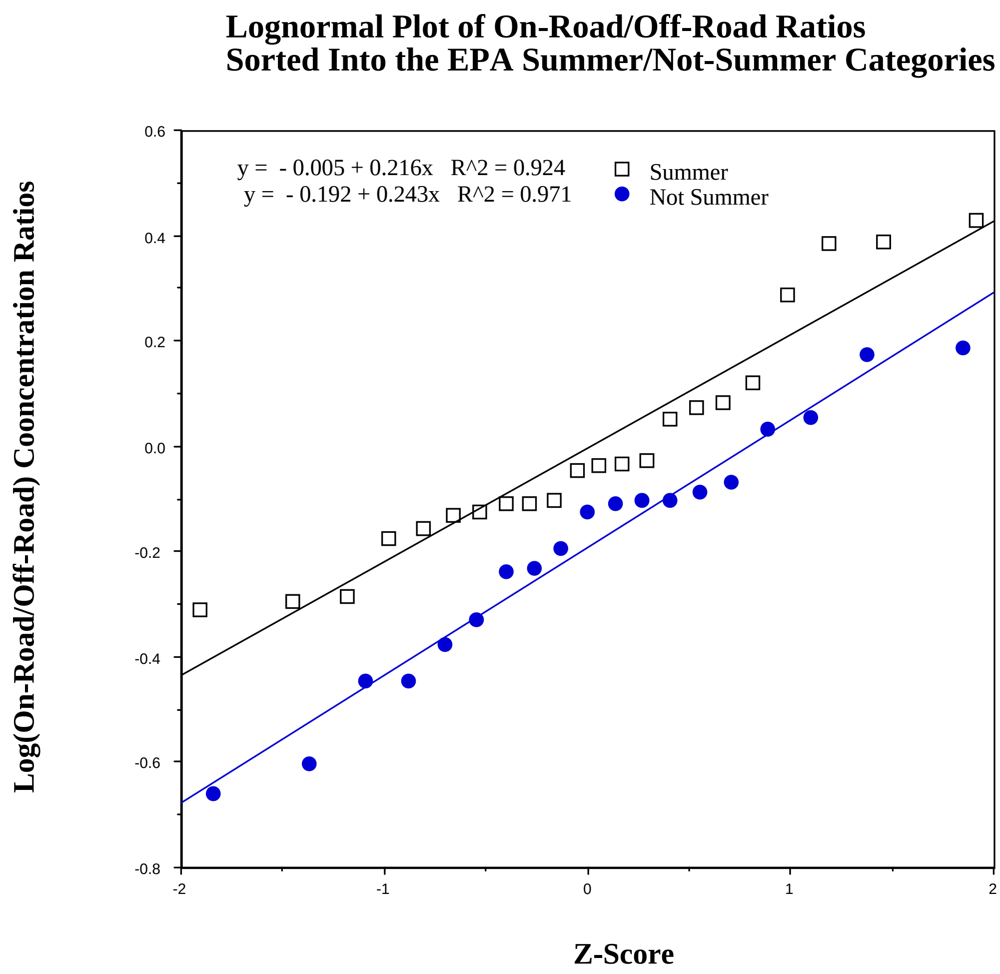
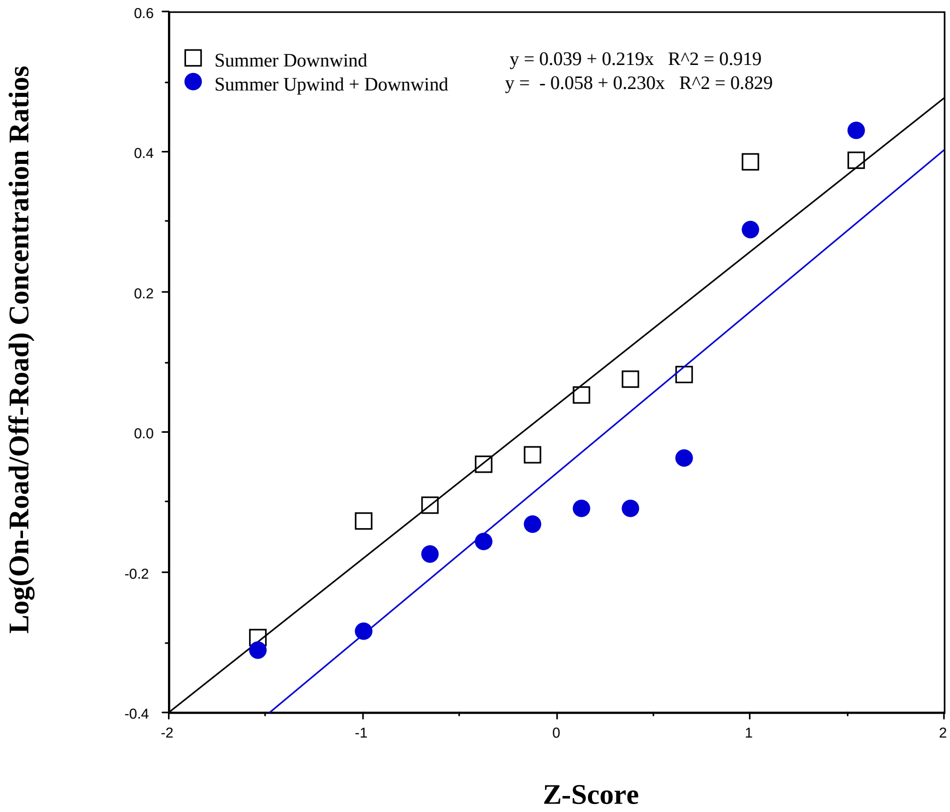
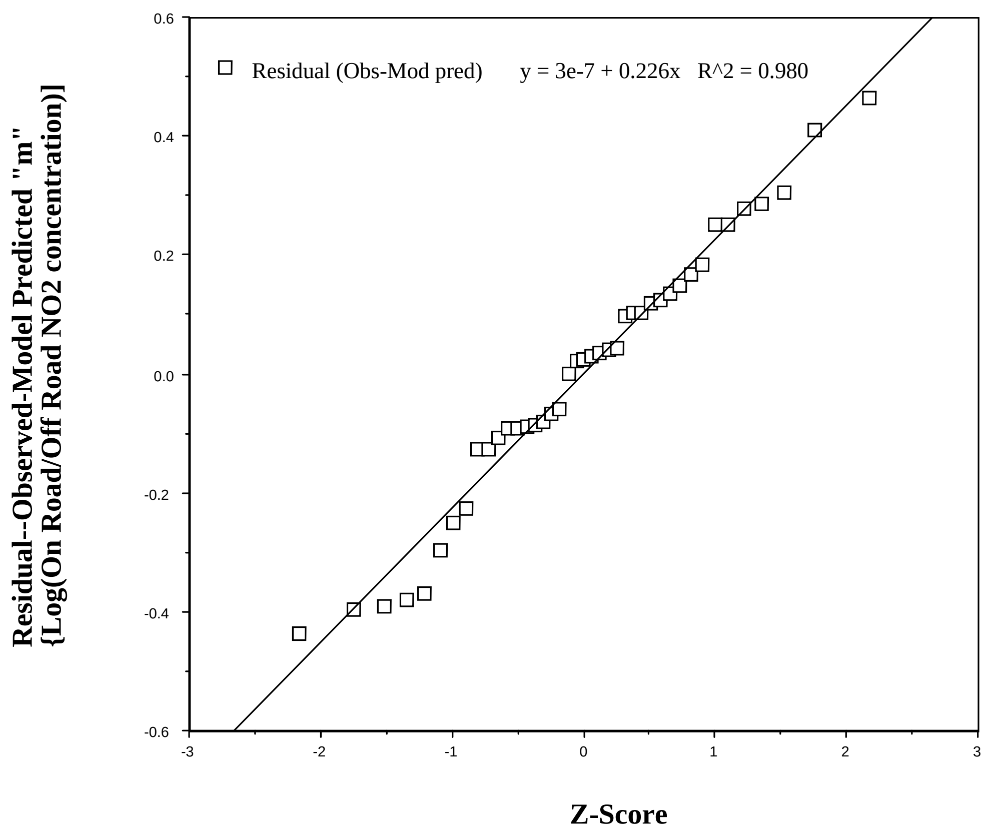

