UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

November 14, 2007

EPA-CASAC-08-001

Honorable Stephen L. Johnson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Subject: Clean Air Scientific Advisory Committee's (CASAC) Consultation on EPA's *Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment* (September 2007 Draft)

Dear Administrator Johnson:

The Clean Air Scientific Advisory Committee (CASAC), augmented by subject-matter-experts to form the CASAC Oxides of Nitrogen Primary NAAQS Review Panel, met on October 24-25, 2007 and has completed its consultative review of EPA's *Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment (*September 2007 Draft*)*. The CASAC uses a consultation as a mechanism for individual technical experts to provide comments to guide the Agency on technical issues early in the development of a document, before the first draft is ready for peer review. Panel members offered oral comments at the meeting and their comments are reflected in the official minutes of the meeting. Written comments provided are attached to this letter. A consultation is conducted under the normal requirements of the Federal Advisory Committee Act, which include advance notice of the public meeting in the Federal Register.

There will be no formal report from the CASAC as a result of this consultation, nor do we expect any formal response from the Agency. We look forward to conducting a peer review of the first draft for the Exposure and Risk Assessment document as part of the primary oxides of nitrogen National Ambient Air Quality Standard review.

Sincerely,

/Signed/

Dr. Rogene Henderson, Chair
Clean Air Scientific Advisory Committee

Attachments

Attachment A: Roster of CASAC Oxides of Nitrogen Primary NAAQS Review Panel
Attachment B: Compilation of Individual Panel Member Comments on EPA's *Integrated Science Assessment (ISA) for Oxides of Nitrogen – Health Criteria* (First External Review Draft, August 2007, EPA/600/R-07/093)

Attachment A: Roster of CASAC Oxides of Nitrogen Primary NAAQS Review Panel

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

CHAIR

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

CASAC MEMBERS

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

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

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

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

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

Dr. Jonathan M. Samet, Professor and Chair of the Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

PANEL MEMBERS

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

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

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

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

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

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

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

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

*Unable to participate in the October 25, 2007 Consultation

Dr. Edward Postlethwait, Professor and Chair, Department of Environmental Health Sciences, School of Public Health, University of Alabama at Birmingham, Birmingham, AL
Dr. Richard Schlesinger,* Associate Dean, Department of Biology, Dyson College, Pace University, New York, NY
Dr. Christian Seigneur, Vice President, Atmospheric & Environmental Research, Inc., San Ramon, CA
Dr. Elizabeth A. (Lianne) Sheppard, Research Professor, Biostatistics and Environmental & Occupational Health Sciences, Public Health and Community Medicine, University of Washington, Seattle, WA
Dr. Frank Speizer, Edward Kass Professor of Medicine, Channing Laboratory, Harvard Medical School, Boston, MA
Dr. George Thurston, Professor, Environmental Medicine, NYU School of Medicine, New York University, Tuxedo, NY
Dr. James Ultman, Professor, Chemical Engineering, Bioengineering Program, Pennsylvania State University, University Park, PA
Dr. Ronald Wyzga, Technical Executive, Air Quality Health and Risk, Electric Power Research Institute, Palo Alto, CA

SCIENCE ADVISORY BOARD STAFF
Dr. Angela Nugent, Designated Federal Officer, 1200 Pennsylvania Avenue, NW 1400F, Washington, DC, Phone: 202-343-9981, Fax: 202-233-0643, (nugent.angela@epa.gov)

*Unable to participate in the October 25, 2007 Consultation

Attachment B: Compilation of Individual Panel Member Comments on EPA's *Integrated Science Assessment (ISA) for Oxides of Nitrogen – Health Criteria* (First External Review Draft, August 2007, EPA/600/R-07/093)

Comments from CASAC Oxides of Nitrogen Primary NAAQS Review Panel on EPA's *Draft Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment* (September 2007 Draft)

Comments from Mr. Ed Avol

Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment (Draft September 2007)

General Comments:
The document provides a useful road map for how the Agency will proceed on the Risk Assessment. If the plan is to only provide a Tier I assessment (air quality characterization) and attempt to argue that insufficient information exists to assess exposure, I believe the Agency will find its own credibility and level of commitment questioned. The annexes provide a wealth of information about the current state of knowledge regarding NO2, and most reasonable and objective reviewers will conclude, I believe, that sufficient information exists to perform the Tier II assessment, and to seriously consider the Tier III assessment. The modeling approaches can provide us with guidance if they are applied appropriately, and we should move forward. Continuing to vacillate and wait for complete and perfect information before deciding that there is sufficient data to proceed (which will ultimately end with an estimate and range, anyway) does not serve the public health or the public's interest.

It would be helpful to have a listing of Abbreviations and Acronyms in this document, to which the reader could refer for clarification.

Specific Comments:
P6, Sec 3.1, para2, line 1 – "Several tools would…" should be "Several tools will…"

P6, Sec 3.1, last paragraph discussing evaluation of uncertainties: This discussion is well-intentioned but not well-constructed. What are the objective criteria by which the exposure assessment will be determined to be worthy of a qualitative or quantitative assessment? How will the magnitude of uncertainty (minimal/moderate/maximal) be assigned? Does a rating of "minimal" (which I would think would be the starting point for every evaluation) lead to qualitative or quantitative determinations? How about two "minimals" and one "moderate" in the matrix of uncertainties, or other possible combinations? And what about over and under-estimates – are over-estimates going to be viewed as more conservative and therefore less uncertain, or vice versa? It is difficult to see how this proposed process will lead to a logical, credible determination, based on what is provided here. Staff may well have a clear understanding and process in mind, but that procedural clarity has not been effectively communicated in writing in the document.

P7, Sec 3.2, para2, last sentence – How will "…those commuting on roadways and persons who reside near major roadways…" be incorporated into the modeled population?

P8, Sec3.3, para 1, first sentence – "All available ambient monitoring data collected since…1995…will be used as is." Presumably what is meant is that all *quality-assured* ambient air monitoring data collected since 1995 will be used?

P8, Sec 3.3.1, para1 (regarding the selection of CMSAs for evaluation) - Presumably some tabular summary will validate this selection of cities, but why Atlanta, Philadelphia, and Chicago over New York, Phoenix, and Denver? Some additional and transparent justification for CMSA selection should be provided.

P13, Sec 3.3.1.3, para 1, last sentence – This summary claim of "…insignificant to limited contribution…" of biomass combustion and ETS toNO2 personal exposure is an over-simplification and over-interpretation of what is presented in the referenced Chapter of the ISA. Please review the referenced chapter and re-evaluate the accuracy of this summation.

P.15, Sec 3.3.3, para 1 (Ambient NO2 measurement), last sentence – In areas like Los Angeles, where significant reductions in NO2 in the past decade have only recently resulted in achievement of NAAQS compliance, the assumption that sources present in the past are the sources present now is almost certainly a poor one; some sources are no longer present, and engine/boiler/source emission reduction controls have changed substantially to achieve emission reductions. How will this be addressed or handled?

P15, Sec 3.3.3, para 3 (Spatial Representativeness), line 6 – Low spatial correlations could be the result of several circumstances other than the presence of local sources (for example, topographical intrusions such as canyons, hills, or slopes between sampling locations leading to local variations in wind direction or wind speed).

