

Compilation of Individual Panel Member Pre-Meeting Comments on EPA's *Risk and Exposure Assessment (REA) to Support the Review of the NO₂ Primary National Ambient Air Quality Standard: Second Draft.*

Comments Received:

Comments from Prof. Ed Avol	2
Comments from Dr. John Balmes.....	7
Comments from Dr. Douglas Crawford-Brown	10
Comments from Dr. Terry Gordon	16
Comments from Dr. Dale Hattis	20
Comments from Dr. Patrick Kinney	22
Comments from Dr. Steven Kleeberger.....	24
Comments from Dr. Edward Postlethwait	26
Comments from Dr. Richard Schlesinger	35
Comments from Dr. Christian Seigneur	37
Comments from Dr. Elizabeth "Lianne" Sheppard	38
Comments from Dr. Frank Speizer	40
Comments from Dr. George Thurston	45
Comments from Dr. Ronald Wyzga	48

Comments from Prof. Ed Avol

Comments on the Second Draft NO₂ Risk Exposure Assessment Document
Ed Avol, CASAC NO₂ Primary Review Panel (04Sep08)

Characterization of Air Quality (Chaps 2, 6, and 7):

This second draft is a marked improvement over the previous version. There is additional detail, explanation, characterization, and continuity of presentation in the revised document. Several alternative approaches to the air characterization analyses have been developed and presented, and provide a useful perspective on consideration of the current standard and the basis for re-consideration.

The discussion of on-road and roadway-related exposures is timely, important, useful, and insightful.

The discussion of uncertainty and variability is an important one and should be in the document, but in my opinion, is somewhat out of place and better located in an appendix, with a shortened summary in the main text.

Characterization of Health Effects Evidence and Selection of Potential Alternative Standards (Chaps 3,4,5):

The health evidence has been presented in a useful and detailed manner, although in some sections (see detailed comments below), it seemed a presentation of the data was being made rather than a summary of findings from the ISA document. The challenging topics of susceptibility and vulnerability were again presented and described, and in this presentation, a clearer differentiation between the two has partially emerged. The challenge and objective for staff should now be to maintain consistent and discrete separation between these working definitions.

Characterization of Exposure (Chaps 6, 8):

It was somewhat surprising that the occasionally presented lower benchmark value of 50ppb NO₂ was not consistently carried through the various presentations of health, alternative standards, and exposure characterizations. If this was an intentional decision, then a justification for this omission and path of action should be presented for evaluation.

The issues of on-road and in-vehicle exposures are important ones thankfully discussed in the revised document. While there are still some portions of the discussion that would benefit from additional comment and data (including consideration of increased on-road speeds leading to higher air exchange rates, and potentially important differences in decay of pollutants with distance from roadside as a function of time of day (interwoven with temporal and spatial activities and opportunities for exposure), the presentation was generally focused, well-done, and appropriate.

(Review of Chapter 8 is delayed until the revised draft chapter is provided by staff)

Characterization of Health Risks (Chaps 7,8,9):

The range of potential health effects benchmarks does seem appropriately discussed, but the appearance/disappearance of 50ppb from consistent data presentation seems odd. The ED visits discussion related to Atlanta appeared to be well-done (but out of my immediate area of expertise).

(I await receipt of Chapter 8 to provide additional comments for consideration)

Specific Editing Comments and Questions

Page ii: Section 5. font changes in listing; make consistent

Page vi: several abbreviations missing from listing: BAL, CAMP study, ECP, EDAA, EDRA, MS, PMN

P4, Figure 1-1: (top right box, “qualitative characterization of US epidemiology studies” – Throughout document, there are references to relevant European/Australian/Mexican/Norwegian research results, so this comment is inaccurate.

(Note – if the notation “NO₂” were changed to “pollutant”, this figure would be generically applicable to virtually all subsequent REA documents)

P9, lines 25-26 (“...model-base estimates indicate that NO₂ levels in many non-urban areas of the United States are less than 1ppb...”): What level of confidence is there for assigning modeled concentrations to areas with little or no actual measurements? Have spot assessments confirmed that levels in non-urban areas are in fact less than 1ppb, or is this a modeling artifice imposed by boundary conditions of the application?

P11, line 21 (“...because most sources of NO₂ are near ground level...”) – I assume you’re actually talking about on-road traffic tailpipe emissions, so why not be explicitly clear?

P11. line 24 (“...levels are likely even higher at elevations below 4 meters...”) – I would certainly agree that this is broadly true, but there are notable exceptions with elevated exhaust, such as from trucks, trains, ships, boilers, power plants, etc, so this is not a universal truth...

P14, line4: “evaluates” should be “evaluated”

P14, lines 24-25 (“...there is only limited supporting evidence from clinical or toxicological studies on potential susceptibility to NO₂ in persons with

cardiovascular disease..."): but aren't there also just limited studies of these people, period?

P15, lines 11-17 (paragraph on criteria that must be met for establishing useful links between polymorphisms and adverse respiratory effects): I agree with these points, but don't these three criteria (involvement in the pathogenesis pathway, observable functional change, and careful consideration of possible confounding) equally apply to every potential adverse health effect?...so it seems to me, this paragraph is either broadly generic and applicable or unnecessary in this specific section.

P17, lines 1-4 (discussion of violence and elevated risk of exposure): what about confounders, such as being more likely to be "out on the streets", or in closer proximity to traffic, or other environmental justice angles (more likely to be closer to higher exposure areas such as train yards, truck depots, shipping docks – less affluent areas)?

P19, line 12 (and in multiple locations throughout the document)" repeated reference is made to the "last review of the NO₂ NAAQS", but it would be more useful to readers and more accurate to refer to the date of the last review (1995?) rather than "the last review", to underscore the time period until now and the opportunity for new and more refined research to have been performed but not yet have been considered...

P19, lines 14 and 15 ("...children and older adults..."): reference in a consistent manner to ages would be helpful to anchor the discussion, e.g. "...children (<18yrs) and older adults (>65yrs)..."

P22, line 2 – spell out what CAMP stands for (and place it in the abbreviation table)

P22, line 5 ("...each subject having an approximate average of two months of data..."): This is poorly worded and inaccurate (you really did not print an approximate average, you used the calculated average; I assume the "approximate" part refers to the two months), and should be re-worded.

P22, line16 ("...intervention study in Australia..."): this is scientifically fine (as far as I am concerned), but it is inconsistent with your earlier declaration about emphasizing US studies.

P24, line6 – "sites" should be "cites"

P25, lines 20-31: This section seems pretty detailed and more suited to the ISA; the findings should be summarized here, not presented in detail. Additionally,

BAL, ECP, and PMN should also be placed on the Abbreviations listing and defined in the text when they first appear.

P26, line 11 – replace “and” with “or”

P29, line12 – “thrombosis”, not thombosis

P32, line 4 – the presentation would be markedly improved with a two-column summary table for short-term NO2 exposure effects, listing the endpoint (respiratory, mortality,...) and the ISA determination (causal, insufficient,...)

P41, line 3-7 – In a previous discussion (P19, section 4.2.2, lines 23+), the Linn study was down-weighted (if not dismissed) for using having employed one-pollutant modeling; here it is identified as appropriate for use and included; isn't this inconsistent?

P42, Figure 5.1: include EDRA and EDAA to Abbreviations listing

P43, Figure 5.2: include MS in Abbreviations listing

P45, line 9 – change “thee” to “three”

P51, lines 8-10: Delete the sentence beginning with “while an individual ambient monitor...” since it adds nothing to the discussion and is a duplication of part of the next sentence.

P58, lines 14 and 20 seem to be contradictory; it seems to say near-road measurements were used to calculate on-road NO2 in Line 14, then say near-road measurements were not used in Line 20...???

P72, lines 16-18: if NO2 levels are generally declining and vehicle exhaust is the primary source, how to explain the increase in exceedances between 2004-2006 and 2001-2003 data sets?

P72, line18: replace “that” with “than”

P91, line 1 (“...if the monitors are not evenly distributed (causing a bias)...”): one doesn't necessarily always follow the other. If the monitoring surface is both smooth and rough terrain, evenly distributed samplers can miss the variability in the rough terrain (in valleys, up hills, etc) and over-report the smooth terrain.

P92, line 15: change “being” to “of this bias would be”

P95, line 16: change “froma” to “from”

P95, line 18: change “selecting monitors these monitors” to “selecting these monitors”

Comments from Dr. John Balmes

John Balmes

GENERAL COMMENTS

Characterization of Health Effects Evidence and Selection of Potential Alternative Standards for Analysis (Chapters 3, 4, and 5)

1. The presentation of the NO₂ health effects evidence is based on the information contained in the NO₂ Integrated Science Assessment. What are the views of the Panel on the overall characterization of the health evidence for NO₂? To what extent is the presentation clear and appropriately balanced?

Overall, I feel that the presentation of the NO₂ health effects evidence in Chapters 3 and 4 is clear and appropriately balanced.

2. The specific potential alternative standards that have been selected for analysis are based on both controlled human exposure studies and on epidemiological studies conducted in the United States. What are the Panel's views on the appropriateness of these potential alternative standards (in terms of indicator, averaging time, form, and level) for the purpose of conducting air quality, exposure, and risk assessments and on the rationale used to select them for that purpose?

I feel quite strongly that the epidemiological evidence supports a short-term averaging time, i.e., an annual standard is not the averaging time appropriate to protecting persons with asthma from developing exacerbations that result in health care utilization. I think the epidemiological data best support a 24-hour averaging time, but controlled human exposure study data better support a 1-hour averaging time. The alternative standard forms and levels that Chapter 5 states will be considered in the subsequent risk analysis are appropriate and reasonably well justified.

Characterization of Health Risks (Chapters 7, 8, 9)

1. Based on conclusions in the final ISA regarding airway responsiveness, we have expanded the range of potential health effect benchmark values to include 0.1 ppm. Do Panel members have comments on the range of potential health effects benchmark values chosen to characterize risks associated with 1-hour NO₂ exposures?

