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 Dr. Ed Avol 2
Comments from Dr. Ellis Cowling 6
Comments from Dr. Dale Hattis 10
Comments from Dr. Timothy Larson 12
Comments from Dr. Donna Kenski 13
Comments from Dr. Armistead Russell 14
Comments from Dr. Christian Seigneur 18
Preliminary comments from Dr. 'Lianne' Elizabeth Sheppard 20
Comments from Dr. Frank Speizer 25

Comments from Dr. 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).

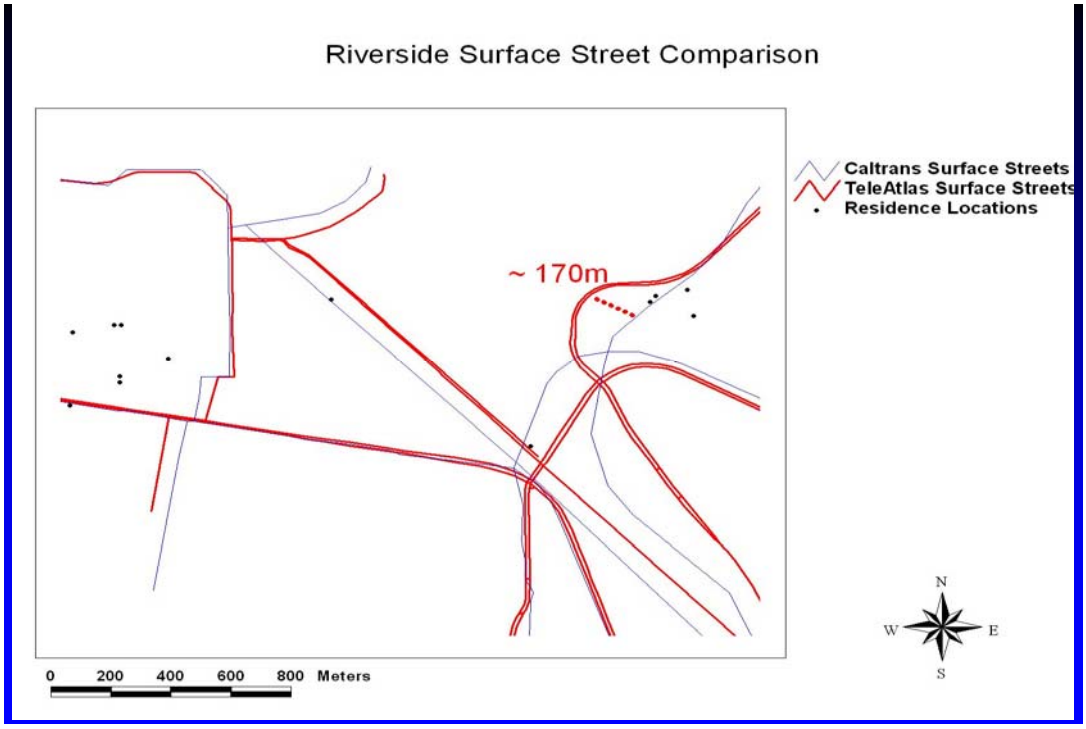
P16, Sec 3.3.3, para 1 (Roadway to Ambient Monitor Relationship), lines 14-19 – Is the implication here that NO2 is a "…reactive pollutant…" and will tend to have a lower I/O ratio? This assertion should be compared to more recent information about in-vehicle measurements. It is my understanding that 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. 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. 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. 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. 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. 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. 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?

Preliminary comments from Dr. 'Lianne' Elizabeth Sheppard

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

Summary comments:

- Overall the tiered approach to exposure and risk assessment seems to be generally reasonable. *However*, modifications to the risk assessment tier structure and additional detail are necessary.
- To the extent feasible, *all* criteria should be specifically stated in advance.
- 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. 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: The 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. 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, a statistical model (objective 3) is only useful if it meets a specific purpose, but the purpose is not stated.
- 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 quite 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 (my third tier, the written second tier) need to be specified in this document in advance of the risk 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 4.5: 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.

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. If there are insufficient resources to accomplish this in the next 15 months than the Administrator deserves to go to jail for not meeting the consent decree!!

Mydoc: RiskAssemnt.NOx.Draft9.07