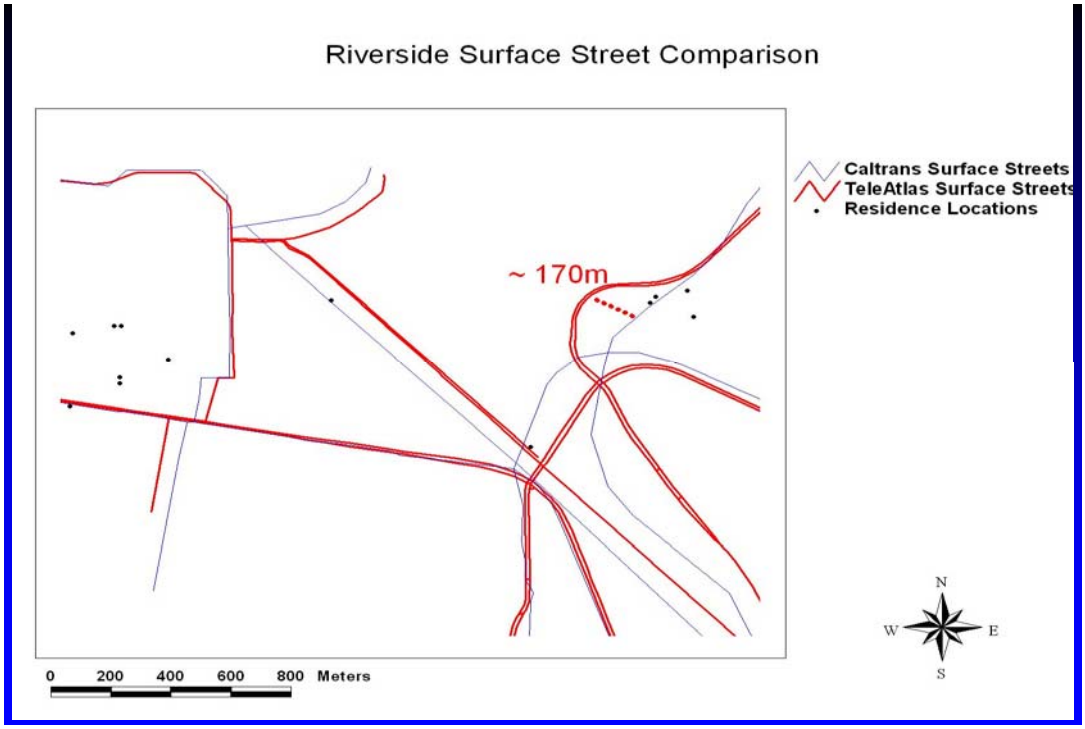
P16, Sec 3.3.3, para 1 (Roadway to Ambient Monitor Relationship), lines 14-19 – Is the implication here that NO2 is a "…reactive pollutant…" and will tend to have a lower I/O ratio? This assertion should be compared to more recent information about in-vehicle measurements. It is my understanding that based on the available information, NO is higher in the passenger compartment (due to the fresh emissions from combustion exhaust being drawn into the vehicle compartment), and that NO2 is somewhat elevated over ambient (reflecting on-roadway conditions), but that NO2 is not as high as near off-roadway (because there has been insufficient time for NO to oxidize to NO2.

P16, Sec 3.3.3, para 1 (Roadway to Ambient Monitor Relationship), last sentence) – Some qualifier must be missing from this statement, because this seems to directly contradict the earlier explanation made in the justification of Equation 3.

P17, Sec3.4 Tier II Screening-Level Exposure…, para 1, bullet 3 – "…factors that contribute to lessened personal exposures to ambient NO2…including time spent indoors and indoor vehicles…" – Doesn't the recent in-vehicle measurements suggest in-vehicle NO2 is somewhat elevated?

P17, Sec3.4 Tier II Screening-Level Exposure…, para 2, lines 6 thru 8 – Is there some protocol for when to apply one of these approaches or the other? Presumably one or two hourly gaps could be filled in using interpolation between valid values at the ends of the missing gap, but this approach would lead to incorrect values if gaps included morning or afternoon traffic hour peaks (since it would not capture or re-construct the peak structure).

P17, Sec 3.4.1 Short-Term Exposure Approach, para1, "…TIGER ROAD network…" If this refers to the road structures based from the highway transportation files, there may be some issues with road placement accuracy, compared to commercially available Tele-Atlas road files. In working with the road files in Southern California to locate streets and residents' homes for Children's Health Study-related research, the transportation files were demonstrated to be occasionally mis-located by 100 meters or more compared to the Tele-Atlas files and the actual location of the roadways. This variation can be critical when considering near-road pollutant exposure (see figure below), given the decay of pollutants with distance from roadways.



P31, Sec 4.1 Risk Assessment Scope Overview…, para2, last sentence – Failing to assign some risk estimate to long-term NO2 exposures runs the risk of not protecting public health from the more potentially more serious and persistent health effects (from long-term, low-level exposures). This sounds akin to ignoring the quantification of the impact because we don't yet fully understand it. At the very least, a statement or discussion should be included discussing this.

P33, Sec 4.3 Tier II Risk… - “health responses reported to be related to NO2 include…” lists several health outcomes, but does not include low lung function (from the Children’s Health Study).

P37, Sec 4.4 Criteria for Determining Approach, last bullet- This undertaking is supposed to lead to the Agency’s best efforts to assess the current information regarding NO2 health effects. The suggestion that there might not be enough time (after allowing 14 years to pass since the previous document release) or insufficient resources to accomplish what the Agency is charged to do is simply not credible; this bullet should either be revised or removed.

Comments from Dr. John Balmes

In general, I think the approaches described in the document are appropriate to provide important information that will be useful in the process of review of the NAAQS for NOx. I urge the agency to try to use emergency department (ED) asthma visit data in the risk assessment. This is especially important given the strength of the epidemiologic evidence of an association between NO2 exposure and asthma outcomes. Asthma hospitalization data only capture the tip of the iceberg of asthma morbidity and the addition of ED data allow a greater proportion of the burden of asthma to be assessed. I made this same point during the ozone NAAQS review and was told that for the cities in which the exposure assessment was done, there were inadequate asthma ED data to use in in the risk assessment. For this review, more ED data should be available – for example, California now has ED as well as hospital discharge data available state-wide.

Comments from Dr. Ellis Cowling

Individual Comments on the *Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment*

My comments are organized below in response to each of the several Charge Questions posed in Karen Martin's September 2007 transmittal letter to Angela Nugent.

Air Quality Considerations

1. Do the Panel members generally agree with using historic air quality data (e.g., pre2000) in certain analyses as a reasonable approach to simulating air quality scenarios with higher NO_2 concentrations, given that current ambient air quality concentrations are lower than the current standard?

Yes, I agree that historical data is a reasonable approach even though some of the historical air concentration measurements may be higher than current ambient air concentrations.

2. Based on the low estimated contribution of policy-relevant background NO_2 to overall ambient NO_2 levels, staff is considering a proportional (i.e., linear) approach to adjusting air quality to simulate just meeting potential alternative NO_2 standards that are lower than current air quality concentrations. Do the Panel members have comments on adopting a proportional approach to simulate just meeting more stringent alternative air quality standards?

Although I am surprised that the contribution of policy-relevant background is as low as it is currently estimated to be, I see no great problem in using a proportional method of adjustment. I have no additional comments to add.

Exposure Analysis

1. In considering the exposure analysis broadly:
a. Do Panel members have any comments on the general structure and overall three-tier approach that staff plans to use for the exposure analysis? Are the criteria that staff plans to use for deciding whether to conduct a Tier II or Tier III analysis clear and appropriate?

The three tier approach seems reasonable to me and the criteria suggested by staff also seem reasonable.

b. Have the most important factors influencing exposure to NO_2 been clearly accounted for and described?

My only major concern is to know whether, and if so, how indoor exposures will be considered and evaluated.

c. The draft plan describes the basis for and selection of population groups of interest (i.e., children, asthmatics (children and adults), and the elderly) for which NO_2 exposure estimates are to be developed. Do Panel members generally agree with the groups of interest identified in the draft plan?

The suggested population groups seem very reasonable to me.

2. In considering the Tier I exposure assessment:
a. Do Panel members agree that an exponential model is appropriate for estimating expected exceedances of short-term health effect benchmarks based on long-term annual average air quality?

I have no experience on which to base an informed judgment in response to this question.

b. Do Panel members agree with the approach to enhance NO_2 air quality data by accounting for the influence of roadway emissions?

Yes, this approach seems reasonable to me.

3. In considering a potential Tier II exposure assessment:
a. Do Panel members agree with the combined emissions/dispersion modeling approach to estimate short-term (hourly) on- and near-roadway NO_2 concentrations?

I have no experience on which to base an informed judgment in response to this question.

b. Is the proposed use of time-location-activity diary data reasonable for estimating short-term exposures for population cohorts?

I have no experience on which to base an informed judgment in response to this question.

c. Do Panel members agree with the use of HAPEM6 to estimate long-term exposures (annual average) and the approach to account for on- and near-roadway NO_2 concentrations?

I have no experience on which to base an informed judgment in response to this question.

4. In considering a potential Tier III exposure assessment:
Do Panel members generally agree that developing individual exposure profiles through the use of APEX is reasonable and appropriate to estimate both short- and long-term NO_2 exposures?

I have no experience on which to base an informed judgment in response to this question.

5. Do Panel members have any comments or advice regarding the general approach to addressing uncertainty and variability in each Tier of the exposure assessment as described in the draft plan?

I have no experience on which to base an informed judgment in response to this question.

Health Risk Assessment

I have no experience on which to base an informed judgment in response to any of these several Health Risk Assessment questions.

Comments from Dr. James Crapo

NOx Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment.

The document provides an appropriate plan for carrying out both and Exposure and Risk assessment for NOx. The tiered approach is appropriate and well laid out in the document. A substantial number of studies have been published during the past decade which suggest that NOx exposures are an important contributing factor in creating adverse health effects from air pollutants at levels of exposure that are substantially below the current NAAQS. The form of the standard is an important question to include in the risk assessment. Current data suggest that short term peak exposures may be more important than long term average exposures as currently regulated in the NAAQS. There are also unique populations at risk both in terms of disease (eg. Asthma patients) and proximity to peak NOx exposures (Those living near roadways or traveling on them frequently).