The range of potential health effects benchmark values chosen as described in Chapter 5 is appropriate, but there is inconsistency in how these values are applied in Chapters 7 and 9. In Chapter 5 (p. 44, lines 16-17), it is stated that 0.05 ppm will be considered in the risk assessment. I support doing so, but in Chapter 7, exceedance data for 0.05 ppm are not presented in the tables. In fact, data for 0.1 ppm are not consistently presented.

To improve clarity and avoid confusion, I suggest being consistent in the application of potential health effect benchmark values.

2. To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized?

The presentation of health risk results is technically sound, clearly communicated, and appropriately characterized.

3. A focused risk assessment has been conducted for emergency department visits in Atlanta, GA. To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized? What are the views of the Panel on the approach taken and on the interpretation of the results of this analysis?

The Atlanta-based risk assessment of the impact of varying NO₂ air quality standards is an appropriate approach to provide useful information to policy makers about the public health impact of alternative standards.

4. What are the views of the Panel regarding the clarity and adequacy of the discussion of uncertainty and variability with respect to the characterization of health risks?

The discussion of uncertainty and variability is reasonably clear, but there is some needless repetition. Chapter 7 contains an in-depth discussion of uncertainty and variability. It seems unnecessary for Chapter 9 (Section 9.6) to repeat this discussion in so much detail (to the point of using some identical text. This section of Chapter 9 would be improved if it placed the key issues regarding uncertainty and variability in the risk assessment in context. Specifically, how important are the various sources of uncertainty and variability in a relative sense, and how should they impact the interpretation of the results? As currently written, these questions are lost in the systematic listing of all sources of uncertainty and variability.

SPECIFIC COMMENTS

p. 13, lines 18-21 Diet is another factor that may affect vulnerability to NO₂. A diet low in antioxidant micronutrients likely increases vulnerability to NO₂.

p. 18, lines 5-11 The REA uses the same five-level hierarchy to assess the level of evidence for a causal relationship that was discussed by CASAC during the SO_x meeting on 7/30-31/08. The panel wrote about the use of this hierarchy as follows: “We concur with using the five levels but recommend that the descriptions be changed to better reflect the level of certainty or confidence in the classification of the level of evidence. The phrasing of the second level is particularly problematic in its addition of the wording ‘likely causal relationship.’ The approach to evidence interpretation should avoid using statistical significance as a criterion for evidence interpretation. CASAC recommends

that EPA reconsider the language used to describe the weight of evidence, particularly for the first three categories which cover a range of certainty or confidence in causal inference that extends from full certainty to lesser degrees. The language used should be consistent with other such schemes used by EPA.” The panel’s previous comments about this hierarchy are relevant here.

p. 23, lines 21-24 The statement, “In the laboratory, airway responses can be measured by assessing changes in pulmonary function (e.g., decline in FEV1) or changes in the inflammatory response (e.g., using markers in bronchoalveolar lavage (BAL) fluid or induced sputum) (ISA, section 3.1.3),” is somewhat misleading in a sub-section entitled “Airway Responsiveness.” Bronchoalveolar lavage and induced sputum cannot be used to directly assess airway responsiveness.

p. 25, line 21-p.26, line 11 Unlike enhanced lung function responses to inhaled allergen after NO₂ exposure, the discussion of airway inflammatory responses to inhaled allergen do not properly fit under the sub-section title of “Airway Responsiveness.” I would move this discussion to the next sub-section (4.2.6) on “Airway Inflammation.”

p. 95, line 3 Ozone would be a better example of a reactive pollutant with a low indoor/outdoor ratio than PM_{2.5}.

p. 95, line 13 Should be “from a major road...”

p. 116, lines 1-4 Should be “For example, changing from a 98th percentile 1-hour daily maximum standard based on **0.1** ppm to one based on **0.05** ppm reduces the estimated incidence of respiratory-related ED visits in Atlanta by about 49 percent in 2007 (from 4700 to 2400);...

Comments from Dr. Douglas Crawford-Brown

**Comments on Risk and Exposure Assessment (REA) to Support the Review of the
NO₂ Primary National Ambient Air Quality Standard: Second Draft**

Douglas Crawford-Brown (3-9-08)

This review is built around the charge questions provided. I focus on Chapters 3, 4 and 5, and then on chapters 7, 8 and 9.

First, a more general comment is appropriate. The overall REA is both a good step forward from the first draft, and contains a wealth of useful material. I will need to withhold complete judgment until Chapter 8 is complete, but the existing chapters are generally in good shape and make for a compelling argument. There is strong consistency between this document and the ISA, with caveats noted later.

There is, however, a need for a global edit of the document. The chapters differ in writing style and even, to some extent, in the way they have used the ISA material. There also seems to me more detail in the Appendices than are reflected in the document. It is appropriate for their to be more technical detail, but I felt that the Appendices suggested that more detailed studies had been done yet not reflected in the main document. I may be wrong about this, but I did get this sense, and it again suggests the need for a global edit to be sure all of the chapters, and the main document and Appendices, are consistent. Finally, the levels of exposure considered don't seem to me consistent throughout. Each chapter and section should use identical ranges of values.

For Chapters 3, 4 and 5:

1. The presentation of the NO₂ health effects evidence is based on the information contained in the NO₂ Integrated Science Assessment. What are the views of the Panel on the overall characterization of the health evidence for NO₂? To what extent is the presentation clear and appropriately balanced?

The authors have done a good job of tracking through the ISA and culling the more important results. These results have then been summarized quite well in the REA and from the basis for the assessment. I found it quite easy to follow the discussion although, as noted in my general comments, I had a sense that there was more detail in the Appendix than is evident in the body of these chapters.

I notice that only the human data seem to play any sort of role in the assessment. I had thought that the animal data were at least going to play some modifying role, although I admit we had decided on the CASAC that the human data would form the primary basis for setting the health benchmarks. At the least, the document should explain the lack of use of the animal data.

I also feel the ISA provided a better understanding of the issue of sensitive, susceptible, vulnerable populations. Why is there no mention of infants throughout? There is significantly more nuance and subdivision of populations in the ISA, and it was just odd to find so little included in the REA. I suppose the decision may have been made to abstract the ISA discussion and use only the most important results for the REA, but if this is the case, it should be explained. I also suppose the results might be the same if they were culled out as a special subpopulation, but can't tell from the analysis.

I agree with the decision to use the studies for which the conclusion is "... the available evidence is sufficient to infer either a causal or a likely casual relationship". However, this judgment is introduced with little supporting justification. As it may set a precedent for further REAs, it deserves a bit of discussion both in the document and in CASAC deliberations. Fortunately, the issue is not so pressing since, even if one uses this set of criteria, the evidence is sufficient to support the EPA assessment and to make a good case for regulatory consideration.

Finally, am I correct that the indoor NO₂ exposure studies are playing no role? This is a bit odd, given that they seem to me to be studies focused largely on NO₂ exposures by

themselves, which is precisely what we are trying to get at here. Or is the assumption made that they, too, are confounded by other exposures? At the least, there needs to be a better explanation as to why these studies are being rejected for use.

2. The specific potential alternative standards that have been selected for analysis are based on both controlled human exposure studies and on epidemiological studies conducted in the United States. What are the Panel's views on the appropriateness of these potential alternative standards (in terms of indicator, averaging time, form, and level) for the purpose of conducting air quality, exposure, and risk assessments and on the rationale used to select them for that purpose?

I first note that I was pleased to see an assessment based on both the clinical and epidemiological studies. This was a good step forward in letting the latter results play a role.

I support the use of NO₂ as the ***Indicator***, if for no other reason than that I don't know how any other indicator would be implemented.

I support the ***Averaging Time*** suggested, although would like to see some discussion of the implications of a daily averaging time, as it links more directly to the epidemiological studies. But I doubt this would change the level of exposure selected. And in any event, I don't think the epidemiological studies can fully support the choice of a daily average value, since there is a lot of intra-daily variability at a location, and this means the appearance of health effects may not be related to the daily average (but rather to shorter-term periods of elevation throughout the day).

On the ***Form***, I wasn't clear how this 98th or 99th percentile issue is to be dealt with methodologically. The document would profit from having an example stream of measurements, then showing how they are to be abstracted down to some sort of cumulative distribution function and the 98th or 99th percentile estimated and then averaged. I support the use of the 98th or 99th percentile as at least a good policy choice (I

am not sure it has any real scientific basis – which is not to say it goes against the science, only that it is not really a scientific issue), but just am not clear what it means methodologically.

As to the *Level*, I am again supportive, both of the lower and upper bounds selected. However, there is a reasonable argument to be made that the incidence of hospital admission for asthmatics may be high even at these levels. And the levels don't reflect, in my kind, any sort of margin of safety. So there is merit in at least considering shifting the range down by factor of 2 or so. Still, I suspect the current range contains a value likely to emerge as the Level for the final NAAQS.

For chapters 7, 8 and 9:

1. Based on conclusions in the final ISA regarding airway responsiveness, we have expanded the range of potential health effect benchmark values to include 0.1 ppm. Do Panel members have comments on the range of potential health effects benchmark values chosen to characterize risks associated with 1-hour NO₂ exposures?

I support this decision, especially since the fraction of asthmatics showing response at between 0.1 and 0.2 ppm (or 0.15) is so high. As I noted in the answer to Question 2 above, there is at least an argument to be made for a further reduction below 0.1 ppm based on a margin of safety, but I am comfortable that the expansion of the range down to 0.1 ppm now encompasses the range of values likely to be considered seriously in regulation.

