UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

June 7, 2001

OFFICE OF
THE ADMINISTRATOR
EPA SCIENCE ADVISORY BOARD

Note to the Reader:

The attached draft report is a draft report of the EPA Science Advisory Board (SAB). The draft is still undergoing final internal SAB review, however, in its present form, it represents the consensus position of the panel involved in the review. Once approved as final, the report will be transmitted to the EPA Administrator and will become available to the interested public as a final report.

This draft has been released for general information to members of the interested public and to EPA staff. This is consistent with the SAB policy of releasing draft materials only when the Committee involved is comfortable that the document is sufficiently complete to provide useful information to the reader. The reader should remember that this is an unapproved working draft and that the document should not be used to represent official EPA or SAB views or advice. Draft documents at this stage of the process often undergo significant revisions before the final version is approved and published.

The SAB is not soliciting comments on the advice contained herein. However, as a courtesy to the EPA Program Office which is the subject of the SAB review, we have asked them to respond to the issues listed below. Consistent with SAB policy on this matter, the SAB is not obligated to address any responses which it receives. Responses are due no later than INSERT DATE.

1. Has the Committee adequately responded to the questions posed in the Charge?
2. Are any statements or responses made in the draft unclear?
3. Are there any technical errors?

For further information or to respond to the questions above, please contact:

K. Jack Kooyoomjian, Ph.D.,
Designated Federal Officer
EPA Science Advisory Board (1400A)
US Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460-0001
(202) 564-4557 Fax: (202) 501-0582
E-Mail: kooyoomjian.jack@epa.gov

United States Environmental Protection Agency
EPA Science Advisory Board (1400A) Washington DC
EPA-SAB-EC-ADV-01-00x June 2001
www.epa.gov/sab

NATA - EVALUATING THE NATIONAL-SCALE AIR TOXICS ASSESSMENT FOR 1996 - AN SAB ADVISORY

AN ADVISORY BY THE EPA SCIENCE ADVISORY BOARD (SAB)

DRAFT REPORT - DO NOT CITE OR QUOTE

F:\USER\PTILLERY\NATApubdft6601.wpd

June 6, 2001 Review Draft

EPA-SAB-EC-ADV-01-00X

Honorable Christine Todd Whitman
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460

RE: National Air Toxics Assessment (NATA) - An SAB Advisory on Review of the National-Scale Air Toxics Assessment for 1996

Dear Governor Whitman:

On March 20-21, 2001 the Science Advisory Board's (SAB's) National-Scale Air Toxics Assessment (NATA) Subcommittee (also referred to as the NATA Review Panel) of the SAB Executive Committee conducted a peer review of the Agency's NATA program. The NATA Review Panel conducted this advisory on the initial NATA of the potential health risks associated with inhalation exposures to 32 air toxics identified as priority pollutants by the Agency's Integrated Urban Air Toxics Strategy, plus diesel emissions.

While a number of the elements of this assessment plan have already undergone scientific peer review, the entire assembly of these elements and application of the full assessment approach have not. The Agency has asked the SAB's NATA Review Panel to comment on the appropriateness of the overall approach, including the data, models, and methods used, and the ways these elements have been integrated, as well as to suggest ways to improve these approaches for subsequent national-scale assessments. The advice and insights contained herein are focused on changes that can be made to the

current (1996) NATA, as well as to the future (1999 and beyond) NATA exercises.

The NATA Review Panel met on February 21, 2001 in a public conference call to provide Panel members and consultants (M/C) with the opportunity to clarify the Charge questions (see brief summary below), request any supplemental materials from the Agency, ask questions on materials already received from the Agency, and discuss preparations for a public review meeting of the NATA Review Panel on March 20 & 21, 2001 held in Research Triangle Park, NC. The Panel M/C met in public conference call follow-up technical editing work sessions on April 24th, May 14th and May 25th where no public comments were solicited. The Panel met on June 13th, and solicited public comments on its June 6th, 2001 Public Draft Advisory.

In summary, the Panel would like to commend the EPA staff on their work on the NATA. This effort represents an important step toward characterizing the relationships between sources and risks of hazardous air pollutants.

We agree that the Agency should not use this initial national-scale assessment directly as a basis for regulating sources for air toxics, and that while the Agency's regulatory priorities will be informed by this and other assessments, the Agency should develop risk-based regulations on the basis of more refined and source-specific data and assessments.

The Panel emphasizes an overarching need for continued, improved data collection in support of the assessment. We stress that collection and compilation of air toxics emissions data is of vital importance to the national capacity for environmental health assessment and management, and encourage continued presentation and enhancements of inventory results to the states, industry and other stakeholders. Furthermore, enhanced collection of ambient concentration data and updating of exposure data for air toxics is needed to support NATA objectives. In the absence of widespread measurements, the 1996 NATA relies on modeling to estimate atmospheric concentrations and exposures. An expanded set of measurements is needed to evaluate and develop confidence in the

models, and to provide independent information about spatial distributions and trends over time.

The following findings and recommendations are presented in response to the charge from the Agency.

Question #1 - National Toxics Inventory: The Panel supports the continued development and presentation of inventory results to the states, industry and other stakeholders for their evaluation and input in order to identify errors, encourage more complete reporting, and data quality assurance. Improvements in the National Toxics Inventory (NTI) would be facilitated through the provision of uniform national reporting protocols and rules; the provision of incentives for industry to measure, validate and report their emissions; and the use of visualization tools (e.g., GIS database and mapping programs) for the NTI. The NATA document should provide a clearer presentation of the methods used for data collection, analysis and interpretation within the NTI. Methods for cross-validation of emission estimates and for development of industry-specific emission factors for use in other applications are needed. For a number of metals, such as chromium and nickel, emissions estimates and calculations in subsequent NATA modules should differentiate between important species.

Question #2 - Model Issues: The Panel is concerned about a number of aspects of the current implementation of ASPEN (the atmospheric transport model used to compute ambient concentrations from HAP emissions) and HAPEM (the time-activity model used to compute human exposure from predicted ambient concentrations) within NATA. In fact, the Agency has acknowledged many of these concerns in the NATA document, and some of these difficulties and concerns can and should be addressed for the current 1996 assessment. Other suggested improvements will require a longer-term effort and should be targeted for the 1999 and subsequent NATAs. Specifically, the ASPEN model is not formulated to predict concentrations of compounds present largely as a result of secondary reactions in the atmosphere or affected by larger-scale, regional transport, as is attempted in the current NATA application. Development of more advanced atmospheric transport models able to address secondary formation of air toxics and larger-scale regional transport is recommended for future

NATAs. Also, the current application of HAPEM4 is significantly flawed in terms of its representation of exposure variability.

As currently formulated, the HAPEM model can be implemented only to compute (and report) the median exposure predictions for each census track (and county). While continued development of HAPEM is encouraged, exposure and risk estimates based on simpler transformations (or direct use) of ambient concentrations should be presented in parallel with those based upon HAPEM results. To demonstrate application of ASPEN and HAPEM4 for a case where the models and available data are adequate to provide for reasonable prediction, we recommend that a full-scale analysis of benzene exposures be conducted across the US. Methods development should also begin for the incorporation of other important pathways of exposure for multi-media pollutants, such as the fish ingestion route for methyl mercury and soil ingestion for lead.