I feel strongly that both tier I and Tier II of the proposed risk assessment should be done. As part of these assessments I concur with the proposed focus on short term exposures at levels of NO2 in the range of 0.2 tp 0.3 ppm. The tier II assessment to provide data on incidence of expected adverse outcomes will be essential for the subsequent consideration of a recommended change in the NAAQS. I also concur with the recommendation that outcomes for both hospitalizations and ER admissions be considered for asthmatic patients.

Finally it will be important to assess the magnitude of confounding in the risk estimates for NOx by co-exposures to particulates and ozone. The agency should explore ways to include this in their models of exposure and risk assessment.

Comments from Dr. Douglas Crawford Brown

Review of the Draft Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment

Doug Crawford-Brown

This review follows the Charge Questions for the chapter on Exposure. A general comment is that I approve of the overall methodology to the extent it is specified in the document. The outline of the methodology comports with past Agency practice and has the potential to generate the kinds of variability and uncertainty characterizations of risk needed for a rigorous setting of a NAAQS for NO2. However, the devil is in the details, and this document does not lay out very clearly how the detailed computational steps will be performed. It mentions the kinds of models and databases that will be considered, and I generally agree with these, but the real question is how they will be employed. This is especially true for the uncertainty analysis, where I think there is a lot of work still needed (not necessarily in this plan) to determine how the more qualitative and quantitative uncertainty results will be combined into an overall measure of uncertainty that can also serve to guide future research. Still, as a plan, this one is reasonable so long as the Agency can decide how to treat the on-going evolution of sources when depicting future exposures and risks..

One other general comment is that the Agency should consider how to balance an assessment based on individual rights, which would focus on setting a NAAQS that protects some upper percentile of the distribution of risk, with one based on cost-benefit analysis, which would focus on the entire variability distribution. The former might be easier to do, and could be a back-stop approach should the complete variability distribution (Tier III) prove infeasible.

1a. Is the three-tier approach appropriate, and are the criteria for deciding whether a given tier is needed clear and appropriate? Yes, I like the three tiered approach. I would just caution that Tier I, while providing some useful insights, is unlikely to yield a scientifically rigorous basis for setting a NAAQS. I would instead view Tier I as producing only a decision as to whether a new NAAQS is needed at all. If the answer is yes, then Tiers II or III would be required to actually generate that NAAQS. And I am confident that the methodologies and databases exist to allow at least Tier II. I did not find the criteria well specified, and kept wondering throughout how a decision would be made to move to Tier II or III. After several readings, I am not certain what the Agency would need to see in Tier I to motivate it to move to Tier II, and from II to III. There are qualitative criteria given, but I don’t know how these would relate to any specific quantitative results.

1b. Have the most important factors influencing exposure been accounted for? For the most part, yes. The largest problem remains in determining how the activity patterns, and mode choice for travel, will be used to relate ambient air levels to personal exposures.

These methods have been employed for other pollutants, however, and so I am confident the Agency staff can obtain at least a first approximation to these issues. On this one point, it will be important not to get too caught up in trying to characterize intersubject variability of risk too exactly, because human movement in an exposure field is an inexact modeling effort at best.

1c. Is the selection of population groups of interest correct? Yes, with the caveat that there should be special focus on groups that intersect these criteria.

2a. Is the exponential model in Tier I justified? I think any model here is only a rough approximation, and the exponential one is as good as any other. As I mentioned previously, it is important not to get caught up in too much detail here, since individuals will tend to average out this spatial curve as they move about.

2b. Is the approach to enhance roadside NO2 concentrations appropriate? I don't know much about this topic, and so cannot comment on it. But it is clear to me that something does need to be done to produce this enhancement, and also to consider in-vehicle exposures.

3a. Is the combined emissions/dispersion modeling approach in Tier II justified? As a general approach, the answer is yes. I am, however, skeptical of the ability to perform such calculations at a refined spatial scale. This will be especially true in road canyons. It will be necessary, therefore, to use the modeling results only as averages over significant geographic areas (not below a census track or block group).

3b. Is the proposed time-location-activity dairy approach correct? This will be a state of the art approach, although it will be difficult to get the kind of spatially accurate estimates of ambient air concentration needed to make a refined diary approach really worth the effort. There will, however, need to be some thought given as to the level of effort put into the diary approach, and not put in more effort than is justified by the spatial resolution of the ambient air field.

3c. Is HAPEM6 the correct approach? Yes, this is the model I would have selected. It has been employed successfully by Agency staff in the past, although the weaknesses noted during the NATA process should be reviewed.

4. In Tier III, is APEX reasonable? My answer here is the same as in 3c. APEX is a good approach, and one with which Agency staff have some experience, but just be sure to match the effort to the level of spatial resolution of the ambient air field.

5. What is our advice on uncertainty and variability in each Tier? Here there is a lot of work still to be done. I agree with the approach of having both qualitative and quantitative aspects. I would not try to force everything into a quantitative framework. The best one can do is a series of conditional uncertainty and variability statements: that conditional upon a certain set of scenarios, or modeling approaches, or databases, or corrections to the data, the following quantitative uncertainty and variability distributions

are obtained. These U and V distributions can be generated for each combination of scenario, modeling approach, etc, and then an overall judgment of uncertainty and variability developed from expert judgment based on these quantitative distributions. But I recommend this for Tiers II and III, not Tier I (where the uncertainty should be more qualitative and where variability should be treated by examining reasonably maximally exposed individuals rather than producing an actual variability distribution.

On a related note, the plan does not yet specify very well how model validation will be preformed. This is an important step required by the uncertainty analysis, and so needs to be rounded out a bit.

Comments from Dr. Terry Gordon

The plan is well written and the tiered approach is appropriate for the task. I will comment only on the health portion of the risk assessment. The conclusion that the adverse respiratory effects are the strongest health findings appears to be valid and substantiated by the ISA. The exposure indices for the majority of the adverse functional and symptomatic effects are, as documented in the ISA, associated with the short term averaging times. The adverse children's lung growth findings, however, although more recent, should also be included in the risk assessment and would be associated with longer term exposure indices.

Comments from Dr. Dale Hattis

Comments on the Risk Assessment Plan

The overall approach for the risk assessment is described as follows:

"health risk will initially be assessed through the identification of concentration levels associated with adverse health effects, termed potential health effect benchmarks. These.. will then be used to determine how often air quality concentrations or estimated exposures exceed concentrations associated with adverse health effects…."

This seems a rather indirect approach that needlessly economizes on helpful theoretical model-building. I think EPA should essentially discard the evident hope that only a "Tier 1" analysis will be sufficient. What is needed are a set of estimates of the entire population distribution of likely exposures[1] and corresponding distributions of population sensitivity to various health effects. These two distributional inputs could then be used to develop estimates of the current burden of adverse health associated with the current exposure distribution, and the capability to estimate how the burden would change with hypothetical changes in the exposure distribution or with possible changes in the NAAQS or other regulatory standards or feasible non-NAAQS technical measures (e.g. standards for auto emissions). The paragraph goes on to say that "an additional characterization of risk may involve use of concentration-response functions…" In my view it is not a question of whether this level of analysis will be needed. It is certainly needed to support the technical and policy choices that EPA needs to make in seriously considering the effects of various options to revise and restructure the NOx NAAQs. The EPA authors need to immediately start their analysis by going about the business of constructing these exposure/response functions, with due cognizance of the need to quantitatively represent uncertainties in the functions used to estimate health endpoints of various types from the various sources of available scientific information.

p. 9, equation 1.

This exponential equation is not discussed in terms of theoretical mechanisms. I am prepared to believe that distributions of concentrations by exposure time are likely lognormal, but it is not clear that this is the basis of equation 1 or how equation 1 is in fact derived from this basis. Equation 1 is simple enough to use, but there should be some comparative testing with data to show it really works for existing NO2 data in the sense of being free of systematic distortions in the incidence of exceedances out to levels that are very far from the mean.

[1] This does not necessarily correspond to the distribution of concentrations/"exposures" at regulatory monitoring sites, as people may live in locations that are differentially represented by the monitoring sites, in addition to the distortion discussed in the document between ground level locations of people and the elevated locations of the air monitors.

p. 10, Table 2. The occurrence of zero's in this table, rather than fractional values below 1, seems unwise and potentially misleading. (As an aside—the assertion that there are absolutely zero places in the U.S. that exceed the current standard of 0.53 ppm annual average also seems dubious.) The equation provided cannot yield true zero incidences.