2. To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized?

I found this a section that was easy to follow, and support the underlying methodology. I will withhold full judgment until Chapter 8 is presented and integrated, as that is where the wheels may fall off. But the methodology for estimating health impacts contingent on

exposure is well developed and sophisticated. The one point that will need clarity is how the health effects can be related properly to the spatial gradient of exposures in the epidemiological studies and in actual population exposures. This is an area that will require more substantive discussion of uncertainties, as I suspect it is an important driver of the results. It is also important to note that exposures close to roadways are likely to be controlled not just by the concentration field (which is considered in the assessment) but by mode choice in travel, which in turn is known to be affected by the quality of the built environment (i.e. how attractive and safe it is to walk or bike near roads). This again is an issue of uncertainty and not assessment methodology, since there is no good way to account for this at present.

3. A focused risk assessment has been conducted for emergency department visits in Atlanta, GA. To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized? What are the views of the Panel on the approach taken and on the interpretation of the results of this analysis?

I am generally supportive of this analysis. It was well written and simple to follow, and makes a good addition to the other analyses in supporting the case being made overall in the REA. Atlanta seems to have been a good choice, leading to (somewhat) clear conclusions). My sole concern is whether the baseline can really be established well. One year is used, but I worry about such a short baseline given the climate in Atlanta, which can swing substantially from year to year. This can change both exposures and the background incidence of related respiratory diseases. I would have preferred to see an average taken over at least several years as a baseline, and suspect the results are driven significantly by the choice of this baseline year.

4. What are the views of the Panel regarding the clarity and adequacy of the discussion of uncertainty and variability with respect to the characterization of health risks.

As with the other REAs, I remain concerned that the analysis of uncertainty is so qualitative. There has been so much work done within the EPA on development of tools for uncertainty analysis, I don't understand why there is the continued reliance on qualitative studies. The reader is left with the impression that all we can say is that the results are uncertain to some unspecified degree. An uncertainty analysis need not be a full Monte Carlo approach, with PDFs for each parameter and nesting of model uncertainty, but there should at least be more quantitative statements about the key parameters and model forms that affect uncertainty, the sensitivity of the estimates to these components, and some idea as to how large of an error might be introduced (and in what direction). This would also greatly improve the utility of the uncertainty results in the Appendix (around page 86) concerning exceedences. These seem to me particularly important results that deserve a bit more reflection and analysis for uncertainty.

Variability is treated a bit better in existing REA, although there is still not a good discussion of how variability and uncertainty are being separated; the purpose behind each kind of analysis; or how variability can at times affect uncertainty in the analyses performed.

To be helpful, the uncertainty analysis should also provide better understanding as to why the uncertainty exists. A summary code could be given for each source indicating whether the issue is one of representativeness (as in monitor locations), measurement methodology, number of samples, incomplete science to back a model, etc. It would then be possible to target future resources on reduction of uncertainty. At some point many years ago, the EPA had a significant program aimed at reducing residual uncertainty, and that work or kind of approach doesn't seem to be showing up in these REAs.

Comments from Dr. Terry Gordon

Terry Gordon

Charge Question Responses

Characterization of Air Quality (Chapters 2, 6, and 7)

1. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary NO₂ NAAQS?

The characterizations and analyses appear to be sound and clearly communicated.

2. In order to simulate just meeting potential alternative 1-hour daily maximum standards, we have adjusted NO₂ air quality levels using the same approach that was used in the first draft to simulate just meeting the current annual standard. To what extent is this approach clearly communicated and appropriately characterized?

To the extent of my somewhat low level of expertise in this area, I understood a great deal more of the approach in this draft of the REA. Therefore, I'd say the approach appears to be clearly communicated and characterized.

3. Because of the impact of mobile sources on ambient NO₂, we have estimated on-road NO₂ concentrations. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized? Do Panel members have comments on the relevance of this procedure for reviewing the primary NO₂ NAAQS?

This is quite relevant to the exposure of individuals to ambient NO_x. Because the time spent 'on-road' may differ significantly amongst MSA's (i.e., traffic-dominated Los Angeles vs. other U.S. cities), it may be appropriate to extend this analysis with different factors for low or high traffic/commuting.

4. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

The uncertainty and variability assessment were thoroughly covered.

Characterization of Health Effects Evidence and Selection of Potential Alternative Standards for Analysis (Chapters 3, 4, 5)

1. The presentation of the NO₂ health effects evidence is based on the information contained in the NO₂ Integrated Science Assessment. What are the views of the Panel on the overall characterization of the health evidence for NO₂? To what extent is the presentation clear and appropriately balanced?

The characterization of the health evidence was quite clear and the decision pathway to choosing the appropriate benchmarks was excellent. The one exception might be the explanation of the different choices of NO_x exposure concentrations in the risk analyses. It took careful and repeated reading to understand why some analyses would range from 50 ppb and up and why others started at higher values. Perhaps, a summary paragraph at the beginning or end of the section (5.5) would provide a clearer explanation of the pathway forward for the choices.

2. The specific potential alternative standards that have been selected for analysis are based on both controlled human exposure studies and on epidemiological studies conducted in the United States. What are the Panel's views on the appropriateness of these potential alternative standards (in terms of indicator, averaging time, form, and level) for the purpose of conducting air quality, exposure, and risk assessments and on the rationale used to select them for that purpose?

The potential alternative standards are appropriately chosen. It is unclear, however, what to make of the current annual standard given the language in the REA that states the "evidence is suggestive but not sufficient to infer a causal relationship between long-term NO₂ exposure and respiratory morbidity".

Characterization of Exposure (Chapters 6 and 8):

1. To what extent is the assessment, interpretation, and presentation of the results of the exposure analysis technically sound, clearly communicated, and appropriately characterized?

Waiting for Chapter 8.

2. The second draft assessment document evaluates exposures in Atlanta. What are the views of the Panel on the approach taken and on the interpretation of the results of this analysis?

Waiting for Chapter 8.

3. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

The assessment of uncertainty was very thorough although presenting it twice in the text and once in a table may be overkill.

Characterization of Health Risks (Chapters 7, 8, 9):

1. Based on conclusions in the final ISA regarding airway responsiveness, we have expanded the range of potential health effect benchmark values to include 0.1 ppm. Do Panel members have comments on the range of potential health effects benchmark values chosen to characterize risks associated with 1-hour NO₂ exposures?

I concur with the selected benchmark and value.

2. *To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized?*

Waiting for Chapter 8.

3. *A focused risk assessment has been conducted for emergency department visits in Atlanta, GA. To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized? What are the views of the Panel on the approach taken and on the interpretation of the results of this analysis?*

The approach is excellent and clearly communicated. If this focus on one city will be used to project risk across the nation, it might be appropriate to include other cities for validation/cross comparison (Philadelphia was done previously?).

4. *What are the views of the Panel regarding the clarity and adequacy of the discussion of uncertainty and variability with respect to the characterization of health risks.*

The assessment of uncertainty was very thorough although presenting it twice in the text and once in a table/conclusion may be overkill.

Minor Comments:

Page 13, line 27 – conduction disorders – heart or nerve?

Page 13, line 28 – hypertension – vascular or pulmonary?

Page 17, line 1 – Is the 4.3 ppb increase 24 hr or annual or 1 hr?

Page 20, line 14 – Delete space after ‘12’

Page 24, line 6 – Change ‘sites’ to ‘cites’

Page 24, Table 4-1 and line 17 – The 76% listed on line 17 does not match the 0.75 fraction in the Table. Also, there appears to be a typo on the 2nd row of data – the fraction for All Exposures (0.68) is higher than both of its components (Exercise is 0.59 and Rest is 0.67)

Page 30, line 12 – chronic is misspelled

Page 31, line 15 – fetal ‘growth retardation’?

Page 42, Figure 5-1 – The legend says 95% CI but such data are not presented in the Figure.

Page 46, line 11 – delete the 2nd ‘to’

Page 52, line 27 – a space is needed between scenario and 4.

Page 53, line 16 – It is unclear what is meant by “A screening...” because the previous sentence already said the ambient air quality data would be screened.

Page 73, line 15 – ‘other time not’ is unclear

Page 73, line 26 – insert ‘are’ before ‘provided’

Page 91, lines 7-17 – It is unclear why the vertical height discussion is included here rather than in the Spatial Representation section 7.4.4. I consider spatial/space to have 3 dimensions.

Page 109, lines 1-4 – The wording is such that someone could grab this reasoning/sentence, even after the earlier careful elucidation of the potential for confounding by co-pollutants, to dismiss the conclusions of the entire risk assessment.

Comments from Dr. Dale Hattis

Dale Hattis—Premeeting Comment on Charge Question 4 for the Analysis of Air Quality

3. Estimation of On-Road NO₂ Concentrations—Technically Sound? Adequately Characterized

In view of the prominence of on-road exposures in the exceedance analysis, I also think that it is important for the authors to document the underlying data they use and do quantitative uncertainty analysis for key parameters that determine the results. In particular they should document their conclusions about the distribution of values for “m” in equation 4 on p. A-107. The authors should document the m values they derived from each study and evaluate the uncertainty in this value they infer for this key parameter from each study. Documenting the individual m values for each study would allow an independent assessment of the authors’ conclusion that the data cannot be well described by a parametric distribution (e.g., lognormal, normal, etc.) whose values are not limited to those directly inferred from the existing studies. The current empirical distribution treatment implies that each observed “m” should be treated as having equal uncertainty/weight and that the differences in “m” values observed by the various authors under various circumstances should be considered to adequately represent the real variability that would be seen in a representative sample of roads in the locations modeled. In fact it is likely that the different measurements carry different uncertainties that could be assessed in a comparative analysis of the likely errors in different measurements. Such an analysis could potentially lead to different weighting of (1) measurements that were done with greater vs lesser accuracy (taking account relative confidence/measurement uncertainties) and (2) measurements that were done under conditions of traffic volume and meteorology that are relatively common vs relatively rare (taking into account the representativeness of the measured conditions relative to the real variability among sites, etc., that is being modeled).