Question #3 - Dose-Response Information: The NATA study makes generally appropriate use of available dose-response information, consistent with currently accepted protocols by federal and state agencies. Since toxicity factors for selected compounds may be updated on an ongoing basis, it is important to have a consistent procedure for incorporating these updates into NATA predictions. Since significant uncertainty is present in chemical dose-response factors, no matter which exposures and risk assessment method is used, care should be taken to isolate and separately report these uncertainties from those introduced through the screening assessment procedures specific to NATA. Ongoing improvements to IRIS are a critical part of the general need for improved scientific capabilities for assessing air toxics. When new values are being considered to replace those currently in IRIS, scenario-based assessment should be used to identify the implications of the possible changes for the NATA results. For some chemicals, the procedures chosen for determining dose-response factors do not follow the standard NATA protocol (e.g., for nickel and 1,3 butadiene). Better justification is needed for the use of alternative methods in these cases.

Question #4 - Risk Characterization: The Panel observed that a number of the methods and

terminology used for the characterization of aggregate, cumulative risks in the NATA study are somewhat novel and unique. While the proposed non-cancer risk aggregation approach is of potential value, it is new, and greater clarity is needed on the assumptions in the current NATA documentation to allow the approach to be assessed in a more thorough and rigorous manner. On cancer risk aggregation, The Panel believes that NATA erred in isolating known human carcinogens (Class A) and combining probable (Class B1) with possible carcinogens (B2). It seems more correct and certainly more precautionary for OAQPS to combine the Class A and Class B1 separately from the Class B2 carcinogens. Ideally, risks based on each of the carcinogen classes should be identified along with the total.

For future iterations of NATA, methods are needed to characterize the full distribution of exposure and risk across target populations and subgroups, so that high-end exposures and risks can also be considered. To accomplish this, the implementation of the HAPEM model needs to be improved to adequately reflect the full range of interindividual variability in air toxics exposures. Adequate characterization of the chronic and acute risks from air toxics resulting from inhalation as well as other routes of exposure, and from both indoor and outdoor sources, is also needed for future NATAs.

Public health scientists in state and county agencies have a good sense for the compounds that pose the greatest concerns to people in their areas. The relative risk predictions from NATA should be compared and evaluated against this source of information. In the next iteration of NATA, this type of 'groundtruthing' should be given high priority. In the current NATA, comparisons may be made using existing reports that compile and summarize national or regional risk drivers. Separate, compound-and location-specific exposure estimates should be reported by NATA, and the exposure assessment should be communicated as a major and valuable product of NATA – not just viewed as a component of the risk assessment. Full characterization and discussion of exposure distributions across target populations can help to focus future data collection needs and planning for exposure reduction.

Question #5 - Diesel Emissions: The inclusion of an assessment of diesel emissions in the current NATA is appropriate, but problematic. The lack of an acceptable unit risk estimate for cancer prevents the treatment of diesel emissions in parallel to the other air toxics. Diesel should be treated in a separate, succinct section of the report in which the calculations for assessing exposures and the present knowledge of risks are described clearly. The set of diesel health risks addressed should be expanded to include the concerns for respiratory disease mortality and morbidity generally associated with fine particulate matter (PM).

Question #6 - Uncertainty and Variability: Given the high degree of conceptual uncertainty in the modeling of air toxic emissions, exposures and risks, and the significant gaps in available data for supporting these, the more aggregate, 'top-down' approach for assessing uncertainty proposed in the NATA report is appropriate. However, the current implementation requires significant further work before meaningful results and insights can be obtained. In particular, the methods and supporting information are not yet sufficient to allow the assignment of probability distribution functions for representing uncertainty in each of the NATA components (emissions, fate-and-transport, exposure, and dose-response) and the combination of these to estimate a probability distribution for the resulting prediction of risk. Instead, a scenario-based approach should be used to capture and discuss key conceptual and data uncertainties in the NATA.

Question #7 - Communications: The NATA document reflects a proper concern with the importance of effective communication of results, to encourage a holistic understanding of air toxic risks and the options available for addressing them. The NATA document also addresses the various information needs of decision makers and stakeholders in the EPA, other federal and state agencies, industry, environmental and other interest groups, and the general citizenry. We recommend that the Agency clearly distinguish between those parts of NATA that are well established versus those which are in an earlier, developmental stage. In developing a web page, the Agency should consider use of a hierarchal set of pages to differentiate between information based solely on reports and personal monitoring, information that is based on relatively simple or highly confident model calculations, and

information based on new model developments, where research is ongoing to improve the basis of prediction. These web pages could be color-coded and titled to indicate the degree of confidence in the information. The NATA document also needs an executive summary focused towards a lay audience.

Question #8 - Benefits Analysis: Basis for a Benefits Assessment: The current exposure methodology and results in NATA are not yet ready for use in the national scale benefits analysis required in Section 812 of the Clean Air Act. Once the needed improvements noted within this advisory are implemented, application to benefits assessment can be considered. In particular, a meaningful benefits assessment must consider the full distribution of exposure and risk (not just median values) and should also address acute health effects. If a full distribution of exposure is estimated for an information-rich HAP, such as benzene, as part of the current NATA, then the 812 study could attempt an initial benefits assessment for that HAP, to illustrate the type of analysis that is envisioned for the future.

Question #9 - Future Research Priorities: An extensive research effort should be mounted to address the wide array of the data and model development needs to significantly improve the scientific foundation of future NATA, as well as regulations based on the health risks of air toxics. The needs (addressed in detail in the NATA document) include both fundamental and chemical-specific research and span the whole of the risk paradigm (i.e., emissions, ambient concentrations, exposures, effects, and risks). Because air toxics research has been under-funded by the Agency for so long, considerable new resources are needed. Fortunately, the NATA allows identification of the uncertainties that are inhibiting the development of reliable quantitative assessments so that the new resources could be well-focused. We understand that the EPA ORD is completing a strategic plan for air toxics research, so there is no need for SAB to duplicate this effort (though specific areas of focus for research are identified by the Panel in this report). We recommend that the Agency's research strategy be developed in concert with external experts on the related topics and that the subsequent draft be reviewed by this or a similar Panel.

We would also like to reiterate a critical comment which was made during the SAB's review of the Cumulative Exposure Project (Phase 1) in 1996, which was the genesis of the 1996 NATA . The current NATA Review Panel still believes this comment is extremely relevant today. " We also encourage the Agency to begin examining ways in which environmental data collected for regulatory purposes might be collected in ways that would make these data simultaneously useful for scientific purposes. With some thought, . . . it should be possible to develop improved guidelines for the collection of some environmental data so that it could be used for the dual purpose of assessing regulatory compliance and advancing environmental science in order to improve the future protection of public health."

We appreciate the opportunity to provide advice on this effort. The Agency staff was open, collegial, cognizant of shortcomings in the document, and accepting of the NATA Panel's suggestions. We look forward to the Administrator's response, particularly to the points highlighted in this letter to you.

Sincerely,

Dr. William Glaze, Chair
EPA Science Advisory Board

Dr. Mitchell J. Small, Chair
NATA Review Panel
EPA Science Advisory Board

NOTICE

This report has been written as part of the activities of the EPA Science Advisory Board, a public advisory group providing extramural scientific information and advice to the Administrator and other officials of the US Environmental Protection Agency. The Board is structured to provide balanced, expert assessment of scientific matters related to problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the US Environmental Protection Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use.