In general I question the whole "exceedance" basis of the key calculations that seem to be aimed at. This framework is probably derived from an implicit threshold theory of the incidence of effects as a function of concentration, and threshold theories seem to have little support in the existing epidemiological data. Another presumption seem to be that long term effects, if any, depend on short term episodes of relatively high concentration. This assumption does not appear to be supported by either empirical observations or theoretical analysis. Rather, I thank what would be more useful is a distributional expression of the total fraction of time spent at various levels of exposure for the population as a whole and for various at groups that are at risk because of either unusual susceptibility or residence in locations with various levels of annual average concentrations. Many of the exposure-response observations seem equally well analyzed

Comments from Dr. Donna Kenski

Comments on Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment
Donna Kenski
October 22, 2007

Air Quality Considerations: Using historic data to simulate scenarios with higher concentrations seems reasonable, as does the proposed choice of a linear approach to adjusting data to lower concentrations. The proposed list of CMSAs did not include New York, which the ISA indicated had the highest mean NO2 concentrations of selected urban areas with multiple monitors (Table 2.5-1). Presumably it will show up in the identification of additional locations of interest?

Exposure Analysis: The 3-tier approach is satisfactory. The important factors influencing exposure have been accounted for. This plan emphasized traffic exposures far more than they were discussed in the ISA, which I thought somewhat neglected this source, so that's a definite improvement. The groups of interest are appropriate.

The choice of exponential model is probably okay, although it would be helpful to see what other approaches were considered and to have some comparative assessment—the McCurdy report is not readily available. Was survival regression considered? How does the change in variance over time (apparent from Fig. 1) affect this model? It is not clear from the text why the predictive equation for each location is lumping all monitors together when, in some locations, significant siting differences exist that will impact the number of exceedances. Why not include a site variable in the model? As above, a comparison of various models or additional rationale for this particular one would be helpful.

Health Risk Assessment: The approach outlined here seems fine. In particular the proposed method of characterizing uncertainty and variability is conceptually appealing. The actual implementation of the Tier I/Tier II risk assessment may uncover issues not dealt with in this document, but it seems like a reasonable approach that can be modified as needed and especially as the data require.

Comments from Dr. Timothy Larson

Comments on NO2 Health Assessment Plan

I have several general comments on the exposure assessment portion of this plan. The multi-tiered approach is a reasonable one, moving from more general to more specific in the exposure assessment. The Tier I approach will provide a reasonable ranking of urban areas for further consideration. However, I have concerns about the Tier II approach. In particular, the use of near road gradient algorithms and Gaussian plume models from line sources will not capture the actual traffic related gradients in many urban areas. The reason for this is the presence of buildings and associated street canyons. There are many urban areas where this is an important factor. It would be useful to identify the presence of canyons as part of this screening procedure prior to using 'flat world' gradients and models. The Danish AirGIS system has this capability if information is available on building footprints and approximate building heights. It would also be useful to develop information on the vertical distribution of personal residences in these same urban areas, given the importance of this parameter. Inclusion of the above factors in a Tier III SHEDS type model seems promising, but only if the Tier II screening is done properly.

Comments from Dr. Kent Pinkerton

Review comments for the Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment

Kent E. Pinkerton, Ph.D.
University of California, Davis
Center for Health and the Environment

To assess risks and exposures using a tiered assessment approach for the level of analysis required and the anticipated utility of the results is a highly logical process, especially in the face of possible future limited resources and budget constraints.

Exposure estimates to compare to potential health effect benchmarks to 1) estimate the number of individuals experiencing exposures of concern and to 2) estimate the range of exposures above levels of concern are appropriate and laudable. Since epidemiologic data appears to be the major driver to establish health effects, it is also important to better define whether uncertainties in the degree of health effects observed are due to NO_2 or an associated co-pollutant.

Since at the present time rare excursions of NO_2 above the current NAAQS occur in the nation, yet numerous health effects due to NO_2 exposure have been reported in the literature, it is highly likely this tiered assessment approach will need to be applied well beyond Tier I assessment. Tier II is a critical and needed parameter, especially for NO_2 exposures to allow for screening-level exposure assessments to establish the relationship between ambient concentrations, local sources and human exposure.

It is my opinion and recommendation that exposure assessment for NO_2 include both short and long-term measurements of ambient concentrations through routine air quality monitoring and modeling analysis. The identification of uncertainties in exposure estimates is also essential to determine.

The populations to be modeled which include children (normal and asthmatic), asthmatic adults and the elderly are the proper groups. It may become essential in the future to further determine the influence of gender and genetic predisposition to respiratory disease as well.

In the Tier I air quality characterization, how were the 5 cities of Los Angeles, Houston, Atlanta, Philadelphia and Chicago chosen? Some justification for city selection would be good. Although it is understood motor vehicles, electric utilities and industrial combustion processes represent the major sources of total NO_2 emissions, why totally exclude rural and areas of high agricultural activity? For example, in figure 2 of the document what is the contribution of agriculture to off-highway emissions of NO_2?

A nice example is provided for Tier I air quality characterization in Table 2. However, it appears Los Angeles would be the only city to experience exceedances in NO_2 levels,

based on the current standard. Therefore, how useful would this model be for other portions of the country to explain potential health effects associated with NO_2 exposure? Perhaps it is important to clarify this model can be adjusted to deal with lower NO_2 levels should the air quality standard be changed to provide greater health protection.

A clear explanation of both short-term and long-term exposure approaches to be implemented in Tier I and II exposure assessment is provided in the document. The Decision Flow diagram for Tier II screening, as well as the basic data required to estimate the numbers of person occurrences of short-term exposures in Tier II exposure assessments are provided and extremely helpful.

Again, the explanation provided for Tier III refined exposure assessments is very helpful to better understand the approach to be used, generated outcomes, as well as variability and uncertainty factors that may be encountered and handled.

Figure 5 is excellent in providing an overview of the entire tiered assessment process. Using a tiered approach as outlined seems very reasonable and highly appropriate to insure the proper assessment of exposure levels to NOx.

Under the overview (4.1) for risk assessment scope and methods, one of the goals of the NO_2 risk assessment is to estimate the number of people exposed at or above potential health effect benchmarks associated with NO_2 exposures at levels just meeting the current standard. This goal could be more specific by estimating subgroups such as children, those with asthma, the elderly and socioeconomic classes.

It is important to clearly indicate what constitutes sufficient scientific data to develop population-based health risks for health effect endpoints in at-risk population groups.

Under the overview (4.1) it is not clear why the EPA would not develop risk estimates for NO_2-related effects associated with long-term NO_2 exposures. Although the evidence is not strong, it has been described as “suggestive” for long-term health effects associated with NO_2 exposure. Mobile, stationary and indoor sources of exposure can clearly be long-term.

Under Tier I health effect benchmarks, susceptible populations composed of asthmatics and allergen-sensitive individuals also factor in children and gender-based differences.

The inclusion of baseline data for emergency department visits and respiratory-related hospital admissions for candidate US locations in Tier II risk assessment to enhance risk assessment seems logical and desirable.

Comments from Dr. Edward Postlethwait
Comments on: Scope and Methods for Exposure and Risk Assessment

1. The draft appears to provide a solid foundation for conducting the planned exposure and risk assessments. It is clear that thorough and in depth considerations were incorporated.

2. Although it might have been dealt with elsewhere, it would seem appropriate to have addressed the issue of the potential for confounding in determining exposure/response relationships based on the current NOx measurement approaches. The issue of PRB and long range transport was addressed on page 4, with the available information suggesting that contributions were negligible. Because NO_2 values represent {NOx – NO}, some effort to speciate the various NOx and how much non-NO_2 NOx (e.g. N_2O_3, N_2O_5, etc) might contribute to the overall NO_2 values would be beneficial in terms of characterizing uncertainties. This may be critical when/if clinical and/or indoor studies are compared to environmental exposures if atmospheric NOx contains appreciable amounts of species other than NO_2.

3. On page 13, indoor source influences suggest that gas cooking etc may contribute to NO_2 exposures but biomass combustion (word stoves, fireplaces) appeared to be discounted under proper ventilation conditions. Would not ventilation conditions also affect the other indoor NO_2 sources?

4. An uncertainty apparently not addressed was measurement accuracy across monitoring stations. This becomes critical when attempting to utilize disparate population-based studies to predict health outcomes based on predicted exposure concentrations. The current chemiluminescent approach for measuring NO_2 can be affected by numerous factors and thus it would seem important to be able to estimate measurement error.

5. On page 22 it is stated that "The ME describes the physical location of an individual, allowing for direct contact with the immediate surrounding air that contains a homogenous pollutant concentration." It is assumed this statement refers to NO_2 per se rather than the admixture of co-occurring pollutants wherein all pollutants would be assumed to be present in equivalent concentrations.

6. On page 25, the document refers to ME being calculated using "a mass-balance." This should be clarified since on a chemical reaction and/or diffusion basis it is unclear how the use of mass balance is intended.