4. What are the views of the panel regarding the adequacy of the assessment of uncertainty and variability?

The assessment at present is entirely qualitative. In my opinion there should be at least some quantitative assessment of some of the important sources of uncertainty. For example in previous comments on an earlier draft of the REA I reported that model-based analyses of the likely effect of distribution of elevations of monitoring heights indicated that NO₂ concentrations observed at those heights were likely to understate NO₂ levels at breathing elevations of about 2 meters by 17-35% depending on the mathematical form of the function used for the decline of concentrations with height. Instead of creating its own more sophisticated analysis of this issue, the current document simply repeats a dismissal of the issue as not likely to be substantial in the light of the fact that most monitors are not as elevated as the monitor studied in the paper that documented the likely bias with monitoring height. Where an approximate quantitative analysis of this bias is possible, I think it should be done.

It is also important to quantitatively analyze and discuss the model biases and uncertainties indicated in Figures A-97 through A-99 on pages A-86 and A-87 of

Appendix A. To my eye, these figures suggest quite a bit of low bias in model predictions of observed exceedances, although the figures are not especially clear on this point. It would be better to put observed exceedances on the y axis and predicted exceedances on the x axis, and also to include an identity line ($y = x$) to better show the reader the magnitude and direction of the biases.

Also please see my comment in response to question 3 on the uncertainties in the present estimation of on-road concentrations.

Comments from Dr. Patrick Kinney

Kinney Comments on NO2 REA
September 8, 2008

I commend EPA staff on a much-improved document for review. My comments are largely either editorial or focus on improving the precision of the writing.

p. 11, lines 4-7: Edit sentence for clarity. This should be understandable to an educated lay person.

p. 11, line 15: at end of sentence, insert examples of oxidation products of concern in parentheses.

p. 12, lines 11-12: change “change the principal conclusions” to “bias conclusions in a positive direction”

p. 17, line 4: It would be helpful to mention here though that there is some evidence that low SES people tend to live nearer to busy roadways than more wealthy people. I think Margo Schwab’s research demonstrated this in Washington, DC.

p. 29, line 13: mention residences too in this context.

p. 18, line 28: the noted range of increased risks (2-20%) means little without information on the exposure range and time scale. Better would be something like, “Effect estimates from epidemiologic studies conducted in the US and Canada indicate that ED visits may increase by 2-20% on days when NO2 concentrations are X ppb higher than on other days.” This same problem recurs on lines 29-30 of p. 19, and elsewhere when presenting epi findings. See also page 27, line 14.

p. 20, line 1: I don’t think the word “detected” is appropriate here. Better would be “inferred” or “estimated.” Any calculation that is based on an arbitrary concentration range (say 30 ppb) is an inferred value, not a finding or observation from a specific study, especially where observed concentrations in the study were typically much lower than 30 ppb, and day to day changes as large as 30 ppb may never have been, or only rarely, observed. This problem occurs again at lines 14-15 of page 20, where it is stated that “The authors **found** a 12% increase in risk per 20 ppb increase in 24-hr ambient NO2.” They may have calculated that, but I doubt they found that.

p. 20, line 26: this statement that “NO2 associations in multicity studies are generally robust” is contradicted by one of the three studies noted. Seems like a more nuanced summary is needed, e.g., in the few multi city studies that have been conducted, more often than not, NO2 was robust.

p. 28, line 5: should the word “generally” be changed to “more often than not”?

p. 42, figure e5-1: I believe the NYC DOH results are for all ages, not just children. Also, caption should include 24 hour.

- p. 43, figure 5-2: caption should also mention 4 and 24 hour results.
- p. 45, line 9: “three” is misspelled.
- p. 59, line 19: change “two year-groups” to “two six-year groups” This occurs elsewhere also.
- p. 70, Table 7-10 and tables which follow: In spite of the footnote, I have a hard time understanding what these numbers mean in terms of spatial and temporal averaging. Need a better explanation to aid interpretation.
- p. 72, line 18: change “that” to “than”
- p. 93, lines 4-6: please expand further on why you think this is likely to be a problem and why it would go in a particular direction. Same goes for line 16. Not obvious to this reader why this would be the case.
- p. 94, lines 19-23: this sentence needs editing for clarity.
- p. 107, line 7: Since some of the observed baseline incidence is due to as-is NO₂ concentrations, do these ED visits need to be subtracted out to get the true “baseline”?
- p. 111, line 6: “three” is misspelled.
- p. 115, line 5: however, as noted above, NO₂ was only moderately correlated with PM₁₀ and O₃ in this study. The text as-written is misleading in this context.
- p. 116, line 2: I think the concentrations should be swapped in this sentence.

Comments from Dr. Steven Kleeberger

I thought the second draft of the REA for NO₂ was very well written, and overall have very few substantive comments. The sections were largely very clearly presented and approaches adequately explained. Minor issues are raised below in response to specific questions.

Characterization of Air Quality (Chapters 2, 6, and 7)

1. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary NO₂ NAAQS?

I thought the air quality characterizations and analyses were appropriately presented.

2. In order to simulate just meeting potential alternative 1-hour daily maximum standards, we have adjusted NO₂ air quality levels using the same approach that was used in the first draft to simulate just meeting the current annual standard. To what extent is this approach clearly communicated and appropriately characterized?

This approach was well communicated and characterized.

3. Because of the impact of mobile sources on ambient NO₂, we have estimated on-road NO₂ concentrations. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized? Do Panel members have comments on the relevance of this procedure for reviewing the primary NO₂ NAAQS?

No comments. As above, I the approach was technically sound and well communicated.

4. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

This is not my particular area of expertise.

Characterization of Health Effects Evidence and Selection of Potential Alternative Standards for Analysis (Chapters 3, 4, 5)

1. The presentation of the NO₂ health effects evidence is based on the information contained in the NO₂ Integrated Science Assessment. What are the views of the Panel on the overall characterization of the health evidence for NO₂? To what extent is the presentation clear and appropriately balanced?

This area was largely well-written, and the presentation was appropriately balanced. However, the general lack of inclusion of animal modeling data was somewhat surprising. Most of the discussion focused on epidemiology and chamber studies. I also thought that a table was going to be created to illustrate/list susceptibility factors known to be, or potentially could be, important in responsivity to NO₂ (perhaps I missed it?). In the "Age" section, perhaps some statement that infants could also be a particularly

susceptible subgroup. Some evidence, though not sufficient to indicate absolute causality, would suggest that this is the case (e.g. increased incidence of SIDS associated with NO₂). A minor point: it is not clear what is meant by "physiological susceptibility" (page 17, line 12).

2. The specific potential alternative standards that have been selected for analysis are based on both controlled human exposure studies and on epidemiological studies conducted in the United States. What are the Panel's views on the appropriateness of these potential alternative standards (in terms of indicator, averaging time, form, and level) for the purpose of conducting air quality, exposure, and risk assessments and on the rationale used to select them for that purpose?

I thought the introduction and explanation of form was appropriate and nicely tied to other criteria pollutants.

Characterization of Exposure (Chapters 6 and 8):

1. To what extent is the assessment, interpretation, and presentation of the results of the exposure analysis technically sound, clearly communicated, and appropriately characterized?

The assessment, interpretation and presentation were clearly presented in Chapter 6.

2. The second draft assessment document evaluates exposures in Atlanta. What are the views of the Panel on the approach taken and on the interpretation of the results of this analysis?

N/A

3. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

N/A

Comments from Dr. Edward Postlethwait

Edward M. Postlethwait, Ph.D.

Comments on the NO_x second draft REA for the Sept 2008 CASAC meeting (in no particular order)

1. The document is clearly improved and the extensive efforts by EPA staff in crafting the REA are evident.
2. Section 3.4 (Susceptibility: Genetics) needs some minor editing to correct inaccuracies (e.g., “absent peroxide activity” which I assume is “peroxidase”; not all ROS are electrophilic).
2. The statements in section 6 regarding the utility of airway hyperresponsiveness versus ED visits are not always clear and thus some editing would strengthen the document. “For example, the public health significance of the effect in question (i.e., ED visits) is less ambiguous in terms of its impact on an individual than in the case of airway hyperresponsiveness.” While the rationale for this is likely evident to those in the field, it does raise questions as to how ED visits are characterized with regard to NO₂-specific related pathophysiological events. Thus, in the absence of more characterizations, counting all ED visits rather than attempting to relate AHR to ambient NO₂ levels does not necessarily appear to strengthen risk estimates.
3. The benchmark NO₂ concentrations are not consistent throughout the document. For example, on pg 49, 50 ppb 1 hr levels are included but do not appear elsewhere.
4. A graphic and/or pictorial representation of the best estimates of the spatial distribution of NO₂ levels around roadways would be useful. Although there appears to be uncertainties in the precise spatial distribution, a consensus best estimate that would give the reader an improved feel for exposures relative to highway proximity and that would help identify where the 100 m boundaries exist (middle of the median outward, edge of road, etc) would seem germane.
5. The estimated risks/exposures appear to be based on Atlanta due to a single publication. However, Atlanta represents neither a worst or best case scenario and has missing data in many of the presented tables. Thus, staff should consider either including additional cities or at least providing a more compelling rationale for the focus on Atlanta. Furthermore, to this reader all the figures/tables in chapter 7 do not lend clarity but rather make this section more difficult to readily decipher. Perhaps final compilations should be presented with the numerous iterations moved to the appendix.
6. Tables 9-1,2,3 suggest that the inclusion of PM₁₀ and O₃ reduce the NO_x related ED visit incidence by factors of > 4 fold bringing into question the strong statements regarding causality and the “robust” NO_x related outcomes in multi-pollutant models. These discrepancies should be resolved to improve internal consistency.