Lognormal Plot of On-Road/Off-Road Ratios
Sorted Into the EPA Summer/Not-Summer Categories
y = - 0.005 + 0.216x R^2 = 0.924
y = - 0.192 + 0.243x R^2 = 0.971
Summer
Not Summer
Log(On-Road/Off-Road) Cooncentration Ratios
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-2
-1
0
1
2
Z-Score

Figure 2

Distribution of Summer Log(On-Road/Off-Road) NO2 Concentration Ratios Subdivided by Those Classified as "Downwind" vs "Both Downwind and Upwind"



Probability Plot of Observed - Model Predicted Ratios of Log(On Road/Off Road) NO2 Concentrations



Comments from Dr. Rogene Henderson

Comments on final Risk and Exposure Assessment to Support the Review of the NO2 Primary National Ambient Air Quality Standard

I thought the document was well prepared and provided the information needed for rulemaking. I compliment the Agency on the clarity of their presentation.

My comments will focus on Chapter 10 and will be organized according to the indicator, form, averaging time and alternative levels for the standard and whether these elements of the current standard need to be changed to protect the public health with a margin of safety.

Indicator:

I agree with the reasoning given in Chapter 10 that the indicator should remain as NO_2.

Averaging time:

The current averaging time is based on earlier studies linking long-term exposure to low levels of NO2 to adverse respiratory effects. More recent studies indicate that there should be more concern for the short-term NO2 exposures. A review of the data in Table 10-3 indicates that a standard based on annual average NO2 concentrations would not likely be adequate to protect against the effects of short-term exposures. I agree with the analysis presented in Chapter 10 that indicates that the standard should be based on either a 24-hr or a 1-hr averaging time. Based on the data in Table 10-2, I agree that a 1-hr standard should protect against 24-hr NO2 concentrations and therefore agree with the staff conclusion to use a 1-hr averaging time for the standard.

Form:

I agree with the reasoning presented in Chapter 10 that a concentration-based form is better than a form based on expected exceedances. I agree that whether you use the 98th or the 99th percentile is dependent on the level one chooses.