Distribution and Availability: This Science Advisory Board report is provided to the USEPA Administrator, senior Agency management, appropriate program staff, interested members of the public, and is posed on the SAB website (www.epa.gov/sab). Information on its availability is also provided in the SAB's monthly newsletter (*Happenings at the Science Advisory Board*). Additional copies and further information are available from the SAB Staff at U.S. Environmental Protection Agency, EPA Science Advisory Board (1400A), 1200 Pennsylvania Avenue, NW, Washington, DC 20460 (Tel. 202-564-4533; FAX (202-501-0256).

ABSTRACT

This advisory provides responses to nine specific questions raised by the Agency to the EPA Science Advisory Board's (SAB) Executive Committee, National-Scale Air Toxics Assessment (NATA) Review Panel as a part of the EPA/Office of Air Quality Planning and Standards (OAQPS) initial (1996) national-scale assessment of risks from exposure to 32 air toxics plus diesel emissions nationwide. The major review meeting took place on March 20 & 21, 2001, with public teleconferences held prior to and following this meeting.

The Panel emphasizes the need for continued and improved data collection in support of the assessment. Collection and compilation of air toxics emissions data is of vital importance to the national capacity for environmental health assessment and management, and we encourage continued presentation and enhancements of inventory results to the states, industry and other stakeholders. Furthermore, significantly enhanced collection of ambient concentration and exposure data for air toxics is needed to support NATA objectives. In the absence of widespread measurements, the 1996 NATA relies on modeling to estimate atmospheric concentrations and exposures. An expanded set of measurements is needed to evaluate and develop confidence in the models, and to provide independent information about spatial distributions and trends over time.

The Panel provided advice and recommendations for the 1996 NATA, as well as for the 1999 and subsequent NATAs. Topics reviewed deal with the National Toxics Inventory (NTI), model issues (specifically on ASPEN and HAPEM4), dose-response information, risk characterization, diesel emissions, uncertainty analysis, communication of results, use in benefits assessment in the future, and future research priorities.

Keywords: hazardous air pollutants, air toxics, monitoring, emissions, transport, fate, exposure, risk, models, ASPEN, HAPEM, NATA

293

U.S. ENVIRONMENTAL PROTECTION AGENCY
SCIENCE ADVISORY BOARD (SAB)
NATIONAL-SCALE AIR TOXICS ASSESSMENT (NATA) REVIEW PANEL FY- 2001

CHAIR

Dr. Mitchell J. Small, Professor, Departments of Civil & Environmental Engineering and Engineering & Public Policy, Carnegie Mellon University, Pittsburgh, PA

SAB MEMBERS*

Dr. Henry A. Anderson, M.D., Chief Medical Officer, Wisconsin Bureau of Public Health, Madison, WI

Dr. Steven M. Bartell, Principal, Cadmus Group, Inc. Oak Ridge, TN

Dr. Calvin Chien, Senior Environmental Fellow, E.I. DuPont Company, Wilmington, DE

Dr. Linda E. Greer, Senior Scientist, Natural Resources Defense Council (NRDC), Washington, DC

Dr. Kai-Shen Liu, Epidemiologist, California Department of Health Services, Berkeley, CA

Dr. Joe L. Mauderly, Director of National Environmental Respiratory Center, Lovelace Respiratory Research Institute, Albuquerque, NM

Dr. Paulette Middleton, Director, RAND Environmental, Inc., Boulder, CO

SAB CONSULTANTS*

Dr. David R. Brown, Public Health Toxicologist, Northeastern States for Coordinated Air Use

Management (NESCAUM), Westport, CT

Mr. Thomas J. Gentile, Chief, Toxics Assessment Section, Division of Air Resources, New York State Department of Environmental Conservation, Albany, NY

Dr. Panos G. Georgopoulios, Associate Professor, Environmental and Community Medicine, UMDNJ-Robert Wood Johnson Medical School, Piscataway, NJ

Dr. Carol J. Henry, Vice President, Science and Research, American Chemistry Council, Arlington, VA

Dr. Jana Milford, Associate Professor, Department of Mechanical Engineering, University of Colorado, Boulder, CO

SCIENCE ADVISORY BOARD STAFF

Dr. K. Jack Kooyoomjian, Designated Federal Officer, US Environmental Protection Agency, EPA Science Advisory Board (1400A), Washington, DC

Ms. Betty B. Fortune, Office Assistant, US Environmental Protection Agency, EPA Science Advisory Board (1400A), Washington, DC

* Members of this SAB Panel consist of the following:

a) SAB Members: Experts appointed by the Administrator to two-year terms to serve on one of the 10 SAB Standing Committees.

b)

u) SAB Consultants: Experts appointed by the SAB Staff Director to a one-year term to serve on ad hoc Panels formed to address a particular issue; in this case, the review of the Agency's National-Scale Air Toxics Assessment (NATA) for 1996 and to provide recommendations for the 1999 and subsequent NATAs.

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

On March 20-21, 2001 the Science Advisory Board's (SAB's) National-Scale Air Toxics Assessment (NATA) Subcommittee (also referred to as the NATA Review Panel) of the SAB Executive Committee conducted a peer review of the Agency's NATA program. The NATA study represents the most current effort by the EPA to provide a nationwide quantitative assessment of health risks associated with the inhalation of 32 priority pollutants and diesel emissions identified as contributing significantly to human exposures and risks in urban areas. The EPA document which is the subject of this review is entitled "National-Scale Air Toxics Assessment for 1996," EPA-453/R-01-003, January 2001.

The Panel found that the draft NATA 1996 document represents an extensive and comprehensive effort to systematically evaluate and link the various components of the risk paradigm relevant to HAP impacts, including emissions, atmospheric transport, human exposure and risk. In the absence of widespread measurements, the 1996 NATA relies on modeling to estimate some elements of the emissions inventory, as well as ambient concentrations and exposures. While some aspects of the current data collection and modeling are advanced enough for confident prediction, others are still highly uncertain. An expanded set of measurements is needed to fully evaluate and develop confidence in the models, and to provide independent information about spatial distributions and trends over time.

As part of our review, we have identified specific areas where the current NATA is especially problematic. Some of these difficulties can and should be addressed for the current 1996 assessment. Others suggested improvements will require a more long-term effort, and should be targeted for the

1999 and future NATA's. In the recommendations that follow in our report, short- vs. long-term targets for implementation are identified.

The development of a nationwide assessment of air toxic emissions, atmospheric transport, human exposure and risk is a daunting task, and the Agency has had to make a number of choices cognizant of the limitations in scientific understanding, available data, and the time and resources available for the assessment. A key choice has involved the selection of the spatial scale of aggregation for conducting the NATA, and for reporting the results. The census tract is utilized as a basis for estimating emissions (at times inferred from information at higher levels of aggregation, such as the county level), predicting atmospheric transport, defining receptor populations, and computing their exposures and risks. The results are then aggregated back up to the county level for reporting purposes. While we agree with this basic strategy for assessment and reporting, there are a number of difficulties that arise in its implementation. The census tract is a good unit for defining the demographic characteristics of receptor populations, but it is not a good geographic unit for air pollution modeling and assessment. In particular, densely populated census tracts are small, while those in sparsely populated areas tend to be large. This tends to misrepresent the allocation of emissions and bias the calculation of representative ambient and exposure concentrations for densely vs. sparsely populated areas. This problem needs to be identified in the current NATA, and addressed in future NATAs through conversion to a regular spatial grid for emissions tracking and the calculation of ambient concentrations, with subsequent conversion back to underlying census tracts for population exposure and risk calculations.