7. While the numerous potential limitations and confounding factors are noted and discussed, it is unclear how they factor into the risk analyses. Based on the presented information, there appears to be a large number of uncertainties that could have substantial impacts on estimating health impacts. Thus, it seems critical that sensitivity analyses are incorporated to evaluate the extent of confounding and how any and all of the uncertainties affect the health/exposure outcomes boundaries.

Comments from Dr. Armistead Russell

Pre-meeting Review of EPA's "Risk and Exposure Assessment to Support the Review of the NO₂ Primary National Ambient Air Quality Standard"
 Armistead (Ted) Russell

General Review:

My overall view of the REA (sans Chapter 8) is positive. I think it begins to lay out the information needed to inform setting of a NAAQS. One aspect I felt missing was that as it focuses on the need, and the respective level, etc., of a short term standard, I was a bit disappointed at the lack of discussion and analysis of "what about the old one"? Will it still be needed and why? What happens if it is removed? The latter question has to go beyond just the health issues, but welfare as well. Certainly, a one hour standard for NO₂ has little relevance to addressing environmental acidification or nutrient enrichment. Given the recent decision by the administrator to not promulgate a separate welfare NAAQS (i.e., for ozone), not having a long term NO₂ standard has major implications. What might be needed in this case is to look specifically at the sources likely to be affected by a short term standard and assess how controls targeted at a one hour standard might impact longer term concentrations and emissions overall.

I might also suggest against the wording on page 5, line 4 "... and of any potential alternative standards." This is not limiting in any way shape or form, and in particular, suggests that standards could be introduced here that have not been subject to any prior relevant analysis. Given the current approach to the NAAQS reviews, it would be appropriate to have this altered to suggest standards for which informative analyses have been conducted and reviewed.

One part that needs to be much more precise is section 5.4 as to the Form of the standard. In particular, you note that it is the 98th percentile averaged over 3 years. By that, I assume one means taking the 98thile from the daily maximum one hour average for each of three calendar years, and taking that average. An alternative is to take all 24-hr measurements, and taking the 98thile, and then averaging. (This was less ambiguous for PM since only 24-hr measurements are used.) This should be spelled out specifically and precisely, along with the mathematical formula.

It seems as though the levels of concern examined vary between chapters, sometimes not including the 0.05 ppb mark (e.g., figures in Chapter 7).

1. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary NO₂ NAAQS?

There is a current weakness here, that raises its ugly little head later in the document. It would help, at this point, to have a more comprehensive description of the range of measured NO₂ levels across the US. As noted below, we are led to believe that there are areas where the monitored mean is 3 ppb. Is this true? What type of area might this be?

The order of cities in Table 7-1 is not apparent at this point, becoming apparent when looking at Table 7-2. I prefer alphabetizing.

Tables 7-3 and 7-4 need to be more precisely described in the footnotes. In particular, what annual means are being described? Are these the annual means of the 98th percentiles of the daily one hour maximums?

2. In order to simulate just meeting potential alternative 1-hour daily maximum standards, we have adjusted NO₂ air quality levels using the same approach that was used in the first draft to simulate just meeting the current annual standard. To what extent is this approach clearly communicated and appropriately characterized?

The approach used is reasonable, though the assumptions (and justifications) should be stated, preferably with some foundation as to why the assumptions are appropriate. This may involve considering specific sources and control approaches and saying “There is no more appropriate approach given what is currently known.” However, this should be done after consideration of specific sources.

As a detail, the mathematical formula description should be tightened up some, i.e., :

F_{ij} = Adjustment factor for location i and year j (unitless)

This tightening should be done for both eq. 6-1 and 6-2 (though j stands for two different things in the two equations). Further, they should specifically state, here, that the PRB is taken as insignificant.

In Table 7-12, I would add a characterization of the F 's for each area.

3. Because of the impact of mobile sources on ambient NO₂, we have estimated on-road NO₂ concentrations. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized? Do Panel members have comments on the relevance of this procedure for reviewing the primary NO₂ NAAQS?

My major concern about this analysis is how well founded is the calculation of “ m ”. While I guess this would come later, this is an area where some quantitative uncertainty analysis should be done.

In equation 7-1, k is the “decay” constant, not the “rate” constant.

4. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

At present, the discussion is largely qualitative, with little information as to even the ordering of major versus minor uncertainties. Section 7.4.2 has much of the discussion on monitor location not technique. This should be moved to 7.4.4. The authors should be specific on the extent of biases and uncertainties. In regards to a standard based on the 1-hr maximum of NO₂, I would argue that the 50% value for the technique bias is

irrelevant.

It is here that we get some discussion on the adjustment factor that should also be present when the process is introduced. Introduce the issues earlier on, noting the more detailed discussion later on. This uncertainty analysis could also be a bit more quantitative (or semi-quantitative).

I appreciate Table 7-19., but would like to see a column characterizing level of uncertainty.

P 93, l 26: Not sure what you mean by “concentration profile”.

Specific Points:

P 20, l 26: Make sure that everyone would concur that the results are “robust”. This word is subjective, and can mean different things to different people. I would be more specific, specifying how many studies still find statistically significant results when multi-pollutant models are used. My take of 3.1-7 does not suggest the word “robust” is quite correct. Further, they only control for PM10, not PM2.5 or, say, EC. Looking at 3.1-10 and 3.3-2 leads me to further suggest the word “robust” should not be used here.

P24, l 6: “cites”

P27, l 14: 2-20% increase per “what”

P32, l 8: This sentence, as written, indicates that adverse health effects can be identified for a mean 24-hr concentration of 3 ppb. This seems a bit far fetched. Indeed, looking at the Table 5.4-1 of the ISA, I think the 3 ppb comes from the Linn study, and if that is the case, it was 3.4 pphm, not 3.4 ppb. I might re-check the units on the various studies summarized in this table. Looking at Figure 2.4-13, it appears as though the minimum mean NO₂ is 5 ppb, and I would have to imagine that is at a pretty sparsely populated location. I am curious, at what locations where epi studies have been conducted is the mean NO₂ under 10 ppb? I will also add, the fact that an adverse outcome associated with NO₂ exposure is found in areas where the mean is, say, 10 ppb, does not mean that the adverse outcome results from an exposure of 10 ppb.

P36, l 11: This line should be “NO_x, for the purpose of this document, includes...”

P41, l 16: “Figures 1 and 2” (Capitalize) . Check capitalization of “Figure”, “Chapter” and “Table”, etc. throughout.

P 45, l 9: three (not thee)

P 67: footnote: Table 7-10.

P72 l18: “than” not “that”

P72, l28: remove “that”

P 73, l26: “) as provided...”

P107, l 17-18. One should note that in the Peel et al., and Tolbert et al., studies, that having only 36 out of 42 hospitals providing usable data does not mean that a similar fraction of potential ED visits are captured (or, conversely that 6/42 are missed). Often, it is the smaller hospitals that are not captured, so a very high fraction of ED's are included. Also, are you sure o the 36 number, particularly for 2004?

Section 9.6 seems out of place. Shouldn't it go later on?

Comments from Dr. Jonathan Samet

Review Comments

Prepared by: Jonathan M. Samet, MD, MS
Risk and Exposure Assessment to Support the Review of the NO₂ Primary National
Ambient Air Quality Standard: Second Draft

Characterization of Health Effects Evidence and Selection of Potential Alternative
Standards for Analysis (Chapters 3,4,5).

1. Charge Question 1:

In general, the document adequately represents the key evidence in the ISA with regard to the overall findings. Chapter 3 covers susceptibility, describing the range of populations found to be susceptible, both to air pollution generally and to NO₂ specifically. The document might be improved by a sharpening of its conclusions. Clearly, one important overall finding is that a large number of people could be susceptible, when considering the full range of groups identified. On the other hand, the experimental and epidemiological evidence would appear to converge in finding that asthmatics are the most susceptible. This conclusion may be stated obliquely on page 14, lines 17-19 in Chapter 3.

The concept of vulnerability, as distinct from susceptibility, is introduced, and appropriately followed through. It will be important to continue to maintain this same distinction between vulnerability and susceptibility in subsequent documents related to the NAAQS for other pollutants.

Chapter 4 provides the relevant findings of the ISA from the literature on the health effects of NO₂. As with Chapter 3, the ISA is satisfactorily distilled for the main points relevant to the REA. There are several important issues to be discussed. This set of evaluations for NO₂ uses the five-level classification of strength of evidence for causation. On page 32, lines 1-3, the staff makes the judgment that it will focus on endpoints for which the ISA “concludes that the available evidence is sufficient to infer either a causal or a likely casual relationship”. This represents a decision that sets a precedent with regard to the level of evidence in support of outcomes that will be considered in the REA. I do not dispute the choice, but I would suggest that it be better justified, given the future significance of this decision.

I also suggest that a stronger justification is needed to set aside the studies of indoor NO₂. The stated rationale acknowledges that these studies focused on NO₂ alone to the extent possible and that the exposure situation indoors differs from that outdoors. On the other hand, the experimental literature is based on exposure to NO₂ alone. Given the emphasis placed on the experimental studies, there does not appear to be a solid rationale for setting aside the studies directed at exposure to NO₂ from indoor sources.

The results of selected epidemiological studies will be considered in the REA. A rather weak argument is made for the appropriateness for this purpose in lines 12-13 on page 33. This sentence, giving staff judgment, follows a reasonably comprehensive discussion of the strengths and limitations of the epidemiological data.

Charge Question 2:

In general, the basis for selecting the indicator, averaging time, form, and level for the NO₂ NAAQS are clearly stated. The averaging time of 1-hour is reflective of the duration of the experimental studies and the finding that there are adverse health effects. Should consideration be given to exploring scenarios for the 24-hour averaging time as well? It might be useful to learn of maintenance of this averaging time would accord the same protection for deeper exposures as would be reached by having the one-hour standard.

With regard to level, the document provides a clear rationale for assessing a lower range extending to 0.05 ppm. The upper end of the range is quite reasonable, due to the experimental findings.