Level:

I agree that the health endpoint of concern is increased airway responsiveness in asthmatics. I agree with the staff arguments given on page 303 that the lower end of the range should be 0.05 ppm. I do not agree with setting the upper end of the range at 0.2 ppm. I would prefer 0.1 ppm. My reasoning is based on the fact a meta-analysis of the controlled human exposure studies in mild asthmatics indicated a LOEL of 0.1 ppm NO2. No severe asthmatics were studied. The Clean Air Act stipulates that we should make our recommendations to protect public health with a margin of safety. Because the endpoint of interest is in asthmatics and because we need a margin of safety, I would not recommend an upper level of the range higher than 0.1 ppm.

Comments from Dr. Donna Kenski

Pre-meeting Comments on Risk and Exposure Assessment to Support the Review of the NO2 Primary National Ambient Air Quality Standard
Donna Kenski
Lake Michigan Air Directors Consortium
December 4, 2008

The piecemeal nature of this review made it very difficult to maintain a coherent picture of the document as a whole. I understand the constraints that EPA is under to meet the court-ordered dates for publication, but I hope that future ISA and REA reviews can be scheduled so that the panel has all the parts together to review at one time. I expect it may be difficult for the panel to come to agreement on the letter in our teleconference Friday, given the particular importance of the final chapter in the REA and the fact that this is the first chance we've had to discuss it. Especially since it is now several months since we've had the benefit of our discussions on the rest of the document.

Despite the unfortunate logistics of the review cycle, I thought Chapter 10 did a great job summarizing the results of the exposure assessment, the risk assessment, and the epidemiological assessment and outlining how each of them provides support in their various ways for a short term NO2 standard. It was thorough, logical, and nicely written as well. Only occasionally did it lose its straightforward tone. For example, on p. 290, in the middle of the first paragraph of Sec. 10.3.3., after a succinct summary of the ISA's evidence for short-term effects, the REA states, "To the extent that these considerations are emphasized, the adequacy of the current standard to protect the public health would clearly be called into question." Surely these are exactly the considerations that MUST be emphasized. To state it this way makes it seem like an option. I don't think the Administrator has the option of ignoring, or even de-emphasizing, the conclusions of the ISA and the REA. The authors are trying too hard to avoid making statements that could be construed as policy judgments, when they are really sound, logical conclusions based on the excellent work presented in both the ISA and REA.

Averaging time: The chapter is a little confusing about the annual vs. hourly standard. Table 10-3 gives a nice comparison of annual averages that would result from just meeting alternative 98th percentile standards. What is somewhat startling is that, under some of the proposed alternative standards, the current annual standard could be exceeded. It begs for more discussion. Even given the conclusion that evidence relating long-term exposure to adverse health effects is suggestive but not sufficient to infer causality, it's hard to accept a new standard that potentially allows exceedances of the previous standard. This is a PR problem, if not a public health problem.

Form: The section on potential forms of the standard seems to have settled on the 1-h daily maximum without much discussion of the implications of multiple high concentrations in one day. I recall reading somewhere in the ISA or REA about inadequate evidence for effects of multiple exposures in a single day. That information would be appropriate to summarize here as well. Similarly, the reasoning behind the 3-

year averaging period is not discussed. Presumably this helps to stabilize the metric? The chapter notes that the 3-year averaging time is discussed in the PM and ozone reviews, but the rationale should at least be summarized here.

I have a modest preference for a 98^{th} percentile form, since it could be set to be sufficiently protective and is likely to be more stable than a 99^{th} %ile standard. Thus it is more likely to avoid the awkwardness of areas that flip-flop in and out of attainment, a situation that creates confusion for the public and for the state governments that need to implement the standard and controls to achieve it. But as the chapter appropriately notes, the protectiveness depends on both the level and the form, so a 98^{th} percentile form must have an adequately protective level. This section could have used some comparisons of, say, 98^{th} percentile at 0.05 ppm with 99^{th} percentile at 0.1 ppm. From Table 10-4, it seems that the 99^{th} %ile is never more than 0.01 ppm greater than the 98^{th} %ile. So might one infer that a 98^{th} %ile standard at 0.09 would be as protective as a 99^{th} %ile standard at 0.1 ppm? Which leads me to:

Level: After presenting an excellent summary on the concentration ranges at which statistically significant NO2 effects have been seen, I was surprised by the staff conclusion that the upper end of the range of levels is 0.2 ppm. I think both the ISA and REA make a very strong case that the upper end should be 0.1 ppm. Setting the standard no higher than 0.1 is all the more critical when one considers the requirement to include a margin of safety, and also the potential of the various analyses to underestimate exposures, which was nicely documented in Sec. 10.3.2.1.

Minor stuff:
p. 300, 1st paragraph: The discussion of averaging time is in Sec. 10.4.2, not 10.5, and the discussion of form is in Sec. 10.4.3, not 10.6

Hyphenation is improved over previous reports, but many instances of 'statistically-significant' are hyphenated in error.

The constant switching between ppm and ppb makes it unnecessarily difficult to compare data in one section of the chapter to another. Please make these consistent.

Comments from Dr. Stephen Kleeberger

I will not reiterate comments made by other panel members regarding adequacy of chapter 10 to summarize indicator, averaging time, level and form except to commend staff on an exceptional effort. I thought all sections were well written and clearly outlined the questions, summarized the existing pertinent investigations, and

With the understanding that chapter 10 will not be revised, I would yet recommend that a table be created to summarize the uncertainties described in the chapter. The text does well to define the uncertainties (beginning on page 284), but inasmuch as chapter 10 attempts to bring together many of the key points in the entire documents, I would have found useful a concise tabulated summary of this important consideration in standard-setting. A summary like this could also serve to better delineate precise research needs.