A major finding of the Panel is that parts of the NATA are based on relatively reliable data and/or well-established scientific estimation and modeling methods, while other aspects are based on more limited data and methods that are in an earlier, developmental stage. This applies to all aspects of the NATA, including emissions estimates, estimates of ambient concentrations based on the ASPEN model, estimates of exposure based on the HAPEM modeling system (or, as suggested in our report, other, simpler methods that should be considered in parallel with the HAPEM predictions), and risk estimates requiring the use of toxicity factors based on different amounts of scientific information and

consensus. To help citizens and other users of NATA better understand the differing bases for NATA results, we recommend use of a hierarchical presentation of results that distinguishes between quantities measured or modeled at different levels of scientific development, and with differing levels of available data and confidence.

While we have attempted to provide specific information and recommendations to improve the 1996 and future NATA studies, we recognize that much of the need for improved information applies generally to the field of air toxics and health toxicology, and is not specific to the NATA. When uncertainties and concerns are apparent in the NATA methodology, we have attempted to distinguish between those endemic to NATA and those more broadly applicable across the field of environmental health risk assessment. We also note that we have focused on the general methodology presented in the NATA document, and not the specific values of inputs and parameters used to implement it (though specific examples are identified to be illustrative of apparent problems and areas of concern). The absence of comment on specific emission, atmospheric transport, exposure and toxicity factors should not be construed to indicate Panel review and approval of these values. Separate peer review is required for the specific parameter values and factors used to implement the NATA.

The Panel addressed the following set of nine charge questions, modified through negotiation from those originally proposed by the Agency. The principal findings and recommendations of the Subcommittee applicable to each question follows.

1. Given the nature of the NTI and the methods by which it was developed and reviewed, have available emissions data been appropriately adapted for use in this assessment? Can you suggest improvements to EPA's application of the NTI for use in future initial national-scale assessments?

a) Can you suggest improvements to the treatment of compound classes (e.g., chromium and compounds), given the nature of the information available in the inventory?

b) Can you suggest improvements to the methods used to spatially distribute area and mobile source emissions?

c) Can you suggest improvements to the methods used to specify default point source emission characteristics in lieu of missing emissions data?

The Panel finds that the continued collection and compilation of air toxics emissions data is of vital importance to the national capacity for environmental health assessment and management. Continued presentation of inventory results to the states, industry and other stakeholders is encouraged, in order to identify errors and to encourage more complete reporting and data quality assurance. Improvements in the National Toxics Inventory (NTI) would be facilitated through the provision of uniform national reporting protocols and rules; the provision of incentives for industry to measure, validate and report their emissions; and the use of visualization tools (e.g., GIS database and mapping programs) for the NTI. While disaggregating emissions estimates to census tracks is necessary for subsequent fate-and-transport modeling, continuing to limit the reporting of emissions to the county level is supported.

The NATA document should provide a clearer presentation of the methods used for data collection, analysis and interpretation within the NTI, in comparison to those used for the National Emission Trends [NET] database for criteria pollutants. Methods for direct cross-validation of emission estimates are needed. Additional approaches that do not depend entirely on ambient concentration measurements and models should be pursued. Comparisons of emission inventories for similar point and area source categories across the States should be made using the 1996 NTI. Comparison of emission estimates from state reporting, National Emission Standards for Hazardous Air Pollutants (NESHAP) information collection requests, and TRI information, should be made when these are available. Diagnostic study of relationships between economic activity (e.g., production, employment) for industrial sectors in an area and the emissions estimated for those sectors, can also be to used to identify possible mismatches or outliers. These relationships may also help in the development of industry-specific emission factors for use in other applications.

For a number of metals, such as chromium and nickel, emissions estimates and calculations in the subsequent NATA modules should differentiate between important species (e.g., Cr^{6+} vs Cr^{3+}) wherever feasible.

There is a need to better validate and document methods used to estimate mobile source emissions, especially for non-road mobile sources. In particular, more information should be provided on the methods used to allocate mobile-source emissions to census tracks. Non-road emission estimates should be further checked and validated where possible, since these are predicted to have a significant impact on ambient concentrations, exposures and risks. For on-road mobile sources, state data based on vehicle miles traveled (VMT) and other state generated input data (e.g., average vehicle speed and vehicle fleet mix) should be used to estimate on-road emissions when available on a county basis.

2. Is the approach taken for the geographic aggregation of ambient and exposure concentrations generated by the ASPEN and HAPEM4 models appropriate in light of the limitations of the models, the available emissions data, and the results of the comparisons of ambient predictions with ambient monitoring data?

The Panel is concerned about a number of aspects of the current implementation of ASPEN and HAPEM4 within NATA. Many of these concerns are already recognized and acknowledged in the Agency report and documentation. For the current (1996) assessment, HAPs should be classified to identify (a) those where ASPEN is expected to provide an appropriate basis for analysis; (b) those for which ASPEN is potentially applicable, but still uncertain, and improvements/refinements are needed; and (c) those for which the model is highly uncertain, and use for these compounds is close to, or even beyond, the range of scientifically defensible applicability for ASPEN. This latter group includes chemicals that occur to an important extent as secondary pollutants (e.g., formaldehyde, acetaldehyde, acrolein), and those for which background or regional areal sources dominate (e.g., lead). Furthermore, geographic regions where ASPEN predictions are likely to provide accurate vs. inaccurate predictions should be identified, based on terrain and climatology. For future assessments,

ASPEN capabilities for NATA should include the ability to address seasonal variations in climatology and emissions. For secondary pollutants, ASPEN cannot be utilized in a reliable manner, and high priority should be given to the local-scale adaptation and application of MODELS-3, or a similar model platform, able to simulate nonlinear chemistry for secondary air toxics and address the larger-scale transport processes important for pollutants with significant background concentrations. Because of these limitations of ASPEN, the NATA report likely underestimates concentrations of these secondary contaminants.

The current implementation of HAPEM4 is significantly flawed in terms of its representation of exposure variability. The selection of different individuals within a cohort in the Consolidated Human Activity Database (CHAD) for each day of a simulation over a year greatly suppresses the individual-to-individual variability between simulations. While this might be an appropriate method for estimating the mean or median exposure in a census tract or county, the subsequent presentation with probability intervals is misleading, since it implies that the presented quantiles represent the population exposure distribution across the targeted area. There are three approaches that can be used to address this problem in the short term (ideally, all three options should be evaluated and their results compared). First, model risk estimates based solely on ambient concentrations can be calculated and reported [as done in the current Cumulative Exposure Project (CEP)]. Second, a simple outdoor-indoor correction factor can be introduced to simulate the effects of inter-individual variability in the fraction of time spent indoors and the overall effective penetration factor for each individual's indoor environments. Third, the HAPEM model can be implemented as currently formulated, but only to compute (and report) the median exposure predictions for each census track (and county). As noted elsewhere, hierarchical presentation of results from all three approaches is recommended, indicating information and estimates based on quantities measured or modeled at different levels of scientific development, and with differing levels of available data and confidence.