7. Only very limited attention was paid to issues regarding biological plausibility and NO_2 acting as a surrogate (ie., pg 34). If health predictions are to be constructed, then it would seem critical that they be placed into an appropriate biological context, especially since NO_2 is generated endogenously via peroxidase activities. Furthermore, it was not clearly evident how co-pollutants would be dealt with in interpreting especially epidemiologic studies wherein frequently there is limited consideration of the dose/effect

relationships unique to individual pollutants. In some studies relative toxicities are poorly considered. Because ozone, for example, is significantly more toxic than NO_2, one would anticipate a need to go beyond simple considerations wherein all pollutants are treated equally relative to exposure concentrations.

8. It is assumed that the impact of uncertainties will be quantified via model output confidence intervals although it was not made specifically clear. Clarifying this issue and discussing the potential magnitude of uncertainty impacts would strengthen the document.

Comments from Dr. Armistead Russell

Review of EPA NO2 Scope and Methods

I am generally pleased with the scope and methods as laid out. It appears, and I hope this to be the case, that it is building upon and building further, EPA's other exposure and risk assessments for reviewing the NAAQS. At the end of one or two more pollutant reviews, it should be almost a well oiled machine (though one that continually improves and considers the unique aspects of the pollutant under consideration).

In regards to its application to nitrogen dioxide, one of the first question that arises is that the ISA considers more than just NO2, but nitrogen oxides in the broad sense (not just NO and NO2). Does the Scope and Methods also have to consider such (e.g., at least consider what the response might be if the determination is that one should look at other components or a sum of components)?

While I generally find that their approaches for assessing the distributions of NO2 exposures are viable and at the level that is appropriate. One could always do a better job, but it is not apparent that for the task at hand it is necessary, with one exception at present. They use an exponential decrease in NO2 going away from a road. They should use the exponential decrease in NOx going away from the road, and then use an appropriate method to split NOx between NO and NO2. In an oxidant limited situation, this could be significant. Also, this will allow them to more explicitly account for changing NO2:NO ratios in the emissions, and assess the overall sensitivity to that split.

Another comment is to try to identify up front the broad levels at which the standard might be set and do some exploratory analyses to show how EOC will very, and the primary sensitivities. While, in the end, the panel will be interested in uncertainties and variabilities, some assessment early on about the sensitivities will be quite useful.

Some other specifics:

In eq (3), the m should be found using linear regression, not as a ratio.

k in eq (2) is not a rate constant. One could call it a dispersion constant, or the like.

Carrying on my comments from the ISA: The monitor uncertainty is overemphasized, and I do believe, mischaracterized. Further, if one is using epidemiologic study results, that bias is built in.

Fairly early on in the process, the results from the exposures and risks in the five cities should be put in perspective of the broader population.

In replying to the given questions:

1. Do the Panel members generally agree with using historic air quality data (e.g.,

pre2000) in certain analyses as a reasonable approach to simulating air quality scenarios with higher NO_2 concentrations, given that current ambient air quality concentrations are lower than the current standard?

   a. **Answer:** It is necessary to know exactly when and how this would be used, but is probably fine. A specific concern is that the older data may have a different NO:NO2 split due to different ozone levels and a different NO:NO2 split in the emissions. The data should be corrected for this if older data is used, and should also be corrected for this when considering future scenarios (this may be a small difference, and if they can show this, great, and then move on).

2. Based on the low estimated contribution of policy-relevant background NO_2 to overall ambient NO_2 levels, staff is considering a proportional (i.e., linear) approach to adjusting air quality to simulate just meeting potential alternative NO_2 standards that are lower than current air quality concentrations. Do the Panel members have comments on adopting a proportional approach to simulate just meeting more stringent alternative air quality standards?

   a. **Answer:** Do you mean proportional or linear? I would prefer linear, though it is recognized there is little difference in this case.

<u>Exposure Analysis</u>:

1. In considering the exposure analysis broadly:

a. Do Panel members have any comments on the general structure and overall three-tier approach that staff plans to use for the exposure analysis? Are the criteria that staff plans to use for deciding whether to conduct a Tier II or Tier III analysis clear and appropriate?

Answer: Yes. (The approach is fine.) EPA should compare and contrast their approach to that used for other pollutants, and document why different methods are used. Again, use each review to make the exposure and risk assessment a more systematic, documented and turn-key. One could see that in about three years (a couple more pollutants) that a system much like that used for air quality modeling is used such that with relatively little effort exposures, risks, variabilities, sensitivities and uncertainties can be calculated, and the system as a whole has been intensely reviewed such that staff need not spend such effort, and the community is more comfortable with the results.

b. Have the most important factors influencing exposure to NO_2 been clearly accounted for and described?

Answer: Not totally… The large role of indoor sources on NOx, and how that gets converted to NO2, needs a bit more work. This issue probably should be picked up more in the ISA as well. Also, the role of NOx in forming and destroying ozone feeds back in to converting NOx to NO2. Further, the discussion here should also deal with the co-occurrence of other pollutants of concern.

c. The draft plan describes the basis for and selection of population groups of interest (i.e., children, asthmatics (children and adults), and the elderly) for which NO_2

exposure estimates are to be developed. Do Panel members generally agree with the groups of interest identified in the draft plan?

Answer: Yes.

2. In considering the Tier I exposure assessment:
   a. Do Panel members agree that an exponential model is appropriate for estimating expected exceedances of short-term health effect benchmarks based on long-term annual average air quality?

Answer: This is fine as long as the model is tested and the appropriate measures of performance are given.

   b. Do Panel members agree with the approach to enhance NO_2 air quality data by accounting for the influence of roadway emissions?

Answer: See discussion above.

3. In considering a potential Tier II exposure assessment:
   a. Do Panel members agree with the combined emissions/dispersion modeling approach to estimate short-term (hourly) on- and near-roadway NO_2 concentrations?

Answer: Yes, as long as the model is evaluated and performance documented.

   b. Is the proposed use of time-location-activity diary data reasonable for estimating short-term exposures for population cohorts?

Answer: Yes, as long as the model is evaluated and performance documented.

   c. Do Panel members agree with the use of HAPEM6 to estimate long-term exposures (annual average) and the approach to account for on- and near-roadway NO_2 concentrations?

Answer: See discussion above (in regards to NO:NO2 splits).

4. In considering a potential Tier III exposure assessment:
   a. Do Panel members generally agree that developing individual exposure profiles through the use of APEX is reasonable and appropriate to estimate both short- and long-term NO_2 exposures?

Answer: Yes, as long as the model is evaluated and performance documented.

5. Do Panel members have any comments or advice regarding the general approach to addressing uncertainty and variability in each Tier of the exposure assessment as described in the draft plan?

Answer: Provide, early on, results of some sensitivity analyses. Do not overestimate uncertainties going in.

Comments from Dr. Jonathan Samet

I write to provide brief comments on the draft version of this document dated September 2007. As indicated in my verbal comments during the October 24-25 CASAC meeting, I found the plan to generally be acceptable in its present form, given the status of the ISA. The document covers the approach that will be used in the three tiers for exposure assessment and for health effects estimation. I have two comments:

- With regard to exposure to NO_2, indoor sources remain substantial contributors. The document is rather general in its approach to describing how microenvironmental concentrations will be estimated. There is a substantial literature on indoor NO_2 concentrations in the presence of various sources that needs to be covered.

- The plan lists the health outcomes that may be considered in the risk assessment, based on the first draft ISA. As indicated in my comments with regard to the ISA, the criteria for finding associations to be causal are not yet sufficiently developed. In general, the health outcomes considered in the risk assessment should be those considered as causal in the ISA. A specific statement to that effect is needed, on the assumption that the plan's authors are in agreement. The current draft plan does not specifically reflect this criterion.

Comments from Dr. Christian Seigneur

Comments on the Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment- Draft - September 2007.

Christian Seigneur
Atmospheric & Environmental Research, Inc.
San Ramon, CA

The three-tier approach for exposure assessment and the two-tier approach for risk assessment appear to be logical ways to proceed. The various steps of each approach are described with sufficient detail for the reader to understand the technical approach and the sources of the data to be used.

QA/QC: One aspect which is not articulated in the document is the Quality Assurance/ Quality Control (QA/QC) procedures that will be followed by EPA. As the assessments proceed to the higher tiers, there will be some very large amounts of data being treated and one must ensure that the proper QA/QC procedures are in place to avoid input or calculation errors.

NO_2/NO speciation – Equations 2 and 3 on pp. 11 and 12: The use of particulate emission control devices on diesel vehicles typically leads to a greater fraction of NO_2 in the NO_x emissions. Such a change in the NO_x speciation for mobile sources could lead to stronger NO_2 spatial gradients near roadways as the NO_2/NO_x ratio will increase at the roadway but the NO_2/NO_x ratio at background sites, which is driven mostly by atmospheric chemistry, may not change. The implication is that the spatial gradient obtained from historical data may not apply (see Equation 2 on p. 11 and associated text). How will EPA address this possible change in the relationship as the vehicle fleet evolves over time?