I would also have found useful another statement in section 10.3.3 (page 290) to re-emphasize that other sub-groups may be as sensitive or more sensitive than asthmatics, and/or that other health endpoints may be important (e.g. birth outcomes)(the points were raised initially on p 279).

Comments from Dr. Timothy Larson

Comments by Timothy Larson on "Risk and Exposure Assessment to Support the Review of the NO2 Primary National Ambient Air Quality Standard Final Draft, November, 2008"

I will limit my comments to the air quality and exposure analyses. The discussion of the representativeness of the exposure results and an analysis of the key uncertainties makes the final versions of Chapters 7 and 8 stronger than previous ones. These key uncertainties are summarized in Chapter 10 including the degree of independence of NO2 distributions with extrapolation to scenarios "just meeting" regulatory levels, the accuracy of empirically-derived relationships between ambient and fixed-site monitors, the representativeness of the APEX activity data, the potential upward bias of the AERMOD predictions of the high end of the NO2 distribution of short-term concentrations, and the choice of Atlanta to represent urban areas within the U.S. The assessment of these specific uncertainties is well-balanced.

However, one source of uncertainty that deserves more discussion involves the amount of exposure that occurs inside residences that are in close proximity to major roads. These exposures are, in principle, captured by APEX's "Indoor-Residence" category. However, the perimeter of the census blocks in Atlanta do not appear in general to cross major roads, i.e., the blocks boundaries are defined in part by these roads. Therefore the assignment of an outdoor exposure level by AERMOD to the centroid of the census blocks means that a number of people reside closer to the road than those living exactly at the block centroid. Given the potentially strong horizontal concentration gradients in NO2 near roads in flat terrain and the strong vertical gradient near roads in built-up areas, the centroid assignment approach used by EPA is probably underpredicting the exposure to those living very near the roads . This is important because 17% of the population in Atlanta resides within 75 m of a major road. It would seem that the actual indoor exposures of this subpopulation due to NO2 penetrating from outside is greater than that predicted by APEX. As it stands, the staff concludes that the AERMOD analyses are "overpredicting the upper percentile exposures" (p. 286).

I agree with the proposed averaging time. One minor point: the near road sites show annual averages that are closer to the 1-hr standards than sites located further from major roads. These relationships may also change if the near-road indoor exposures are considered.

I agree with the proposed upper and lower range of the level. One additional factor to be considered in the ANPR is the possibility that future monitoring sites may be located nearer to the roadway than they are at present, i.e., the NO2 siting criteria may change from "urban scale" to "microscale". There is certainly precedent with carbon monoxide monitoring (short-term standard, major mobile source contribution).

I do not have a strong opinion as to which of the proposed forms of the standard (98th or 99th percentile) are appropriate. I think that stability of the target should be given extra weight by EPA over other arguments.

Comments from Dr. Armistead Russell

Post-call Comments on the Revised REA, with specific focus on Chapter 10 and response to the question "What scientific evidence and/or scientific insights have been developed since the last NOx review to indicate if the current public-health based based NAAQS need to be revised or if alternative levels, indicators, statistical forms, or averaging times of these standards are needed to protect public health with an adequate margin of safety?"

Armistead (Ted) Russell

First, I was pleased by the progress made between our meeting and this current version of the REA. A number of issues have been addressed, but some linger. In particular, I am still concerned about the high bias in the AERMOD results, particularly for simulating higher end concentrations, and it is these concentrations that are driving the exposure analyses and the interpretation of the results of these analyses in Chapter 10. The performance is really not that good, but I am persuaded that the results should be used, but the performance issues should be considered more objectively, with particular attention to the biases in simulating upper end concentrations, and the on-road-to-monitor ratios. Uncertainty, error and bias are very different beasts when discussing and using model results. Here, bias is the major issue, and I am not persuaded by the staff arguments for not dealing with it more thoroughly. Also, it would be very good to give a pdf of simulated on-road concentrations. In regards to Table 8-7, it is important to note those concentrations labeled as AQ monitors are not measured, but are adjusted measured concentrations (see issues below).

A second concern is that the exposures over specified benchmark levels dominantly occur on road unless that benchmark level is 100 ppb AND indoor sources are included, in which case, indoor sources play a major role. In that case, Fig. 8-25 suggests that a 100 ppb benchmark level has virtually no effect on the fraction of asthmatics impacted even with a 50 ppb, 99^{th}%ile standard. Given that, we must be very concerned with how well we can approximate on-road exposures (which I would suggest is not very well) given the high bias in the AERMOD results (Fig. 8-8 is actually rather alarming in this case, which shows the on-road to monitor values ranging up to over 100, though the study of in-vehicle to monitored values that is cited in the ISA and REA (Riediker et al., 2003) , note that "NO2 concentrations inside the cars were always low." , noting that the average in vehicle level was 31.4 ppb when the mean ambient was 30.4 and roadside was 49.9, when one high extreme value that they could not explain was removed.). Further, the AERMOD results for on-road vs. non-road levels are virtually all greater than one (except at the 0%ile), while observed on-road-to-monitor factors are less than 1 at the 50%ile. This strongly suggests that the modeled on-road exposure is biased high, likely significantly. I also note that the (albeit limited) comparison of observed to modeled NO_2 exposures provided further suggests a potentially large high bias in exposures at the higher end of the distributions (Fig. 8-11 vs. Fig. 8-12). Countering the concerns, above, but not quantitatively addressed in the REA, is that there are urban areas that are not as spread out as Atlanta, and the residential or other near-road exposures could be proportionally higher. Providing a pdf of concentrations on the road, as well as, say 20m and 50 m from the road, might provide some information to how this issue might influence the interpretation of the results to other locations.