To demonstrate application of ASPEN and HAPEM4 for a case where the models and available data are adequate to provide for reasonable prediction, we recommend that a full-scale analysis of exposure to benzene, or another well-studied, -monitored and –characterized compound, be

conducted across the US. This would include the development of improved activity pattern selection methods to allow a reasonable simulation of interindividual variability in long-term exposure. These improved methods would then be available for application to other compounds in future NATA studies. Methods development should also begin for the incorporation of other important pathways of exposure for multi-media pollutants, such as the fish ingestion route for methyl mercury and soil ingestion for lead.

3. Has available dose-response information (e.g., different sources of information, a different prioritization scheme) been appropriately used in this assessment? Can you suggest methods that could improve upon the use of available dose-response information?

The NATA study makes generally appropriate use of available dose-response information, consistent with currently accepted protocols by federal and state agencies. Since toxicity factors for selected compounds may be updated on an ongoing basis, it is important to have a consistent procedure for incorporating these updates into NATA predictions. Differences in NATA predictions should be illustrated when current potencies or benchmark dose factors are used vs. different values that may be under consideration or proposed for change. Since significant uncertainty is present in chemical dose-response factors, no matter which exposure and risk assessment method is used, care should be taken to isolate and separately report these uncertainties from those introduced through the assessment procedures specific to NATA. Significant uncertainties in IRIS and other chemical toxicity databases indicate the high priority of ongoing research to update and improve the knowledge base for dose-response assessment of air toxics. For some chemicals (either in the initial screening phase or in the subsequent, more-detailed assessment), the procedures chosen for determining dose-response factors do not follow the standard NATA protocol (e.g., for nickel and 1,3 butadiene). Better justification is needed for the use of alternative methods in these cases.

4. What are the strengths and the weaknesses of the overall conceptual approach to risk characterization used in this assessment? Given the underlying science and the intended purposes of the assessment, can you suggest ways in which the risk characterization could be

improved?

a) *Is the method used to aggregate cancer risks appropriate? The aggregation of carcinogenic risk within two categories, based on weight-of-evidence classifications, is of particular interest.*

b) *Is the method used to aggregate non-cancer hazards appropriate? The summation of hazard quotients within target organs, the categorization of sums by ranges of uncertainty factors, and the inclusion of all target organs (as opposed to only the organs associated with the critical effect) are of particular interest.*

The overall conceptual approach to the risk characterization is reasonable. It generally follows the guidelines and procedures of risk assessment (with exceptions noted later for mixtures). However, as detailed subsequently, some of the key specific elements in implementation of the conceptual approach are not consistent with current assessment guidelines or best practices. In spite of the inadequate scientific foundation for a quantitatively reliable risk assessment of air toxics (due to uncertainties in the existing health and exposure assessments), specific revision to the 1996 NATA could bring the methods used more in line with current best practice, and help to better fulfill the goals of the document. The following discussion focuses on areas of the NATA needing improvement, recognizing the need for a prompt revision for the 1996 NATA and the need to build in as much science quality as possible at this point. In our opinion, considerable effort is needed in making the revisions, but the recommendations are created to be feasible for experts dedicated to the work over the short-term.

We are certain that EPA will seek to incorporate the lessons learned from this evaluation in their development of the 1999 NATA. The most important overarching need for the 1999 NATA is to conduct more comprehensive risk assessments. By that time, IRIS dose-response assessments and the foundation of exposure assessments will have been improved. Future assessments should more

carefully evaluate the literature on observed exposures and health effects, as well as the overall underlying health science and information, rather than relying exclusively on the numerical results of NATA model exposure estimates and the IRIS (or similar database) dose-response values. For example, alternative analyses based on scientifically sound databases should be compared to those based on the routine application of standard models.

The current NATA includes only chronic inhalation health effects from exposure to outdoor sources air toxics. The document is quite clear on this, but the resulting limitations of the assessment need to be more explicitly discussed. Effects from less-than-lifetime exposures and total exposure to air toxics are key issues requiring further evaluation. The 1999 NATA needs to incorporate these issues, especially assessments based on the multiple pathways of exposure to outdoor sources of air toxics.

In the current EPA cancer guidelines, chemicals are classed according to the weight of evidence in support of the inference that they are carcinogenic. The classes for known or suspected carcinogens include:

A: "Known" carcinogens based on sufficient evidence from epidemiologic studies to support a causal association between exposure to the agents and cancer;

B1: "Probable" carcinogens based on limited evidence from epidemiologic studies, but sufficient evidence from animal studies; and

B2: "Possible" carcinogens based on sufficient evidence from animal studies, but inadequate evidence or no data from epidemiologic studies.

Known human carcinogens are summed separately from probable human carcinogens in the NATA document. Probable human carcinogens are lumped with possible carcinogens. This is not conventional. The only difference between the known and probable classes of carcinogens is the extent

of available data from human studies, and human studies of these compounds are relatively rare. Thus, it seems more correct and certainly more precautionary for OAQPS to combine and report the Class A and Class B1 separate from the Class B2 carcinogens. Also, OAQPS should provide an estimate for all types of cancers summed together and then break them out by group. These revised calculations should be feasible for the 1996 NATA.

A Hazard Quotient (HQ, equal to the exposure divided by the RfC) and Hazard Index (HI, equal to the sum of HQs) approach are common means of assessing noncancer risks. As everyone agrees, there is a high degree of uncertainty in this approach. However, the means of doing this calculation in this draft NATA document do not follow current EPA guidelines and are scientifically questionable, and therefore need to be improved.

The HI methodology is commonly accepted for chemicals having a common mode/mechanism of action. In the absence of data, some assessors default to using a common organ (in accordance with EPA mixtures assessment guidelines). The key phrase is in the absence of data. In some cases, chemicals having known different modes/mechanisms were added together in computing an HI (e.g., formaldehyde which produces nasal effects was added to cadmium which produces lung effects through different mechanisms). This needs to be corrected.

The calculation of greatest concern is the target-organ-specific-hazard index (TOSHI). This HI was calculated by taking the RfC for a chemical based upon the critical effect and dose to one organ and transferring this RfC to all other organs affected by that chemical. The RfC is based on the most sensitive indicator of effects, to which conservative uncertainty factors are applied. To take this value and apply it directly to other organs (deemed inappropriate by EPA for the original RfC calculation) is scientifically questionable. If EPA wishes to use a TOSHI approach, it is essential that EPA goes back to the database for each chemical and actually develops TOSHI's with a high level of scientific rigor. Without that effort, they should be eliminated from the document. It is recognized that the IRIS database for many of these substances is out-of-date, but timing considerations for revision of this version of NATA may restrict the TOSHI reevaluation to this IRIS database. Although this would

compound any errors due to the status of the underlying database, it is preferable to the approach used in the current document. It is also important that problems in computing HI's (due to uncertainties in both the methodology and the supporting data) be clearly identified in the text as a significant limitation.