Estimates of NO_2 concentrations, p. 18: EPA proposes to use the steady-state Gaussian dispersion model AERMOD to calculate the NO_2 concentrations near roadways. AERMOD is a dispersion model that was designed for point sources (Cimorelli et al., "AERMOD: An dispersion model for industrial source applications – Part 1", *J. Appl. Meteorol*., **44**, 682-693, 2005) and which has been evaluated with data from point sources (Perry et al., "AERMOD: An dispersion model for industrial source applications – Part 2", *J. Appl. Meteorol*., **44**, 694-708, 2005). Emissions from roadways differ from those from point sources as vehicle traffic induces some additional turbulence. The use of a simple chemical scheme to account for the rapid titration of NO by ozone to form NO_2 appears appropriate here (although it is not clear what is meant in footnote 8 on p. 18 by "simple reaction rate constant"). However, it is unclear why EPA would want to use a point source dispersion model that is not designed for roadway emissions when roadway dispersion models (such as CALINE4) are available. CALINE4 has been subjected to performance evaluation with measurements made near roadways (Benson, "A review of the development and application of the CALINE3 and CALINE4 models", *Atmos.*

Environ., **26B**, 379-390, 1992) and would seem more appropriate for use here, particularly if AERMOD has not been evaluated for near-roadway estimates.

Example calculation of Table 4: It is not clear how the in-vehicle concentrations are calculated. One person appears to be in a vehicle at more than 75 m from the road but the concentration within the vehicle does not appear to be a function of the distance from the road. Is the NO_2 concentration within the vehicle assumed to be constant regardless of the location of the vehicle?

Comments from Dr. 'Lianne' Elizabeth Sheppard

Comments on the Scope and Methods Plan for the Exposure and Risk Assessment

Summary comments:

- Overall this is a very thoughtful and structured approach that should help inform policy. The tiered approach to exposure and risk assessment seems to be generally reasonable. *However*, modifications to the risk assessment tier structure and additional process and detail are necessary.
- To the extent feasible, *all* criteria should be specifically stated in advance. This should reduce the workload by defining a clear path. It will also make the decisions in the process more transparent.
- Clarify whether the tiers of each assessment are purely conditional on the results of the previous tier or whether they use different information to address details that overlap. If tiers are modified to be a pure progression (i.e. conditional on the results of the previous tier), this may reduce effort.
- Throughout the document the word "would" is used when "will" is more correct. This is after all a plan, not a hypothetical plan.
- In order to help assure the process is open and transparent, all tiers of the exposure and risk assessment need to be covered in the risk and exposure assessment document. In order to make the document as transparent as possible, this policy should be followed even if the final judgment is that the data are insufficient to conduct a specific tier of the assessment. This policy should be stated in the introduction.

Section 2: Shouldn't this section be folded into Section 3? It seems premature to be discussing simulated air quality data when the purpose of the simulation hasn't been stated.

Section 3: My suggested general improvements for this section are to clearly state criteria and to make equations more explicit by adding indices.

Section 3.1: Clearly define each tier in the overview section. Add the tier numbers to Table 1. Clarify whether the information used in each tier is conditional on the data, information, or choices made in the previous tier.

Section 3.2: Why are no population groups defined based on exposure? I suggest adding people living or working near roads.

Section 3.3:

- Clearly state the intended use of the air quality characterization. Different simplifying assumptions will be realistic for different purposes (e.g. prediction of long-term averages vs. traffic-dependent hourly-resolved exposure). Without the intended use stated, it is difficult to evaluate the objectives of the analysis (paragraph 2). For instance, why would the analysis be limited to areas of potential concern, and what are the criteria for "potential concern". As another

example, with regard to a statistical model (objective 3), clarifying all purposes of the analysis can be followed by identification of reasonable simplifying assumptions, thus reducing the modeling effort.

- It is also necessary to distinguish long-term from short-term metric objectives. This distinction needs to be revisited *throughout* the section.

Section 3.3.1:

- The first sentence is good, but now this summary statement needs to be made clear.
- The second sentence appears to be missing a word at the end.
- As an example of clarifying "aggregating data", it should be stated that the objective is to create a single daily (hourly?) time series over space of monitors that are similar. Note that criteria are needed for "similar".
- The criteria for the selected cities are generally listed, but the reader is not informed why those cities were selected.
- Shouldn't site characteristics be included in the list of criteria used to identify additional locations? This is alluded to with the motor vehicle traffic density criterion, although the reference to "by location, not monitor" is cryptic.
- State how the aggregation will be done and what are the criteria for including monitors in the aggregation. Make it clear whether this is temporal or spatial aggregation, or both. I don't understand the purpose of all the statistical tests that are planned and what criteria will be used to determine if additional aggregation is appropriate. (p 8-9).
- p 9: The first full paragraph confuses me. What are the purpose and the outcome of the comparisons within and between locations? What data are to be used?
- p 9: Please add indices to all variables in all equations and define these indices!! Are these data indexed in time by year, day, or hour? What are the spatial indices – site within location?
- p 10: This document is very short on specifics. For instance, *how* will "regression models, parameters, and respective concentration exceedance estimates" be compared?
- p 10: I don't understand how the two parts of the sentence fit together: "The regression model is highly dependent on the prevalence of concentration exceedances, justifying the aggregation of particular (and similar) locations."
- Footnote 4 suggests a valid year could have an entire season missing.

Section 3.3.1.1:

- Equation (2) is a general equation. None of the parameters have values. Will they be estimated from data? What time scale is being considered? Add indices.
- Are C_v and C_b data or predictions? How are they obtained?
- How will the equation (2) result be used to derive (3)? Add indices.
- The entire plan for this section is wide open and subject to many interpretations.
- Why is the goal to obtain on-road estimates of NO_2 instead of characterizing NO_2 as a function of distance from road?
- Note that "on-road NO_2 concentrations" are *predictions*, not data.

Section 3.2: Descriptive statistics should include measures of spread as well as central tendency.

Section 3.3:

- Paragraph 2: Restate sentence to say the tiered approach uncertainty assessment is done with the goal of identifying the best supported quantitative analysis.
- Paragraph 2: Presumably the "identified components are, in a broad sense, also relevant to subsequent exposure analyses" because this tier I analysis is the input to the tier II analysis. Correct? Please state clearly.
- Add "Choice of NO_2 as the index compound" as one of the components of uncertainty.
- Temporal representativeness: State what the "temporal profiles" are. Are these estimated hourly average air quality over a multi-year period for a given spatially aggregated location with specific spatial features?
- Spatial representativeness: The purpose of the predictions really matters when deciding how to proceed with limited spatial data. State the purpose. What prediction equations are being referred to? What kind of correlations will be evaluated?
- Monitor to exposure representativeness: Why is personal exposure even being mentioned in the Tier I estimates? Isn't it more important that the AQ characterization is done in locations that are representative of population exposure to ambient concentration?

Section 3.4:

- p 17 line 2 – add "ambient-source" to describe the possible lower bound estimate.
- Gas stoves are an important factor in greater personal exposure and should be listed in the example to indicate home characteristics will also be considered.
- In doing spatial interpolation of exposure, it will be important to only include monitors that are representative of usual population ambient source exposure (as opposed to those highly influenced by local sources that won't apply to the entire census tract or adjacent tracts). I am concerned that some factors could be counted twice if the local source monitors aren't removed first, since local sources will be added in with the planned adjustment.
- Following the previous comment, I suggest discussing locations represented by ambient monitors as a function of monitor siting criteria and/or GIS covariates.
- Insert the word "predicted" to clarify the complete set of concentrations won't necessarily be data.

Section 3.4.1:

- Organizationally, why not define the on-road concentrations as the 0 m road proximity class?
- Why do indoor sources need to be identified as important contributors to ambient air concentrations to be considered?
- Figure 4: Why can there be significant on-road concentration but little elevated concentration at <75m? (see the first site)

Table 4: While it is clear that the total column is a weighted average, it is completely unclear what the average concentration total row means (particularly given the numbers provided).

Section 3.4.1.2:

- There is an assumption that the spatial and temporal contributions to NO_2 are relatively simply related, i.e. temporal estimates from one location can be linearly transformed to get estimates at a new spatial location. Ideally this assumption should be checked. At least it should be discussed.
- I think uncertainty in model structure can be evaluated with sensitivity analyses.

Section 3.4.2:

- Does the term "long-term exposure" mean "annual average exposure"?
- Add an introductory paragraph and start a new section subheading for the material already at the beginning of this section.
- Equation (4) needs indices for time, space, and microenvironment type. Clarify the range and units of the indices.
- Note a different approach to roadway contribution is being used here.