If one, instead, uses on-road-to-monitor ratios, there is again a problem. It is not apparent that the derived ratios account for a limitation on the ability to oxidize NO to NO_2 during the periods with the highest NOx levels. If the ozone levels are significantly reduced, as they will be in locations of high NOx emissions, NO (the primary form of NOx emitted) will not be rapidly oxidized to NO_2. Thus, using a distribution of on-to-off road ratios (m) that does not take this in to account (e.g., using lower m's during periods of high NOx levels, if the underlying data suggest such) could lead to a high bias in predicted on-road NO_2 at the higher end of the distributions. Again, this leads to a bias in estimating the number of exposures of concern (e.g., over specific benchmark levels). Associated with this issue is that the ratio of monitor to "on-road" NO_2 levels will be very sensitive to where the monitor is located. Using near-road monitors to assess attainment would require less of a buffer to protect against high on-road NO_2 exposures than one located distant from a major road (though studies have found lower in-vehicle NO_2 levels).

A next concern in conducting this risk assessment is that I believe that the concern over confounding is downplayed. NO_2 is a product of combustion, so non-laboratory exposures typically will co-occur with other products of combustion. Of particular concern is PM, and more specifically, fine PM rich in carbonaceous compounds. However, combustion byproducts also include a large range of organic gases as well. Some of these gases are also of concern (e.g., formaldehyde). Thus, I consider using multi-pollutant models critical for assessing if there is an association of concern (both single and multi-pollutant models should be used, and results considered). Unlike the REA (and the ISA), given the number of studies that are no longer statistically significant when using multi-pollutant models, I would not characterize the results as robust. Further, they have not measured, and thus do not control for, many of the other combustion byproducts of concern. Why this is of additional concern is that controls to reduce NOx emissions might actually lead to increases in those other combustion byproducts (e.g., VOCs, such as formaldehyde, and constituents of PM), and this risk is not explored in the REA. Further, increases in the VOC emissions would lead to increased ozone in cities, particularly those where NOx emissions appear to lead to decreases in average ozone exposure. This is also not considered in the REA.

In regards to Chapter 10, the document does clearly present EPA's rationale behind their choice of an indicator, averaging time, form, and range of levels for a modified primary NO_2 NAAQS. This chapter was very well laid out and presented, and was one of the best summary chapters of this type I have read. I agree with the rationale behind the choice of indicator (NO_2), and do not believe that current instrumental issues are a major problem, particularly if one uses one hour (or 24-hour) maximum levels. As for averaging time, the evidence is clear that a shorter averaging time is called for, and I agree with the rationale for a one hour average. I do not believe a 24-hour average is as strongly supported. As for form, a concentration-based standard, using a 98^{th} or $99^{th\%}$ile, averaged over three years, is appropriate, and that choice should be influenced by the choice of level. I tend to prefer the less extreme for stability between years. They should be clear in how this is stated, i.e., that the maximum 1 hour average is taken for each day, and that the 98th%ile is taken from the distribution of 1-hour averaged, daily maximums.

EPA staff suggests a range of level for consideration between 50 to 200 ppb. The upper end could actually lead to an increase in exceedences of benchmark levels, and there is ample clinical evidence to suggest that health effects would be found at concentrations that would occur if 200 ppb were the level chosen, particularly given the meta-analysis suggesting an lowest observable effect level of 0.1 ppm (which I take as having only one significant digit). Given the issues discussed above, the lower end is less strongly supported, and has issues that have not been addressed in the REA. As such, I would be most comfortable with a 98%ile, 1-hr daily maximum, NAAQS of about 100 ppb.

Comments from Dr. Christian Seigneur

Comments on final version of the REA

Christian Seigneur
CEREA, Université Paris-Est

My comments pertain to the discussion of the uncertainties. The first paragraph of Section 8.12.1.6 provides a good introduction to the uncertainties associated with the air quality modeling by clearly stating the types of uncertainties and their causes (e.g., AERMOD is essentially a model designed to identify maximum impacts from elevated point source; it is applied here to evaluate a distribution of concentrations due to emissions from mobile sources, i.e. a problem that differs significantly from AERMOD's original purpose). However, the second paragraph lacks objectivity and sounds more like an a posteriori justification of the use of a model with bad performance than a true discussion of uncertainties, as exemplified by the following statements (all line numbers refer to the first full paragraph of p. 236).

Lines 2-4: "the evaluation of modeled air quality…shows overall good agreement between AERMOD… and…monitored NO_2 concentrations". AERMOD overpredicts the mean NO_2 concentration at all three Atlanta monitors. The REA only mentions explicitly the smallest overprediction (10%), conveniently avoiding to mention quantitatively the larger overpredictions at the other two receptors. It appears clearly in Figures 8-6 and 8-7 that AERMOD overpredicts NO_2 concentrations overall and it is inappropriate to state that AERMOD shows "overall good agreement" with the measurements.