The integration of an exposure assessment with a health assessment is extremely difficult, even under data-rich circumstances. Errors in the exposure assessment (such as the improper characterization of inter-individual variability in estimated population exposures, described above), combined with the significant uncertainty in toxicity (dose-response) factors, can lead to considerable errors in estimated risk, of unknown magnitude and direction. While somewhat more reliable estimates of county mean or median exposures and risks (and the distributions of these across the nation), and resulting inferences about the *relative* (rather than absolute) risk between counties and compounds, may be achieved, considerable uncertainties will remain in these estimates as well. As discussed later in response to Charge Questions 6 and 7, this creates a considerable challenge to the Agency, as how they should characterize and present the uncertainty and confidence that can be placed in the resulting risk estimates. To help characterize the level of confidence that is warranted, the Agency should implement some selective groundtruthing exercises for the predicted exposures and risks for some of the selected air toxics. EPA can identify a data-rich air toxic that would then be used to compare various risk characterization approaches in the 1996 NATA. Benzene could serve as such a test compound, but others should also be considered. The 1999 NATA should include more such comparisons, as well as consideration of different scenarios that would facilitate understanding of whether the uncertainty drivers are the health assessment or the exposure assessment factors and estimates.

The NATA appropriately raises the issue of the importance of understanding risks to children. It furthermore implies that a lack of assessments of children is a matter of concern. In so doing, it does not give EPA credit for already taking actions to address and consider children through the assessment process. For example, all chronic assessments are based on a 70-year lifetime. In addition, the RfC methodology has an uncertainty factor of 10 for sensitive populations (including children). Since childhood is a portion of the 70-year exposure period considered, children are incorporated in the

assessments of *chronic effects.* For example, if a 70-year exposure is likely to cause a chronic problem, a 21-year exposure would likely cause less of a problem (or no problem).

Concern over the need for additional, special consideration and assessment designed to protect children *does* arise when more *acute* health endpoints are considered. Since we do recommend that acute health effects be considered in future NATA's, the data collection, research and assessment activities necessary to develop exposure and susceptibility estimates for children relevant to acute effects should begin now. Until such time that results from these more-targeted efforts are realized, greater uncertainty is likely to be present in both acute and chronic exposure assessments for children (and acute health effects estimates). The current NATA document has addressed some (but not all) of the uncertainties and issues related to children in describing the key data collection, modeling and characterization issues for exposure calculation.

5. Although EPA has concluded that available data are not sufficient to develop a reliable quantitative estimate of cancer unit risk for diesel emissions, it is clear that this pollutant class may be of significant concern in a number of urban settings. The risk characterization in this report includes a discussion of diesel particulate matter to help states and local areas frame the importance of this pollutant compared to the other air toxics. In the context of this assessment, is the discussion in this report regarding making risk comparisons among other air toxics appropriate? Can you provide any suggestions that would improve upon this approach to comparing the toxic health effects of diesel particulate matter with other pollutants?

The inclusion of an assessment of diesel emissions in the current NATA is appropriate. Furthermore, the caveats used in the report to describe the current state of knowledge about diesel particle health risks are reasonable and generally consistent with the latest CASAC findings and recommendations. The exposure assessment is especially valuable. However, the attempt to treat diesel emissions in a fully integrated and step-wise manner, in parallel to the other air toxics addressed in the report, is awkward, and the required frequent repetition of the Agencies "belief statement" that

diesel particles are (or may be) among the most significant health risks among air toxics is not adequately supported in the report. The current status of our knowledge of the risks from diesel emissions should be summarized more clearly in a separate and succinct section of the report, and the calculations used for computing diesel exposures and risks expounded upon in that section. The set of diesel health risks addressed in this section of the report should be expanded to include the concerns for respiratory disease mortality and morbidity generally associated with fine particulate matter (PM).

6. Given the limitations inherent in this preliminary assessment, have uncertainty and variability been appropriately characterized?

a) *Can you suggest ways that the characterization of uncertainty and variability could be improved, made more transparent, or integrated more effectively into the risk characterization?*
Can you suggest methods for quantifying individual as well as composite uncertainties associated with the emissions inventory, dispersion modeling, exposure modeling, dose-response assessment, quantitative risk estimates, and accumulation of risk across air toxics?

Given the high degree of conceptual uncertainty in the modeling of air toxic emissions, exposures and risks, and the significant gaps in available data for supporting these, the more aggregate, 'top-down' approach for assessing uncertainty proposed in the NATA report is appropriate. However, the current implementation requires significant further work before meaningful results and insights can be obtained. In particular, the methods and supporting information are not yet sufficient to allow the assignment of probability distribution functions for representing uncertainty in each of the NATA components (emissions, fate-and-transport, exposure, and dose-response) and the combination of these to estimate a probability distribution for the resulting prediction of risk. Instead, a scenario-based approach should be used to capture and discuss key conceptual and data uncertainties in the NATA. This would allow the focus to be upon the assumptions and data-gaps that might contribute to inaccuracies in the assessment, rather than a focus on imprecision implied by the current probabilistic method and results (with the implication that the central tendency of the estimate has a degree of

reliability that in many cases may not be justified).

For each of the components of NATA, summary tables should first be developed summarizing the amount of available vs. missing data for the assessment. A sequential outcome (or 'event') tree, with different branches to represent the adoption of each of the major conceptual or data-source assumptions could then be constructed. For the emissions component, the alternative scenarios could consider use of information from the different available sources and databases. For the fate-and-transport model predictions of the ratio of ambient and exposure-unit concentrations to emissions, the scenarios can address compounds and conditions where ASPEN is applicable, vs. those where it is not. As noted above, the current implementation of HAPEM is inappropriate for representing variability in the target population exposures, and alternative approaches (when developed) could also form the basis for different scenario evaluations in the assessment. For the dose-response component of the model, reliance on different databases or the use of currently accepted vs. proposed (or 'under review') toxicity factors would allow insight into the impact of these assumptions.

When combined, this scenario tree would provide insight into which combinations of assumptions lead to the most important differences in predicted exposure and risk (and air toxic prioritization), and which assumptions in turn require further discussion with stakeholders and improved resolution through further data collection and model development. This would also help to provide insight as to which sources of uncertainty are specific to the NATA and which are common to all health risk characterization efforts, suggesting specific needs for NATA improvements as well as more general priorities for air toxics research in ORD.

A major output of the NATA may involve lists of counties estimated to be among the top X (or top Y%) of counties in terms of computed exposure and risk for all compounds, or selected HAPs. Should such lists be developed as part of NATA, it will be very important to identify the sensitivity of the results to differences in assumptions, using the scenario tree approach described above. Readers should be able to identify the specific reasons why a county is included in any list, for example, due to high estimated emissions of a particular type (facility, area, mobile on-road or off-road) for particular

sets of compounds; low ambient dilution and dispersion (due either to local meteorology or the presence of small census tracks with high emissions); or specific demographic or time-activity factors. The presentation should also indicate the plausible scenarios under which the county is *not* included in the list.

The use of a detailed ('bottom-up') Monte Carlo simulation for characterizing uncertainty in NATA predictions is not recommended at this time, though such an approach should be used as part of the ongoing studies to explore the sensitivity of the component models to different parameter inputs.

7. Have the results of the assessment been appropriately and clearly presented? Can you suggest alternative methods or formats that could improve the presentation and communication of these results?