Section 3.4.2.1:

- Give an equation to show the relationship described in the first sentence.
- On what time scale will the additional exposure metric be calculated? (p 24 top)

Section 3.5.1:

- p 26 first sentence first full paragraph: The approach to predicting hourly NO_2 from monitoring data and dispersion models is a major research topic in itself. The approach taken here is relatively simple, and thus it should be mentioned as a limitation and source of uncertainty.
- Should in-vehicle estimates be separated by road type?

Section 3.5.3: Instead of relying solely on informed judgment, why not compare estimates from plausible models formulated differently?

Section 3:6: Define number of peak concentrations. Discuss Figure 5 in more detail, and possibly move it to the beginning of Section 3.

Section 4:

- I think the risk assessment needs to be reorganized to have 3 tiers. The first tier should be a qualitative assessment of the health evidence. This will list and consider all important health effects based on human and animal studies. Not all of these can be used for benchmark calculations or quantitative risk assessment, but it will be important to review them all first and get a sense of the scope of the risk qualitatively. Then a narrower list will be used for the second and third tier assessments. Not only does this proposed new tier structure allow for better progression in the treatment of the health results, but it also elevates the

importance of the qualitative risk assessment in the document and protects against it being treated as an afterthought.

- Criteria for acceptable outcomes to use in the risk assessment (as well as other aspects of the RA such as choice of city for the analysis) needs to be specified in advance for each tier.
- Criteria for even conducting a quantitative risk assessment (the proposed second tier) need to be specified in this document in advance of the risk assessment.
- The quantitative risk assessment should move towards incorporating an integrated uncertainty assessment as an integral part of the process. This would extend uncertainties in underlying assumptions into the estimated risks by incorporating a probabilistically weighted range of assumptions into the risk estimates and uncertainties estimated. This will be most appropriate for the highest tier of quantitative assessment.

Section 4.2: State the criteria for selecting health effects to be used for the benchmark analysis.

Section 4.2.1: State the planned health effect benchmark levels or criteria for selecting these levels.

Section 4.2.3: Third paragraph: In addition, a distribution of benchmarks could be applied rather than sensitivity analyses of a set of single values. I'm confused by the end of this paragraph (starting "From a directional perspective…"). My understanding suggests either the wording is backwards or I am confused. Perhaps an example will help the reader's comprehension.

Section 4.3: The criteria for what is sufficient information to develop credible exposure-response relationships *must* be stated. I note there is information about such criteria in later subsections. Restating the criteria in another form, such as a list, may be helpful.

Section 4.3.1.1 (and 4.4): I believe that the last two additional factors (2: availability if sufficient C-R data in locations relevant to the US and 3: availability of baseline incidence data) should be given less weight in the decision to proceed. Both can be evaluated with sensitivity analyses.

Section 4.3.3. The analysis should move towards an integrated representation of all the most important uncertainties. Sensitivity analyses are typically done by altering one assumption at a time. This approach can be useful for identifying the most influential of the uncertainties. It is possible statistical uncertainty (e.g. estimated variance of a parameter) will be the least influential; rather the more influential uncertainties may be due to the form and structure of the risk assessment such as the choice of model, rollback assumptions, etc. The integrated analysis would extend the Monte Carlo analysis to jointly assess a range of model assumptions by putting probabilities on different assumptions, such as several choices of the concentration-response model coefficient and/or its functional form. The difficulty with the integrated approach is there is another layer of assumptions and reasonable ranges that need to be elicited and incorporated into

the analysis. The advantage is the quantitative risk estimates will have additional sources of uncertainty incorporated directly into the result. Moving towards an integrated analysis of these uncertainties will give a fresh perspective of the range of possible effects and the probability of effects at different levels.

Section 4.5: Restating an earlier comment: Summary of the health effect data should precede the quantitative risk assessment as the first tier risk assessment. The qualitative risk assessment should do more than just provide the "broad context for the quantitative risk estimates". It should be the foundation.

Response to charge questions: Exposure assessment

1. *Broad exposure analysis evaluation*: a. The 3-tier structure is fine. In the final document all tiers should be covered, even if the only text is justification for why the tier was not conducted. b. See comments regarding Charge questions 1-3 of the ISA. c. Add populations at higher risk due to high exposure.
2. *Tier I assessment*: a. Without seeing an analysis of the data, I can't comment. I suspect sites that differ by important NOx sources (e.g. distance to road) may fit this model more or less well. b. The roadway emissions approach seems to be overly simplistic, but may be the most reasonable approach given this is a lower tier assessment.
3. *Tier II assessment*: a. Again, I would like to see an analysis of the data before I can form an opinion. I am concerned that near road monitors should not be used for the spatial interpolation model. Regardless, there may be too few monitors over space to do a reasonable spatial interpolation. It is not clear why the outcome generated should be a measure of exceedance, or how the criteria for exceedances will be chosen. b. This appears reasonable, although it treats people-hours as exchangeable. For this purpose that appears to be acceptable. c. No comment without further data analysis.
4. *Tier III assessment*: APEX is a reasonable approach. Presumably the ambient monitoring predictions are conditional on the Tier II exposure assessment. In fact, coming up with the "modified and interpolated hourly NO2 concentration measurements" is a major research topic in itself. The approach taken is relatively simple and thus needs to be recognized as a source of uncertainty.
5. *Advice on general approach to addressing uncertainty and variability*: Our best collective scientific understanding should help us prioritize sources of uncertainty and variability and focus on the most important ones. For those, I suggest comparing estimates for plausible models formulated under different assumptions (sensitivity analyses) and working towards extending the uncertainty analysis to jointly assess multiple different sources of uncertainty simultaneously (integrated uncertainty analysis).

Response to charge questions: Risk assessment

1. *Overall structure and two-tier risk assessment approach:* I think there should be 3 tiers for the risk assessment with the first tier being a qualitative risk

assessment. The criteria for selecting the endpoints should flow from the ISA, particularly the Chapter 5 summary rewritten to use open and transparent criteria. I think the Pilotto et al (2004) study estimates are based on a strong enough study design that even though they haven't been directly replicated, they should be considered for the risk assessment. I don't consider the same location or no outside US location criteria to have sufficient weight to limit a Tier II risk assessment.

2. *Tier I risk assessment:* a. Criteria for identifying benchmarks must be clearly stated in advance. b. I think the lung growth effects are also important to evaluate. c. While the Pilotto et al (2004) study is not a controlled study, it is an intervention study and a much stronger design than a pure observational study.
3. *Tier II assessment:* a&b. See section 4.3 comments.
4. *Addressing uncertainty and variability:* An integrated uncertainty assessment should be considered. See section 4.3.3. comments.

Reference: Additional thoughts on how to proceed with the integrated uncertainty analysis may be found in this report:

Committee on Estimating the Health-Risk-Reduction Benefits of Proposed Air Pollution Regulations, National Research Council. Estimating the Public Health Benefits of Proposed Air Pollution Regulations. NRC 2002 (see Chapter 5)
http://www.nap.edu/catalog.php?record_id=10511

Comments from Dr. Frank Speizer

Nitrogen Dioxide Health Assessment Plan: Scope and Methods for Exposure and Risk Assessment (September 2007 draft)

Answers to Charge Questions (paraphrased)

Submitted by Frank E. Speizer

Date: October 17, 2007

Air Quality Considerations

1. Use of historic air quality data pre 2000.

This is not an unreasonable use of historical data. Figure 1 on page 5 suggest a marginally significant decline in the annual average NO2, but the variation seems to have changed substantially with a marked drop in the 90%tile level starting around 1997. In table 2.5-1 in ISA on page 2-52 spatial variations are wide in some cities. Thus, for the last 10 years may want to inflate the variance to better take into account the individual city variation.

2. Use of a proportional approach to modeling alternative air quality standards.

I think the same observation made above applies to the use of proportional adjustments. Somehow the drop in the 90%tile values along with the variation across regions (cities) needs to be dealt with. If proportional models works that is fine.

Exposure Analysis

1. Broad considerations.

General Structure. This seems reasonable but I would be disappointed if Staff concluded that they could not get past Tier I. For factors influencing exposure perhaps there needs to be some discussion on how the interaction with Ozone will be handled. In some of the regions there are likely to be competing interaction, with quenching affecting what is being measured and difficulties attributing risk. (This may all come up later). In addition, it might be indicated as to how, at least in a general sense indoor exposure, will be considered. Population groups of interest. If possible I think it would be useful to consider children broken down somewhat differently. The text in section 3.2, page 7, suggests birth to age 18. I think it would be better to consider birth- preschool (near home); 4 or 5 to 9 (local community); and 10-18 (active outdoor physical activity). I recognize that the data may not exist but at least the breakdowns for exposure might be considered. The other grouping seem appropriate, except might want to consider those adults carrying a cardiovascular disease diagnosis as a separate (potentially more susceptible) group.