Lines 8-10: "the degree of bias…is within the factor 2 commonly used to indicate relatively unbiased model performance". OAQPS seems to confuse "error" and "bias". The factor of 2 is generally expected for ground-level impacts of elevated point sources; however, it does not imply that remaining within this factor of 2 means that there is no bias. If all modeled values are within a factor of 2 of the measurements but overestimate the measurements, there is a bias toward overestimation.

Lines 10-12: "In considering that this upward bias occurs mainly in the early morning hours, it is possible that there may not be a large proportion of the simulated population exposed at these times of the day". The high NO2 concentrations are due in part to the heavy traffic at those hours, which implies a large number of people driving those cars. Unless this statement can be quantitatively substantiated, it should be deleted.

In summary, the REA document should present the uncertainties associated with the AERMOD simulations in a more objective fashion instead of using the uncertainty discussion as a justification for using a model with poor performance. It would also be appropriate to highlight the need for the future development of a model that would be more appropriate than AERMOD for simulating the air quality impacts of roadways.

Comments from Dr. Richard Schlesinger
Comments on REA, specifically Chapter 10:

Overall, this final draft presents in a clear manner the scientific bases for recommending the NOx standard.

The few specific comments are as follows:

P. 283. First bullet in second paragraph. The term "control" may not be the best word to use in this context. Perhaps the statement should read "negatively affect asthma treatment protocols."

P. 287. The last sentence in the first paragraph of section 10.3.2.2 notes that 1 hr levels above 0.2 pm are unlikely to occur in locations that meet the current annual standard. Levels at or above 0.2 ppm have been associated with adverse health outcomes. Thus, this line of thought seems to contradict the earlier statement in the last sentence on page 282 of section 10.3.2, which states that results of studies support a relationship between NO2 exposures and respiratory endpoints at ambient concentrations in areas that meet the current standard.

P. 309. Section 10.4.4.3 This section concludes that the standard level should be set between 50 and 100ppb. However, the prior discussion seems to indicate that if the current level of 53 ppb as annual average is maintained, that shorter term excursions would occur that may result in adverse health outcomes in some groups of the population. Furthermore, the last sentence implies that the REA is recommending a change from the current standard. Thus, this reviewer is somewhat confused as to what is being recommended here.

Comments from Dr. Lianne Sheppard

Final draft NO2 REA comments

Summary:
The major changes to this document are an added chapter 10 and revisions to chapter 8 in response to CASAC comments. I am concerned that the difficulties with receptor locations in AERMOD have not been corrected or even fully acknowledged. This is has serious implications for the results and our inference from them. Several other important changes have been introduced, such as a table characterizing key uncertainties in chapter 9.

Organization and placement of policy discussions:
Chapter 10 provides a bridge between the conclusions of the ISA and REA and the policy recommendations to appear in the ANPR. The discussion contained in chapter 10 is an essential component of the review process and has been well crafted. However, the outline of the current process suggests to me that the ANPR is supposed to cover policy assessment while the REA should be confined to risk-based considerations to inform standard setting (e.g. see figure 1-1). Therefore I suggest duplicating section 10.4 in the ANPR. Policy inferences in the REA are a helpful precursor to the ANPR because there is an opportunity for CASAC feedback with time for EPA staff to integrate this feedback into their work. The description of the process should be modified to acknowledge the incorporation of this preliminary policy assessment into the REA and the name of the REA should be changed to reflect its broader scope.

Important omissions:

- The assumption regarding receptor locations for AERMOD predictions that are fed into APEX is not adequately justified and is completely omitted in the Chapter 8 discussion of uncertainty. Therefore this likely major source of underprediction of exceedances is completely ignored in chapter 10.
- The data have a great deal to tell us about monitor siting and levels of NO2. The monitoring network design, particularly with respect to proximity to roads, will have huge implications for whether or not exceedances are measured. Chapter 10 needs to discuss the data in the context of insights they provide into network design. Given the move towards a short-term standard, it will be important to add regulatory monitors near roads. (Consider the NO2 sources and the precedent set by regulation of CO.) As is evident in Table 10-3, monitors that are closer to roads have annual averages that are higher than monitors far (>100 m) from roads. The implications of network design on the annual standard also must be discussed.

Uncertainty presentation:

- Consider using recent advances in the assessment and presentation of uncertainty presented by the Intergovernmental Panel on Climate Change (see e.g. the IPCC 4th Assessment reports, such as the AR4 Synthesis report and the Physical Basis report; http://www.ipcc.ch/ipccreports/assessments-reports.htm).

- Make sure to clearly delineate suspected biases as part of the uncertainty presentation.
- The presentation should explicitly recognize the CAA requirement to protect the population with an adequate margin of safety. Given the need for a margin of safety, uncertainty should not be used as an argument for increasing the levels under consideration for a standard (e.g. p. 304).