The REA states that alternative long-term standards to the current annual value will not be considered. On the other hand, the evidence does not provide certain evidence that there are long-term consequences. Would a short-term standard alone be sufficient? Are there areas that would be in compliance with a short-term standard but not with a long-term standard?

Specific Comments:

Page #	Line #	Comment
9	5	Correct spelling anthropogenic
9	11	“ “ “
10	19	Can't the concentrations vary over the time of exposure?
10	28	“On average.(add in the United States) people spend...”
11	21	Delete “this produces” insert there is.
11	23	“2.5-fold (delete increase in) substitute higher
12	11	“However, (insert the possible consequences of this exposure error do not) delete “is not expected to”
13	3	Delete subpopulations (insert groups within the general population) are at increased risk for suffering (adverse effects from NO ₂ exposure.)
13	25	Delete are believed (insert have been found)
14	19	What is meant by “most sensitive”?
15	1	Delete agreed (insert had comparable findings)
15	13	“First the product “ Necessarily? What about regulations?
16	9	“the vicinity” (Insert of roadways)
18	21	“recent studies” (or the entire body of evidence?)
23	12	Delete lead to the type of outcomes assessed (Insert be the basis for the effects observed) (insert increased before respiratory illness
24	6	“In addition the ISA” delete sites (insert cities)
24		Table row 3 column 2 Should be 66%?
27	15	In association with? What exposure?
27	22	Redundant phrasing
27	27	Adverse (insert respiratory) health
28	10	Add (and the potential for confounding)
33	12	This sentence seems overly strong in view of the prior discussion of limitations of the epidemiological findings.
33	19	But don't these studies address NO ₂ alone as in the exposure studies?
34	20	“levels is (increased) airway hyperresponsiveness”
39	7	Would it be reasonable to consider the 24-hr averaging time as well?
45	9	“...we have employed thee approaches” Change to three.

Comments from Dr. Richard Schlesinger

RICHARD SCHLESINGER

Overall, this draft is an improvement over the first. Certain concepts that were not clear in the first draft have been clarified in this one. Some comments on the specific questions follow:

To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary NO₂ NAAQS?

From my perspective, these are clearly communicated.

In order to simulate just meeting potential alternative 1-hour daily maximum standards, we have adjusted NO₂ air quality levels using the same approach that was used in the first draft to simulate just meeting the current annual standard. To what extent is this approach clearly communicated and appropriately characterized?

This is explained much better in this than in the first version of the document.

Because of the impact of mobile sources on ambient NO₂, we have estimated on-road NO₂ concentrations. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized? Do Panel members have comments on the relevance of this procedure for reviewing the primary NO₂ NAAQS?

No specific comment.

.What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

These are clearly addressed in this version of the document.

Chapters 3, 4, 5

The presentation of the NO₂ health effects evidence is based on the information contained in the NO₂ Integrated Science Assessment. What are the views of the Panel on the overall characterization of the health evidence for NO₂? To what extent is the presentation clear and appropriately balanced?

The health outcomes evidence was clearly presented. However more animal toxicology studies should have been included as basis to support findings in epi studies.

The specific potential alternative standards that have been selected for analysis are based on both controlled human exposure studies and on epidemiological studies conducted in the United States. What are the Panel's views on the appropriateness of these potential alternative standards (in terms of indicator, averaging time, form, and

level) for the purpose of conducting air quality, exposure, and risk assessments and on the rationale used to select them for that purpose?

I agree with the chosen alternatives

Chapters 7, 8, 9

Based on conclusions in the final ISA regarding airway responsiveness, we have expanded the range of potential health effect benchmark values to include 0.1 ppm. Do Panel members have comments on the range of potential health effects benchmark values chosen to characterize risks associated with 1-hour NO₂ exposures?

I agree with this approach. However, there does seem to be some inconsistency in the specific benchmark concentrations in different sections of the document.

A focused risk assessment has been conducted for emergency department visits in Atlanta, GA. To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized? What are the views of the Panel on the approach taken and on the interpretation of the results of this analysis?

I agree with this approach

What are the views of the Panel regarding the clarity and adequacy of the discussion of uncertainty and variability with respect to the characterization of health risks.

The discussion is quite sound.

Comments from Dr. Christian Seigneur**Comments on the NO₂ Risk and Exposure Assessment- Second Draft – September 2008.**

Christian Seigneur
CEREA – Joint Laboratory ENPC/EDF R&D
Université Paris-Est
Champs sur Marne, France

The second Draft shows some clear improvement over the first draft. For example, the uncertainty in the measurement of NO₂ ambient concentrations is now qualified as being small near emission sources (p. 11, lines 17-19). Furthermore, EPA has added benchmark concentrations as low as 0.05 ppm and has added a new analysis based on the epidemiological studies where the effect of NO₂ concentrations on emergency department visits are compared for different NO₂ concentration benchmark values. These latter revisions required a significant amount of new work by EPA staff and they clearly demonstrate that EPA listened and responded in an effective manner to the CASAC Panel's comments.

Most of my comments on the previous draft focused on the air quality modeling. It appears that the modeling of NO₂ with AERMOD is more challenging than expected, possibly because NO₂ is both a primary pollutant (emitted directly from vehicle exhaust and power plants, for example) and a secondary pollutant (formed by oxidation of NO by HO₂ and O₃ for example). The simplest approach to calculate secondary NO₂ concentrations near an emission source is to use an "oxidant-limited" approach, which assumes that the reaction of NO with O₃ is fast and stops when all O₃ has been depleted. The reaction of NO with HO₂ radicals and other peroxy radicals can generally be neglected near a NO_x source because the concentrations of those radicals are typically negligible when NO concentrations are high. Nevertheless, such a simple chemical scheme must be implemented carefully to correctly represent the atmospheric chemistry (for example, it is implemented incorrectly in the current regulatory version of CALPUFF).

Other air quality modeling related comments made on the previous draft pertained to the model performance of AERMOD, taking into account the uncertainty in model formulation in the uncertainty analysis, and the assumption made for the NO₂/NO_x emission ratio. I will wait to see Chapter 8 and related material before providing additional comments on the revised air quality modeling of this second REA draft.

Minor editorial comments:

p. 45, line 9 and p. 109, line 41: three instead of thee

p. 57, line 17; 0 meter instead of 0 meters (since it is less than 1)

Comments from Dr. Elizabeth “Lianne” Sheppard

General comments: This second draft REA is significant improvement over the first draft. I commend EPA staff for their success. The document is appropriate in its length and depth of coverage of the material, and it gives a concise yet clear discussion with improved flow. Chapter 5 is a good addition. It uses the ISA results to justify the standards to be considered in the risk and exposure assessment. Chapter 6 discusses the risk and exposure assessment approaches used in the remainder of the document. It is easier to follow this revised discussion. [Notably, there is no mention of the tiers described in the Health Assessment Plan. This language should be revisited in future rounds of the review process so there is more cross-document integration.] Chapter 9 has been added in response to CASAC comments and appropriately presents an example scenario to put the epidemiological results in context. There is no concluding chapter. The document may further benefit from this addition in order to integrate the conclusions from the three risk and exposure assessment approaches.

Air quality characterization: Chapter 7 is much improved. I have several major comments as well as a host of specific comments [to be] noted below.

- I appreciate the additional stratification of monitors by distance from major roads and the ability to see the differences in the analyses as a function of this stratification. More refinement is needed prior to presenting this analysis to describe the stratification and also the features of the two strata. Specifically, Table 7-1 should be expanded to show the number of complete monitors in each time period that fall into each stratum. In addition, it would be valuable to know how much heterogeneity there is across monitors in each of the strata on important characteristics such as the population density in a buffer surrounding the monitor, the dominant land use within the surrounding buffer, the types of major roads nearby, and the distribution of distances to major roads. An entire subsection should be added to Section 7.2 to further describe these features.
- This chapter still suffers from the assumption that the monitoring network actually represents a sample of monitors that provides meaningful information about population exposure. This assumption should be stated up front and evaluated in the beginning of the chapter. I suggest adding a section that discusses the monitoring objectives of the NOx regulatory monitoring network, presents an analysis of the proportion of monitors sited to meet each objective, and provides a qualitative discussion of how these objectives align with the intended inference from data analysis presented in chapter.

Health effects evidence characterization and selection of potential alternative standards: This appears appropriate.

Characterization of exposure: Chapter 8 is still not available.

Characterization of health risks: The addition of the health risk assessment for ED visits in Atlanta is generally appropriate and appears to be sound. I appreciate this added perspective.

Specific comments:

- p. 11 14-7: I think this statement is too general. The consequences of exposure misclassification will depend upon the use of the data.
- p. 11: The distinction between exposure and concentration are blurred. The focus is on exposure but the discussion is about concentration.
- p 12 11-13: Again I think this statement is too general here. In the context of personal exposure-response vs. concentration-response estimates from epidemiological studies, I think an important distinction is that the two types of studies are estimating different parameters. This is more important than a focus on measurement error bias (which I would agree shouldn't be neglected).
- [To be added]

Comments from Dr. Frank Speizer

Comments on the NO₂ Risk and Exposure Assessment Document dated August 2008

Submitted by: Frank E. Speizer, MD

Date: Sept. 1, 2008

Chapter 3:

Discussion of Susceptibility by dependent variables: The REA document appropriately presents the conclusions of the ISA as to the adequacy of the evidence (and were appropriate) lack of adequacy of the evidence for a wide variety of risk groups. For preexisting disease it appears that only asthma appears to be supported by both epidemiological and clinical or toxicological studies. For cardiovascular disease and diabetes only epidemiological studies are cited. Further comment needs to be presented that the reason there is not more supporting evidence is because the studies really have not (and maybe cannot) been done, rather than there being negative clinical or toxicological studies to report. For age the only data appear to be epidemiologic and this point although indicated by the studies cited could be more specifically stated. The discussion on the genetics of susceptibility is appropriately qualified. One genetic marker is identified and though plausible the studies do not provide specific support. Finally, a little more could have been said to rule out gender as a specific susceptibility factor as it seems illogical not to indicate that it really must relate to degree of exposure rather than anything inherent in being male or female.