2. Tier I exposure assessment

For exponential model and accounting for emissions this seems to be appropriate, however, what will need to be discussed later is how this model deals with the time-varying patterns of exposure that might occur as people "move through" their approximate exposures. This forms the basis of the discussion of uncertainties in section 3.3.3. Although the potential issues that might vary exposure and uncertainties are well

described, it is not made clear just how these will be handled. (Perhaps there will not be a variable added to Equation 3 and residential time within x number of meters of a roadway value will only be discussed qualitatively at this level of analysis but such should be stated. Alternatively, if there would be a way to incorporate residential time (or other modifiers of exposure) in the equations that would be useful.
3. Tier II. This is a exceeding well written description of what needs to be done, and if accomplished should satisfy the numbers needed for any risk assessment. As I read through this I am wondering if all the comments above on Tier I are irrelevant as many of the comments above are answered in this section. Therefore should consideration be given to combining the two Tiers into a more expanded discussion, since much of the uncertainty in Tier I and specified again on page 17 are dealt with here . (Leaving it own uncertainties.) In picking the distance to roadways (<75 m, 75-200 m, >200 m) some justification needs to be added on page 18. Particularly since on page 11 the spatial drop off exposure levels gives a range of 200-500m to get to ambient. Note also the footnote on page 18 on age distributions is more in line with my comments above on age groups. What is not clear in the discussion of uncertainty is the how the nature of the monitoring station (residential, commercial, industrial) as well as nature of residence near roadway (single family houses, large apartment blocks) get taken into account, for example in table 4. Will sites that are used for regulatory control at the edge of a factory be excluded? What if all the people in the tract live more than 200m from the monitor? In addition I assume that "fraction of the population in each location" somehow gets factored in when census tract is used (as total population, and age distribution within in each census tract, are not all the same). Again, many of these issues are discussed in the Tier III section, and again it makes me question whether the separation into separate Tiers in imposing more criticism than is necessary. Figure 3 on page 29 outlines the criteria needed, it seems likely that sufficient data are available to proceed, so no more than a descriptive discussion of how staff gets to Tier 3 is needed and they should get on with doing the assessment as proposed.

Risk Assessment Scope and Methods

On page 33 section 4.3 I would recommend to staff that they reverse the order of discussion on the credible exposure response relationships for controlled human exposure studies and the epidemiological studies, particularly sine the section title is Risk Based on Epidemiological Studies. Clear most if not all of the controlled studies have been carried out in normal healthy volunteers; whereas the epidemiological studies are in general observations on free-living population groups that obviously contain people with vary levels of risk. Judgment on risk assessment should be made on the latter group with the controlled human exposures experiments mostly designed to assess and understand potential mechanisms for the risk observed in free-living populations.

The plan as outlined a two Tier effort, like for the exposure assessment seems somewhat arbitrary as to whether it is called a two tier effort or a logical progression in gathering the data necessary (and I believe from the draft ISA) available to do all that is proposed. The short term exposure assessment is well documented to move forward, particularly for the respiratory outcomes described. With regard to the long term assessment particularly for hospitalizations and mortality by sub-regions, this may have to await the assessment of the draft ISA.

With regard to the criteria for determining the approach to tier ii on page 37, I accept that the thinking of these steps are necessary to get to the data but as indicated above I believe there are sufficient evidence, particularly for the short term effects, that doing the risk assessment for respiratory outcomes should be straight forward. However, I totally reject the placement of the last bullet on page 37, in section 4.4 as a criteria for doing what is needed. There should be sufficient resources to complete the task in the next 15 months. It is a priority for the task to be done in the next 15 months and if resources must be borrowed from other activities to get it done, that should be the decision.

Comments from Dr. George Thurston

COMMENTS FROM GEORGE D. THURSTON

Comments on Nitrogen Dioxide Health Exposure and Risk Assessment (September 2007)

General Comments

This plan successfully outlines the scope and approaches for, and highlights key issues in, the estimation of population exposures and health risks posed by NOx under: 1) existing air quality levels; 2) upon just meeting the current NO_2 primary NAAQS, and; 3) upon just meeting potential alternative standards under consideration by the Administration.

The main concern I have is in the selection of the key CMSA's for evaluation. More justification for these choices needs to be presented. In particular, it is surprising to me that New York City, which has a high concentration of those whom the ISA is pointing to as most affected (e.g., poor children with asthma living near roadways). I think consideration of both LA and NYC would give balance to the document, as LA has a high concentration of NO_x exposures, while NYC includes a high concentration of especially affected individuals. The other three cities selected should provide more "typical" U.S. urban exposures/populations.

I also feel that the U.S. EPA should consider emergency department (ED) visits to the extent possible. These data are becoming more widely available on a routine basis throughout the U.S., and may more accurately reflect the extent of effects, given recent cost-saving efforts to treat problems like asthma in the ED, rather than as a hospital admission.

Specific Comments

P8, Sec 3.3.1, par. 1 (regarding the selection of CMSAs for evaluation) – Why Atlanta, Philadelphia, and Chicago over New York, Phoenix, and Denver? More justification for CMSA selection should be provided.

P8, Sec. 3.3.1, par. 2. I feel this list of criteria should include some involving exposed populations in each CMSA in addition to the exposure concentration profiles discussed here (perhaps %poor, ssthma prevalence, % living within 200 yards of a highway, % minority, or the like),

P35, Sec 4.3.1.2, par. 3. I agree with the balanced discussion of the multi-pollutant issue presented here. My only addition is that, when possible, the models considered should consider no more than two pollutants at a time to help clarify the role of each pollutant relative to each of the others. When more than that are included, it becomes much harder to sort out real effects from multi-collinearity effects.

Comments from Dr. George Thurston

General Comments

This plan successfully outlines the scope and approaches for, and highlights key issues in, the estimation of population exposures and health risks posed by NOx under: 1) existing air quality levels; 2) upon just meeting the current NO_2 primary NAAQS, and; 3) upon just meeting potential alternative standards under consideration by the Administration.

The main concern I have is in the selection of the key CMSA's for evaluation. More justification for these choices needs to be presented. In particular, it is surprising to me that New York City, which has a high concentration of those whom the ISA is pointing to as most affected (e.g., poor children with asthma living near roadways). I think consideration of both LA and NYC would give balance to the document, as LA has a high concentration of NO_x exposures, while NYC includes a high concentration of especially affected individuals. The other three cities selected should provide more "typical" U.S. urban exposures/populations.

I also feel that the U.S. EPA should consider emergency department (ED) visits to the extent possible. These data are becoming more widely available on a routine basis throughout the U.S., and may more accurately reflect the extent of effects, given recent cost-saving efforts to treat problems like asthma in the ED, rather than as a hospital admission.

Specific Comments

P8, Sec 3.3.1, par. 1 (regarding the selection of CMSAs for evaluation) – Why Atlanta, Philadelphia, and Chicago over New York, Phoenix, and Denver? More justification for CMSA selection should be provided.

P8, Sec. 3.3.1, par. 2. I feel this list of criteria should include some involving exposed populations in each CMSA in addition to the exposure concentration profiles discussed here (perhaps %poor, ssthma prevalence, % living within 200 yards of a highway, % minority, or the like),

P35, Sec 4.3.1.2, par. 3. I agree with the balanced discussion of the multi-pollutant issue presented here. My only addition is that, when possible, the models considered should consider no more than two pollutants at a time to help clarify the role of each pollutant relative to each of the others. When more than that are included, it becomes much harder to sort out real effects from multi-collinearity effects.

Comments from Dr. Ronald Wyzga

In general the approach appears to be reasonable, but ambitious. The challenge will be to implement this plan with available resources in a way that is scientifically credible and in a way that illuminates the available information to inform further discussions about the NO2 primary NAAQS.

Within this approach there are two points that I want to raise:

The input data for the assessment are subject to uncertainties (multiple studies with different results, missing information, etc.). I would urge the investigators to consider ways to internalize the uncertainties within the analyses rather than to undertake base analyses accompanied by sensitivity analyses. The outputs of such analyses are likely to ranges or distributions of estimates rather than point estimates, but such ranges more accurately portray reality than do point estimates.

Secondly I am disturbed by the approach taken in evaluating the risks and exposures associated with ambient levels of NO2 associated with just meeting the current standard. Since the entire US is currently within compliance, the entire country has NO2 levels below the current standard. To assume under this scenario that the entire country has NO2 levels at the standard will clearly lead to overestimates of risk and exposure, and there is concern that the resulting estimates will be accepted at face value and misinterpreted. I urge the staff to consider ways to confront this issue in a way that more realistically portrays the exposures/risks associated with current and alternative standards.