Specific comments chapter 8:

- The addition of Figure 8-1 is useful. However it does not show receptor locations as inputs into AERMOD, a major oversight.
- p 150: There should be some discussion of the link between the receptor locations used in the AERMOD predictions and the selected census blocks used by APEX.
- p. 167: There has been no assessment of the distribution of population as defined by receptor locations related to the actual percentage of the Atlanta population living close to roads. Footnote 12 is helpful additional info, but is insufficient. The buffers for evaluation should be selected based on scientific considerations, not regression model fits to predicted data (description p. 216). The locations of the centroids used in this analysis do not appear to have changed since the last version. I am concerned that too few centroids are within 50 or 100m of a major road relative to the population of Atlanta. Since NO_2 concentration varies dramatically as a function of distance to nearby major roads, if the distribution of distances to major roads does not reflect the Atlanta population, the results could be seriously biased. Given the different buffers in the descriptions on pp 167 and 168, it is difficult to make any sense of the limited statistics that are reported.
- P. 168-9: The description of the distributional calculations on p 169 was clearer to me this time. Justify the selection of the 4 km buffer for evaluation of predicted relative to monitored concentrations with e.g. additional data analysis. I am concerned that the summarizing features of within-location distributions across locations could give a misleading focus of attention. Since figures 8-6 and 8-7 are based on pointwise summarizations across locations, evaluation of the approach is important.
- p. 190: Quantify the number of non-road block centroids located near a major road and pay particular attention to short distances (not just the 75 m buffer). Stratify the AERMOD on-road CDF in Figure 8-9 by road proximity category. Possibly pass this through to APEX to estimate the indoor distributions also.
- Figure 8-17,18: Provide details that clearly defines what locations, individuals, and activities that make up each category in the pie chart and provide this reference with the figures. It is too easy to make assumptions about what is contained in each category. I'm also concerned that near-road exposures (e.g. at residences) are underreported in AERMOD and similarly in how they pass through to APEX indoor exposures. To that end, consider adding at least one additional micro-environment in Table 8-9: Indoors – Residence near major road.
- P. 216: Distance to roadway bins should be based on scientific considerations, not regression analyses of predicted data. Based on Zhu et al (2002), most of the

effect is within the first 100 m. It would be entirely appropriate to have 2-3 bins completely within that first 100m.

- p 216: The description of estimation of population percentages within buffers is not transparent. It should also be made very clear that Table 8-14 is based on a model that uses unverified assumptions of uniform distributions of populations over space. How were major roadways at the boundary of the block handled? Are boundaries "within the block"?
- p 217: Be careful about circular reasoning. It is possible that most exceedances were associated with in-vehicle exposures because the spatio-temporal NO2 predictions relied on receptors that were systematically farther from roads than some of the population. If the model did not properly locate residents sufficiently close to major roads, the at home predictions of NO2 would have been biased low.
- Section 8.12: This section does not even mention receptor locations as a source of uncertainty.

Specific comments, Chapter 10:

- p 281, bottom 1st paragraph: Statistical significance should not be used as a criterion for scientific importance.
- p 286: Receptor locations affect predictions and may not be aligned with the population distribution, particularly at the closest distances to major roads where the NO2 will be the highest.
- p 294: Table 10-2 presents ratios, not correlations.
- Table 10-3: Add N of monitors. Note that the annual average is closer to the 1-hour standard for monitors close to a major road (<100m). This has implications for network design as well as the choice of short- and long-term standards. Add a discussion of network design and interpret the data in that context.
- Keep ppb or ppm throughout. I prefer ppb because more significant digits can be retained.
- p 302: Not sure I follow the argument of why the plausible relationship between NO2 and health events would be more uncertain at the lower range of NO2. Is there data to suggest the mixture changes at lower concentrations?
- p 303: Recast discussions to focus on estimates and intervals, not statistical significance or positive associations.
- p 304: I wasn't convinced by the 0.2 argument.

.

Comments from Dr. Frank Speizer

Pre-Conference Call Review Comments on Chapter 10 of REA- Primary NO2 document. Dated November 2008

Submitted by: Frank Speizer

Date December 2, 2008

I have focused my comments on Chapter 10 with specific considerations of the questions posed in the email dated Nov. 21, 2008.

General Comment:
Generally I found the Chapter one of the best I have read over many years. It follows a logical format, outlines the issues clearly and presents the Staff position well. (I will take issue with one conclusion, and have indicated that below).

Specific Comments:

Page275, para 2, line 6 and spilling over to the next page. There is a degree of emphasis here on "uncertainty" vs. "margin of safety". The issue is clarified on page 276, but starting as it does it suggests a bias that comes though later in the Chapter, where the issue of margin of safety is all but forgotten.

Page 284, 1st bullet, sentence beginning line 5: Not clear what is being assumed here. May require further spelling out.

Page 285, 1st bullet, sentence beginning line 4: Although I agree with the statement, it is conceivable that more severe asthmatics (because of treatment or chronic symptomatic state) might actually be less susceptible when actually measured.

Page 290, top discussion of two pollutant model and then effect of adding O3 or PM10: I t is not clear what makes up the two pollutant model to which O3 or PM10 is added. I would have thought that one or the other of these pollutants was added to NO2 to make up the 2 pollutant model. Please clarify.

Page 290, Section 10.3.3: Para 1, sentence beginning on line 7. I applaud the Staff's conclusion here, but believe it could be even more direct, rather than "called into question"

Potential Alternative Standards

Indicator: Summarized briefly and well. Agree that NO2 remain the indicator.

Averaging time: The argument summarized in Tables 10-1 and 10-2 is very convincing that moving to 1 hour is appropriate. I don't know if another table or some example of the text on the page 293 that uses the table to convert ratios to actual estimates would be useful. Alternatively table 10-1 could be expanded to show the hourly maximums.

Form: Clearly the discussion of the 99th and 98th percentiles is appropriate. However, Staff does not seem to reach a conclusion and it is at this point "uncertainty" seems to dominate the thinking rather than "margin of safety". I would agree that the form is partly related to the level, and thus move on to that discussion.