The NATA document reflects a proper concern with the importance of effective communication of results, to encourage a holistic understanding of air toxic risks and the options available for addressing them; and to address the various information needs of decision makers and stakeholders in the EPA, other federal and state agencies, industry, environmental and other interest groups, and the general citizenry. A problem facing EPA staff in this task is finding a means to clearly communicate which pieces of the assessment are understood and characterized with a relatively high degree of confidence, and which require further data gathering and model improvement before reliable estimates can be assured. Given the importance of environmental pollution information such as this (e.g., the widespread use of the TRI and the current NTI data by business, environmental groups and citizens), we recommend that the Agency clearly distinguish between those parts of NATA that are well established, vs. those which are in an earlier, developmental stage, based upon less certain science and models, and more limited data. In developing the web page for communicating results, the EPA should consider use of a hierarchical set of pages to differentiate between:

1) Information that is based solely on data or data reports, e.g., emissions data sets and ambient concentration and personal monitoring datasets for different compounds in different

locations;

2) Information that is based on relatively simple or highly confident model calculations, such as ambient air concentration values computed by ASPEN for well-characterized air toxics that are not affected by secondary pollutant formation processes, in areas (terrain and meteorology) where ASPEN can provide reliable prediction, or total exposures to ambient pollutants computed assuming a simple indoor-outdoor penetrating factor; and

3) Information based on new model developments, where research is ongoing to improve the basis for prediction.

These pages could be color coded and titled to indicate: a) existing NATA data (using, for example, a blue background); b) existing NATA models (pale green background); and c) models undergoing research and development (yellow for caution). Graphic representations, such as a thermometer type graph, could be used to display the levels at which different health effects are seen, or to present different cancer risk levels.

The current NATA document was written to some extent for this Panel, with a number of the discussions directed towards an SAB review. A more general report for a broader audience should be written. This revised report should include an executive summary to prime readers to key findings and issues from the beginning. Many of the graphics used for summarizing risks across the multiple compounds and in different locations are very clear and effective (though this does make the responsibility even greater for ensuring that these results are accurate and reliable).

Members of the Panel held differing opinions as to whether model exposure and risk estimates or rankings should be presented for specific counties in the U.S. Such information might include an alphabetical list of the 100 counties with the highest exposures and risks (or the top Y% of counties). As noted in response to the previous charge question, such a listing should include information to help readers discern the particular reasons why (and the set of assumptions under which) the county is

included in the list. Some members of the Panel felt strongly that states, citizens and other stakeholders would greatly benefit from this information and that, in any case, other organizations will be able to access and manipulate the NATA results to produce it. Others felt just as strongly that the uncertainty in NATA estimates is too great to justify identification of specific "hot-spot", high-risk counties, and that even if others could generate such a list, this was preferable to the EPA itself producing it (with the implied "official support" that this would entail). We note this disagreement within the Panel and hope that we have clarified (here and in the main report) the advantages and disadvantages to the Agency of producing a list of counties with high estimated NATA exposures and risks.

8. The exposure methodology in NATA is being considered as one candidate for providing the basis for a national scale benefits analysis (as required in Section 812 of the CAA). Please comment on the strengths and weaknesses of this approach, recognizing the limitations outlined in the NATA report.

The current exposure methodology and results in NATA are not yet ready for use in the national scale benefits analysis required in Section 812 of the Clean Air Act. Once the needed improvements noted above are implemented with a few more iterations of the approach, application to benefits assessment can be considered. In particular, a meaningful benefits assessment must consider the full distribution of exposure and risk (not just median values) and should also address acute health effects. Once exposure predictions are improved and validated, the cost-effectiveness of alternative toxics management strategies (for emissions and exposure reductions) could be compared, stopping short of a full benefits assessment (that would be based on health risks, mortality and morbidity avoided). If a full distribution of exposure is estimated for an information-rich HAP, such as benzene, as part of the current NATA, then the 812 study could attempt an initial benefits assessment for that HAP, to illustrate the type of analysis that is envisioned for the future.

9. Do you have suggestions for research priorities that would improve such air toxics assessments in the future?

An extensive research effort should be mounted to address the wide array of the data and model development needs to significantly improve the scientific foundation of future NATAs, as well as regulations based on the health risks of air toxics. The needs (addressed in detail in the NATA document) include both fundamental and chemical-specific research and span the whole of the risk paradigm (i.e., emissions, ambient concentrations, exposures, effects, and risks). Because air toxics research has been under-funded by the Agency for so long, considerable new resources are needed. Fortunately, the NATA allows identification of the uncertainties that are inhibiting the development of reliable quantitative assessments so that the new resources could be well-focused. We understand that the EPA ORD is completing a strategic plan for air toxics research, so there is no need for SAB to duplicate this effort. We recommend that this strategy be developed in concert with external experts on the related topics and that the subsequent draft be reviewed by this or a similar Panel. The Health Effects Institute is also preparing an Air Toxics Strategy, so ORD might also derive benefit from their activity. Research needs for diesel particles can be obtained from EPA's recent diesel health assessment.

Developing a research strategy takes considerable time and then more time is required for implementation. Using the information developed in research programs is just as important as generating the information. Thus, no air toxics research program can be useful until it is incorporated in Agency models for assessments and until, for example, the new dose-response assessment information is entered into IRIS. These activities also need appropriate resources. The Panel recommends that EPA develop a plan that describes what work (information collection, research, and assessments) it will perform with existing resources over the next few years that will directly improve the 1999 and future NATAs.

2.0 INTRODUCTION

2.1 Background

The air toxics program was authorized under the 1970 Clean Air Act and reauthorized through the 1990 Amendments to the Clean Air Act (CAA). Since 1990, EPA and its regulatory partners, including State, local, and tribal governments, have made considerable progress in reducing emissions of air toxics through regulatory, voluntary, and other programs. To date, the overall air toxics program has focused on reducing emissions of air toxics from major stationary sources through the implementation of technology-based emissions standards. These actions, as well as actions to address mobile and stationary sources under other CAA programs, have achieved substantial reductions in air toxics emissions. The EPA expects, however, that the emission reductions that result from these actions may only be part of what is necessary to protect public health and the environment from air toxics. The Agency's approach to reducing air toxics risks consists of four key components: 1) source-specific and sector-based standards (e.g., risk-based standards, under the Residual Risk Program[1]; area source standards, through the Integrated Urban Air Toxics Strategy)[2]; 2) national, regional, and community-based initiatives; 3) National Air Toxics Assessment (NATA) activities; and 4) education and outreach.

As a primary component of our national air toxics program, NATA activities include all data gathering, analyses, assessments, characterizations, and related research needed to support the other components of the EPA air toxics program. More specifically, NATA activities include: expanding air toxics monitoring; improving and periodically updating emissions inventories; periodically conducting national- and local-scale air quality, multi-media and exposure modeling; characterizing risks associated with air toxics exposures; and continuing research on health and environmental effects of, and

1 The Residual Risk Report to Congress was reviewed by the Residual Risk Subcommittee of the SAB on August 3, 1998. See EPA-SAB-EC-98-013, Review of the US EPA's Report to Congress on Residual Risk (U.S. EPA/SAB. 1998.)

2 The Integrated Urban Air Toxics Strategy is documented in 64 FR 38705

exposures to, both ambient and indoor sources of air toxics. The EPA plans to use these technical support activities to help set program priorities, characterize risks, and track progress toward meeting overall national air toxics program goals, as well as specific risk-based goals such as those of the Integrated Urban Air Toxics Strategy.