Issues of vulnerability: The Staff has made the definition of vulnerability to essentially mean increased risk of higher exposure. This is an acceptable definition but not clear that all would agree. It is also not clear that it will define groups separate from those that are at increased susceptibility. Their own example of selected ethnic groups at higher risk of asthma indicates that susceptibility cannot be separated from vulnerability.

It is not clear why in section 3.7 Staff has focused on the study of Clougherty et al where violence is mentioned as a modifier. Surely figure 5.3.1 of the ISA gives the strong impression that most of the studies of asthma ED admissions are positive, without suggestion that these are all associated with chronic violence exposure.

Chapter 4.

The table presented on page 24 (table 4.1), and text on page 25 provides an important summary statistic that needs to be carried forward in analysis (I hope I find it). Two-thirds of asthmatics appear to be responsive as defined as airway responsiveness with exposure between 0.1 ppm- 0.15 ppm.

Top of page 32: Staff appears to have made a decision that only outcomes that result in a consensus that the overall effect is either causal or likely causal would be considered for formal quantitative risk assessment. This is reasonably well supported but could be considered too conservative. Since the differences between likely causal and suggestive but not sufficient to be considered causal are for the most part related to the numbers of studies rather than the potential plausibility of results, it seems to me to include this category at least initially would be warranted.

Choice of studies: Preferred US based studies. OK but would add or at least consider some of the European studies, in which NO₂ exposure in particular may very well be better documented.

Ambient over indoor. Again, OK but the blanket rejection of studies that include an indoor component seems inappropriate.

ED visits and hospital admission, OK there are more data but this works for me. Single and multiple pollutant models—good.

Range of 0.1-0.3 ppm chosen for constructing health risk estimates appropriate. However, some of the latter tables go down to 0.050 ppm.

Chapter 5

Indicator of NO₂ still reasonable.

Averaging time: Issue here is whether suggestive is enough to make consideration of a standard appropriate. This is one of the concerns brought about by the formalization of causation status. It must be remembered that whatever criteria are used they are a “suggested plan” for formalizing reasoning and should not be considered as “written in stone”.

The fact that previous work led to the conclusion that a long term standard was needed and new work in the ISA only indicates there is a suggestive relationship cannot be the sole criteria for dropping the standard. This issue will need to be discussed further. A better argument for dropping the standard will have to do with the degree to which the

public is protected by an alternative short term standard and whether the levels of exposure for a long term standard simply do not exist.

Form of standard: consideration of the concentration level rather than frequency of exceedences seems reasonable. The issue will be how this concentration is determined in the 1 hour averages and if a longer time period is required. Although the Staff appears to want to be consistent with what was done with PM, my recollection is that the final choice of 98%ile over 3 years was an administrative decision and not a scientific one. The fact that both 98 and 99%ile is being looked at is good.

Level: Upper bound of 0.2 ppm is appropriate, however data presented in Figure 5.1 suggest a lower upper bound should be considered. Given that 5% excess risk for asthmatics for admission is occurring with a 98 or 99%ile level of 0.09-0.1, respectively, to allow the upper bound to go higher will most likely double the risk if one speculates from Figure 5.2 that 30-40% excess symptoms of asthma occurs with a 98 or 99%ile at 0.05. This becomes too high. With regard to the lower bound Staff seems to be accepting this relatively high risk of asthmatic symptoms as the norm. This is not reasonable and some margin of safety or measure of uncertainty must be factored in. I would suggest at least cutting the lower bound by a factor of 2.

Chapter 6

A brief description of apparently what is being done in Chapter 8. If I understand it what is being done with the “role down” is to bring the calculated estimates of benchmark levels to below currently found levels. Given the change from Annual averages to 1-24 hour averages, I would like to be assured, and it will be interesting to see, if we have sufficient short term measure at Background levels to do this.

(Unfortunately, Chapter 8 did not appear in time for the meeting).

Chapter 7

Approach: The selection criteria of sites seems reasonable Unfortunately, ~20 of sites selected did not meet criteria and in looking over Table 7-1 as a summary of the sites it is apparent that the sites at altitude (Denver, Colorado Springs) as well as the hot dry climate of Phoenix in the 2001-06 period did much worse than the sea level sites in terms of completeness of data. This will need to be discussed and considered in terms both of exposure representativeness as well as potential for health risk at altitude. This is

particularly troublesome since two of these three sites had the most exceedences in the study period summarized in table 7-2.

Although I may have missed it, it needs to be made clear in discussion of tables 7-5 to 7-8, and 7-10- 7-11, that the number of exceedences at 150 contains the total of those at levels over 150 (if that is the case). Otherwise looking at the tables can be confusing and misinterpreted.

Not clear why in tables 7-12, 7-13, 7-16 p98 is not included in the tables. Up to this point this value is included and is likely to be part of the discussion on the form for setting the standard.

Uncertainty analysis: Exposure discussion is good with consideration of a number of issues. More could be said of the “missingness” and potential biased representativeness of the available data (for example, the concern about the selective nature of the missingness at altitude). With regard to the statement on page 95 of the overestimated effects of in-vehicle levels, it seems a little too strong for the available limited data. It is conceivable that under particular traffic conditions (heavy traffic resulting in idealizing) the in vehicle values could be considerably higher than roadside measures.

Table 7-19 is particularly troubling. Taken out of context it suggests that there are very defined bias directions that can be accounted for in each of the source questions. The text provides much of the discussion of just how complex these issues are and seems to stand alone. Even though the Table is titled “Summary of qualitative uncertainty...” it doesn’t add very much and might be poorly used to simply confuse the administrator in assuming that the results for each category of source can be defined better than is really the case.

Chapter 9

It would appear that the choice of Atlanta to model the risk assessment is quite appropriate. (It might also be useful to consider another city with more sustained excesses or more traffic like Los Angeles if time permitted). However, it seems to me that the selection of one year (2004) as the baseline incidence may not be appropriate. Although it is probably true that the general incidence of respiratory ED visits does not change over a period of a few years, the potential for a unique year to be influenced by events such as

a high or low flue season is great and therefore averaging an incidence figure over even 2 consecutive years would have been better.

In tables 9-2 and 9-3 it is not clear what is used in the calculation of the figures for the “current annual standard”. (I may have missed the explanation of how the daily max hourly is derived). Is a footnote necessary to clarify this for the reader?

Although it is useful to summarize the uncertainties at the end of the chapter, there is considerable confusion and variability of the nature of the kinds of uncertainty expressed. For example uncertainty about the level of certainty of causality is quite different for the uncertainty of the representativeness of Atlanta to other urban areas. Similarly the statistical uncertainty due to sampling error is quite different from the estimate of baseline incidence. The more thoughtful discussion of each of the topics as presented in Section 9.6 is there is not helped by this effort to put it all in one paragraph.

Comments from Dr. George Thurston

Prof. George Thurston's Initial Pre-Meeting Comments on CASAC Oxides of Nitrogen Primary NAAQS Review Panel on EPA's Second Draft of:
Risk and Exposure Assessment to Support the Review of the NO₂ Primary National Ambient Air Quality Standard

Overall, I find the document much improved, and feel that the EPA staffers are to be congratulated on an excellent job responding to CASAC's concerns regarding the first draft. I am especially pleased to see the epidemiology-based risk assessment provided in Chapter 9, and am hoping to see something similar provided in the next draft of the SO_x REA.

My specific responses to the assigned questions for this 2nd draft NO_x REA are provided below.

Characterization of Air Quality (Chapters 2, 6, and 7)

1. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary NO₂ NAAQS?

RESPONSE: This is done well, in general. However, the reference benchmarks considered seem to vary in different parts of the document: Page 46 (line 18) says 100, 150, 200, 250, and 300 will be considered, but Table 7-8 considers only exceedances of 150, 200, 250, and 300 (no 100?). In addition, Chapter 5 says that a standard of 50 ppb will be considered (on line 16), so shouldn't EPA also present the number of exceedances of that level in Table 7-8 etc., too? Similarly, Table 7-5 and 7-6 only consider 150, 200, 250, and 300 ppb. Shouldn't these all consider the same benchmarks to allow maximum comparability?

Also, with regard to term usage in the text, the word "historic" is used sometimes applied to describe data collected in the past (e.g., footnote on pg. 52, line 15 of page 54, line 14 on pg. 62, etc.), while sometimes it is call "historical" data : I believe the latter is the more correct usage. Similarly, "less than" is occasionally used (e.g, line 25, pg. 67), but I think "fewer than" is the more accepted usage when referring to counts.

2. In order to simulate just meeting potential alternative 1-hour daily maximum standards, we have adjusted NO₂ air quality levels using the same approach that was used in the first draft to simulate just meeting the current annual standard. To what extent is this approach clearly communicated and appropriately characterized?

I think that the explanation of, and justification for, the rolling upward and rolling downward could be worded a bit more clearly on page 47, lines 21-30. I think it has to be made clearer that each roll-up/back model is seeking to allow us to see how much more acute NO_x effects protection a potential new short-term standard might provide, vs. the existing annual standard, and vs. other potential short-term standard options.

3. Because of the impact of mobile sources on ambient NO₂, we have estimated on-road NO₂ concentrations. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized? Do Panel members have comments

on the relevance of this procedure for reviewing the primary NO₂ NAAQS?

I think this is a very important analysis, as it is said that people spend something like 6% of their time commuting, but get something like 60% of their daily outdoor air pollution exposures during that time (Source: CARB). I think the approach taken is a reasonable one.

4. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

This is a very good qualitative assessment of the factors affecting uncertainty and variability, but I would have liked a more quantitative assessment by which to inter-compare the importance of each factor: are they of differing orders of magnitude? A figure with a range of possible uncertainty effect (e.g., max-min range of % change) would be more helpful, if it could be developed.

Characterization of Health Effects Evidence and Selection of Potential Alternative Standards for Analysis (Chapters 3, 4, 5)

1. The presentation of the NO₂ health effects evidence is based on the information contained in the NO₂ Integrated Science Assessment. What are the views of the Panel on the overall characterization of the health evidence for NO₂? To what extent is the presentation clear and appropriately balanced?

This section is concise and well done: sufficiently clear and balanced.

2. The specific potential alternative standards that have been selected for analysis are based on both controlled human exposure studies and on epidemiological studies conducted in the United States. What are the Panel's views on the appropriateness of these potential alternative standards (in terms of indicator, averaging time, form, and level) for the purpose of conducting air quality, exposure, and risk assessments and on the rationale used to select them for that purpose?

The choice of 50, 100, 200, 250, and 300 is appropriate, but not all levels are consistently considered in all sections of the document. For example, Table 7-5 considers only 15, 200, 25, and 300, ignoring 50 and 100 ppb. Similarly, on page 55, line 7, the minimum benchmark is stated to be 200, not 50. The lower benchmarks should also be addressed throughout to allow comparison with other Tables and Figures in the document.

Characterization of Exposure (Chapters 6 and 8):

1. To what extent is the assessment, interpretation, and presentation of the results of the exposure analysis technically sound, clearly communicated, and appropriately characterized?

Chapter 6 does a good job of this. Chapter 8 remains to be seen.

2. The second draft assessment document evaluates exposures in Atlanta. What are the views of the Panel on the approach taken and on the interpretation of the results of this analysis?

This remains to be seen, as Chapter 8 not provided yet.

3. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

I assume that this will be provided in Chapter 8.

Characterization of Health Risks (Chapters 7, 8, 9):

1. Based on conclusions in the final ISA regarding airway responsiveness, we have expanded the range of potential health effect benchmark values to include 0.1 ppm. Do Panel members have comments on the range of potential health effects benchmark values chosen to characterize risks associated with 1-hour NO₂ exposures?

I feel that the benchmarks should also include 50 ppb across all analyses, as it is already used in the epidemiological risk assessment.

2. To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized?

The application of the epi-based risks are sound and clearly and appropriately characterized.

3. A focused risk assessment has been conducted for emergency department visits in Atlanta, GA. To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized? What are the views of the Panel on the approach taken and on the interpretation of the results of this analysis?

I highly commend Staff for this effort: this work is excellent, and highly useful in the standard-setting process. I strongly recommend that this approach be emulated in future REA's (e.g., SO_x).

4. What are the views of the Panel regarding the clarity and adequacy of the discussion of uncertainty and variability with respect to the characterization of health risks.

This seems adequate, though the addition of quantitative estimates (e.g., order of magnitude) would help in inter-comparing the potential importance of each source of uncertainty.

Comments from Dr. Ronald Wyzga

Comments on CASAC Oxides of Nitrogen Primary NAAQS Review Panel on EPA's
Second Draft of:

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

Comments on Second Draft of:

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

Overall comment:

There is a very impressive amount of work that has gone into this document; I am clearly impressed by the volume of materials and the thought behind much of the written material. My major problem is that it appears as if the various chapters were written by different individuals and it is unclear how much of the material fits together. Indeed at times the material in the various chapters/appendices appears to contradict material in other sections. I personally believe that the authors of this document would have benefited substantially if additional time had been available to them before the release of this draft.

Characterization of Air Quality (Chapters 2, 6, and 7)

1. To what extent are the air quality characterizations and analyses technically sound, clearly communicated, appropriately characterized, and relevant to the review of the primary NO₂ NAAQS?

The air quality characterizations are well presented and summarized. I have two major questions, however. The document uses a dichotomy of >/<100 m of a roadway in order to characterize monitors and their resulting data. Yet it also states that under most circumstances there is a 90% reduction of the NO_x roadway emissions within 10m of a roadway. Given these information, is the 100m breakpoint appropriate? Should an alternative breakpoint have been considered? Are there any monitors within 10 m of a roadway? What do they show? Also I would like to see data on the diurnal patterns of NO_x levels, especially around roadways. APEX says it includes such information, but it is not explicitly covered elsewhere; hence I am confused. Intuitively it seems to me that the highest exposures could occur around roadways during rush hour, but I don't see anything in this document that addresses this issue.

It would also help me to understand the data if Tables 7-3 and 4 explicitly indicated the number of monitors considered for each city; this would give some indication of the extent there was coverage in a given city.

Tables 7-5 through 7-8, Tables 7-10-7-18 also indicated the total number of opportunities for there to be exceedances; in pother words, what is the denominator for the rate of exceedance. It is not clear from the text.

2. In order to simulate just meeting potential alternative 1-hour daily maximum standards, we have adjusted NO₂ air quality levels using the same approach that was used in the first draft to simulate just meeting the current annual standard. To what extent is this approach clearly communicated and appropriately characterized?

The material is clearly presented; I remain uncomfortable with a concept that shows far greater risks and exposures resulting from conformance to the current standard than with an “as is” scenario. I am also uncomfortable with the great variability in multiplier constants across different geographic areas; I have no solution for this problem, but it is disturbing and suggests to me that there are considerable uncertainties introduced by this approach.

3. Because of the impact of mobile sources on ambient NO₂, we have estimated on-road NO₂ concentrations. To what extent is the approach taken technically sound, clearly communicated, and appropriately characterized? Do Panel members have comments on the relevance of this procedure for reviewing the primary NO₂ NAAQS?

See my response to question 1. I would like to see more data on diurnal variability near roadways and from monitors closer than 100m roadways.

I am also unclear about how values of the on-road factors (m) were randomly assigned? What distributions were assumed for m ? p. 58, ll. 23-24.

4. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

See responses to the above questions. Some of my questions may need to be addressed in the uncertainty section.

Characterization of Health Effects Evidence and Selection of Potential Alternative Standards for Analysis (Chapters 3, 4, 5)

1. The presentation of the NO₂ health effects evidence is based on the information contained in the NO₂ Integrated Science Assessment. What are the views of the Panel on the overall characterization of the health evidence for NO₂? To what extent is the presentation clear and appropriately balanced?

I believe it is accurate and fair.

2. The specific potential alternative standards that have been selected for analysis are based on both controlled human exposure studies and on epidemiological studies conducted in the United States. What are the Panel’s views on the appropriateness of these potential alternative standards (in terms of indicator, averaging time, form, and level) for the purpose of conducting air quality, exposure, and risk assessments and on the

rationale used to select them for that purpose?

There is inconsistency within the document about the level chosen for analysis. Personally I believe that a shorter-term exposure (e.g., 1 hour) is more appropriate than a longer-term one (24-hour); this also allows greater congruence between the human clinical and epidemiological studies. If I recall correctly, didn't the Clean Air Act Amendments of 1977 suggest that a standard of 3 hours or less be considered; is this still relevant?

Characterization of Exposure (Chapters 6 and 8):

1. To what extent is the assessment, interpretation, and presentation of the results of the exposure analysis technically sound, clearly communicated, and appropriately characterized?

I have mixed feelings here. I read the Chapter 6 and feel that I have a good understanding of the analytical approach taken. Then I read the Appendix, which is also well-written, but seems to take a more detailed approach than is conveyed in the Chapter. For example, I am unclear how and the extent to which APEX is applied from reading the document per se.

2. The second draft assessment document evaluates exposures in Atlanta. What are the views of the Panel on the approach taken and on the interpretation of the results of this analysis?

I am basing my comments on Appendix C, which I basically like.

In the discussion of the Tolbert et al. and Peel et al. studies, it should be noted that the authors considered an a priori lag structure for hypothesis testing purposes. Peel et al does present lag-specific results for asthma and upper respiratory illness, which show that most response is for lags 0-2 although there is also some response for NO₂ for lag 3 as well.

I would ask whether the risk assessment should consider time of day. Near-road exposures could be considerably higher during certain periods, and I wonder given the capabilities of APEX can this be addressed?

It should be noted the Peel et al. and Tolbert et al. papers were part of a larger study (ARIES) that considered other health endpoints as well. In particular, daily respiratory mortality and unscheduled physician visits were considered in papers by Klemm et al and Sinclair and Tolsma. Cardiovascular responses were also addressed in other papers. It should be noted that NO₂ was highly correlated with some components on PM, especially EC, in this dataset.

Extrapolation of the underlying incidence data for emergency department visits to other areas could be impacted by socio-economic considerations. For those without health insurance or the ability to pay for healthcare, ED visits are more frequent as

those with other options seek them out before visiting an ED.

3. What are the views of the Panel regarding the adequacy of the assessment of uncertainty and variability?

See responses to the above. It should be noted that the statistical significance of results is impacted by the presence of other pollutants in the model. The Peel et al. results particularly found ozone to be of concern for respiratory endpoints.

.

Characterization of Health Risks (Chapters 7, 8, 9):

1. Based on conclusions in the final ISA regarding airway responsiveness, we have expanded the range of potential health effect benchmark values to include 0.1 ppm. Do Panel members have comments on the range of potential health effects benchmark values chosen to characterize risks associated with 1-hour NO₂ exposures?

There appears to be inconsistencies in the document about benchmarks or exposure levels utilized.

2. To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized?

See above comments. I also await the results of risk analysis for human clinical study data.

3. A focused risk assessment has been conducted for emergency department visits in Atlanta, GA. To what extent are the assessment, interpretation, and presentation of health risk results technically sound, clearly communicated, and appropriately characterized? What are the views of the Panel on the approach taken and on the interpretation of the results of this analysis?

See above comments.

Minor comment:

P. 45, l. 9: "three"