Level. Here I would disagree with the Staff. The argument is well presented that the effect level is seen at or below 0.1 ppm. Staff indicated "little evidence of any effect threshold"…"key studies …with a range of 98th/99th percentile 1-hr daily maximum levels from 0.05 ppm to 0.21 ppm". The several bullets on pages 303 and 304 clearly summarize that health relevant effects are noted, including the fact that the ISA does not draw distinctions between levels within the range of 0.1-0.3 ppm. The fact that there appears to be more uncertainty at the lower levels means to me that the margin of safety issue is important. Thus I take issue with including on page 304, that there is reasonable supported by evidence to go to 0.2 ppm. In fact as the series of bullets on pages 307 and 308 suggest at 0.2 there would be little improvement either in air quality or health effects from the level set in 1971 and little reason to change the standard. That would clearly be wrong, as Staff points out, and I would agree (as they conclude on page 309) that the range should be between 0.5-0.1 ppm.

Comments from Dr. George Thurston

Prof. George D. Thurston comments on final "Risk and Exposure Assessment to Support the Review of the NO2 Primary National Ambient Air Quality Standard"

My comments will also be organized according to the indicator, form, averaging time and alternative levels for the standard, and whether these elements of the current standard need to be changed to protect the public health with a margin of safety.

Indicator:
I agree that the indicator should remain as NO2.

Averaging time:
I agree that the annual standard is not sufficiently protective of health, and with the recommendation to use a 1-hr averaging time for the NO2 standard.

Form:
I agree that a concentration-based form is better than a form based on expected exceedances, and that the absolute level of the standard chosen will depend on whether choose use the 98th or the 99th percentile. It will also depend on how many years are averaged to compare with the standard, if any multi-year averaging is done.

Level:
I concur with Dr. Henderson's argument that the upper level of the range should be no higher than 0.1 ppm.

Comments from Dr. Ronald Wyzga

Comments of Ron Wyzga on Chapter 10

Three approaches are used in Chapter 10 to consider the implications of alternative standard formulations: one based on analyses of data from 18 monitors throughout the US; one based upon an Atlanta case study using alternative benchmark levels; and one based upon the does-response taken from an epidemiological study in Atlanta. All three approaches are reasonably well-described, and no one approach is emphasized to the detriment of the others. There is a fundamental issue, however, that is not discussed as well as it should in this chapter; namely, the issue of differences between exposures and monitored concentrations of NO2. The document (and the ISA) notes that exposures are likely to be significantly higher near roadways, and monitors may not capture these high exposures. The first two approaches give greater consideration to this issue than the third. I therefore would place less weight on this approach for understanding this risks associated with NO2 exposure. There needs to be a greater discussion of the use and value of this approach than currently exists on p. 300 given the disparity between personal exposures (especially peak exposures) and monitored levels. The current discussion is based upon comparisons of personal exposures and ambient monitoring results for extended periods of time (>24 hours). For a much shorter period of one-hour, these comparisons are likely to be of little value. Highest exposures likely occur when individuals are near roadways; it is unlikely that they remain outdoors near roadways for periods of 24 hours or more. Unfortunately there are no measurements of short-term personal NO2 exposures. In the REA adjustment is made to estimate peak levels near roadways from monitored levels. The need for this adjustment contradicts the statement made at the bottom of p. 300. This issue also identifies the need to expand the discussion of how to interpret the results from the epidemiological analyses in the context of standard setting. A related issue is that of where regulatory monitors are likely to be placed; if they are placed near major sources or where ambient levels are thought to be higher (e.g., near roadways), the discussion in Chapter 10 would be different.

There are four attributes of the standard discussed in Chapter 10. The REA and supporting evidence provide no support for any alternative to NO2; hence the choice of this indicator is clearly justified. The ISA gives greatest support for a short-term standard, averaging time of one to 24 hours. Epidemiological evidence supports both averages; clinical studies demonstrate responses for shorter periods of time. The REA suggest an hourly standard based largely upon the human clinical study results. Exposure considerations buttress this choice. Individuals are much more likely to be exposed to outdoor levels of NO2 for shorter periods of time than longer periods of time because most individuals spend less than 10% of their time outdoors; hence the probability of being exposed to a peak level for shorter periods of time, such as one hour, is greater than that of being exposed for longer periods of time (e.g., 24 hours). Further consideration of potential exposures from the behavioral and exposure pattern literature should be incorporated into this document.

The discussion of the form of the standard vis-à-vis the 98th or 99th percentile issue is well-discussed and well-presented. More discussion could be given to the issue of using data over a 3 year period; no support for this time period is given in the document.

Finally the REA considers a range of standard levels. The range is large (50-200 ppb), but is appropriate for illustrating the potential impacts of alternative levels; the consideration of this range facilitates the choice of an appropriate level. The analyses are helpful, but could be more explicit with respect to the uncertainties associated with estimated impacts of alternative standard levels. The presentation of "central estimates" only, for example, on page 308 is potentially misleading. Although statistical significance is not the only consideration that should be given to interpreting study results, it is an important consideration that should not be obscured. The setting of a standard for NO2 is a difficult issue, especially given the collinearity of NO2 levels with both measured and unmeasured pollutants. Those who set the standards should be aware of the full uncertainties associated with the estimated impacts of a standard.