As part of the NATA activities, the EPA Office of Air Quality Planning and Standards (OAQPS) has completed an initial national-scale assessment that demonstrates an approach to characterizing air toxics risks nationwide. This initial assessment provides preliminary information for characterizing, on a national scale, potential health risks associated with inhalation exposures to 32 air toxics identified as priority pollutants in the EPA Integrated Urban Air Toxics Strategy. In addition, the assessment examines the inhalation exposure resulting from emissions of diesel particulate matter. The primary stated goals of the initial national-scale assessment are to assist in:

1) Identifying air toxics of greatest potential concern, in terms of contribution to population risk;

2) Characterizing the relative contributions to air toxics concentrations and population exposures from different types of air toxics emission sources;

3) Setting priorities for the collection of additional air toxics data (e.g., emission data, ambient monitoring data, data from personal exposure monitoring) for use in local-scale and multipathway modeling and assessments, and for future research to improve estimates of air toxics concentrations and their potential public health impacts;

4) Establishing a baseline for tracking trends over time in modeled ambient concentrations of air toxics; and

5) Establishing a baseline for measuring progress toward meeting goals for inhalation risk reduction from ambient air toxics.

2.2 Charge

In the months leading up to the SAB NATA Review Panel meeting, the Agency and the Board negotiated a Charge consisting of the nine questions below as follows:

1. *Given the nature of the NTI and the methods by which it was developed and reviewed, have available emissions data been appropriately adapted for use in this assessment? Can you suggest improvements to EPA's application of the NTI for use in future initial national-scale assessments?*

a. Can you suggest improvements to the treatment of compound classes (e.g., chromium and compounds), given the nature of the information available in the inventory?

b. Can you suggest improvements to the methods used to spatially distribute area and mobile source emissions?

c. Can you suggest improvements to the methods used to specify default point source emission characteristics in lieu of missing emissions data?

2. Is the approach taken for the geographic aggregation of ambient and exposure concentrations generated by the ASPEN and HAPEM4 models appropriate in light of the limitations of the models, the available emissions data, and the results of the comparisons of ambient predictions with ambient monitoring data?

3. Has available dose_response information (e.g., different sources of information, a different prioritization scheme) been appropriately used in this assessment? Can you suggest methods that could improve upon the use of available dose_response information?

4. What are the strengths and the weaknesses of the overall conceptual approach to risk characterization used in this assessment? Given the underlying science and the intended purposes of the assessment, can you suggest ways in which the risk characterization could be improved?

a) Is the method used to aggregate cancer risks appropriate? The aggregation of carcinogenic risk within two categories, based on weight-of-evidence classifications, is of particular interest. Is the method used to aggregate non-cancer hazards appropriate? The summation of hazard quotients within target organs, the categorization of sums by ranges of uncertainty factors, and the inclusion of all target organs (as opposed to only the organs associated with the critical effect) are of particular interest.

5. Although EPA has concluded that available data are not sufficient to develop a reliable quantitative estimate of cancer unit risk for diesel emissions, it is clear that this pollutant class may be of significant concern in a number of urban settings. The risk characterization in this report includes a discussion of diesel particulate matter to help states and local areas frame the importance of this pollutant compared to the other air toxics. In the context of this assessment, is the discussion in this report regarding making risk comparisons among other air toxics appropriate? Can you provide any suggestions that would improve upon this approach to comparing the toxic health effects of diesel particulate matter with other pollutants?

6. Given the limitations inherent in this preliminary assessment, have uncertainty and variability been appropriately characterized

a) Can you suggest ways that the characterization of uncertainty and variability could be improved, made more transparent, or integrated more effectively into the risk characterization?

b) Can you suggest methods for quantifying individual as well as composite uncertainties associated with the emissions inventory, dispersion modeling, exposure modeling, dose_response assessment, quantitative risk estimates, and accumulation of risk across air toxics?

7. Have the results of the assessment been appropriately and clearly presented? Can you suggest alternative methods or formats that could improve the presentation and communication of these results?

8. The exposure methodology in NATA is being considered as one candidate for providing

the basis for a national scale benefits analysis (as required in Section 812 of the CAA). Please comment on the strengths and weaknesses of this approach, recognizing the limitations outlined in the NATA report.

9. Do you have suggestions for research priorities that would improve such air toxics assessments in the future?

2.3 SAB Review Process

The SAB Panel was recruited following nominations received from SAB Members and Consultants, the Agency, and outside organizations. The group met in public session on March 20 -21, 2001 at the Radisson Governor's Inn in Research Triangle Park, NC. Written comments were prepared before, during and after the meeting by Panel members and consultants, and made available at the meeting, which formed the basis for this report. A more detailed description of the SAB process for this review can be found in Appendix A.

3. EVALUATION OF THE DRAFT 1996 NATA

3.1 General Findings

The Panel found that the draft NATA 1996 document represents an extensive and comprehensive effort to systematically evaluate and link the various components of the risk paradigm relevant to HAP impacts, including emissions, atmospheric transport, human exposure and risk. In the absence of widespread measurements, the 1996 NATA relies on modeling to estimate some elements of the emissions inventory, as well as ambient concentrations and exposures. While some aspects of the current data collection and modeling are advanced enough for confident prediction, others are still highly uncertain. An expanded set of measurements is needed to evaluate and develop confidence in the models, and to provide independent information about spatial distributions and trends over time.

As part of our review, we have identified specific areas where the current NATA is especially problematic. Some of these difficulties can and should be addressed for the current 1996 assessment. Others suggested improvements will require a more long-term effort, and should be targeted for the 1999 and future NATA's. In the recommendations that follow in our report, short- vs. long-term targets for implementation are identified.

The development of a nationwide assessment of air toxic emissions, atmospheric transport, human exposure and risk is a daunting task, and the Agency has had to make a number of choices cognizant of the limitations in scientific understanding, available data, and the time and resources available for the assessment. A key choice has involved the selection of the spatial scale of aggregation for conducting the NATA, and for reporting the results. The census tract is utilized as a basis for estimating emissions (at times inferred from information at higher levels of aggregation, such as the county level), predicting atmospheric transport, defining receptor populations, and computing their exposures and risks. The results are then aggregated back up to the county level for reporting

purposes. While we agree with this basic strategy for assessment and reporting, there are a number of difficulties that arise in its implementation.

The census tract is a good unit for defining the demographic characteristics of receptor populations, but it is not a good geographic unit for air pollution modeling and assessment. In particular, densely populated census tracts are small, while those in sparsely populated areas tend to be large. This tends to misrepresent the allocation of emissions and bias the calculation of representative ambient and exposure calculations for densely vs. sparsely populated areas. This problem needs to be identified in the current NATA, and addressed in future NATAs through conversion to a regular spatial grid for emissions tracking and the calculation of ambient concentrations, with subsequent conversion back to underlying census tracts for population exposure and risk calculations.

A major finding of the Panel is that parts of the NATA are based on relatively reliable data and/or well-established scientific estimation and modeling methods, while other aspects are based on more limited data and methods that are in an earlier, developmental stage. This applies to all aspects of the NATA, including emissions estimates, estimates of ambient concentrations based on the ASPEN model, estimates of exposure based on the HAPEM modeling system (or, as suggested in our report, other, simpler methods that should be considered in parallel with the HAPEM predictions), and risk estimates requiring the use of toxicity factors based on different amounts of scientific information and consensus. To help citizens and other users of NATA better understand the differing bases for NATA results, we recommend use of a hierarchical presentation of results that distinguishes between quantities measured or modeled at different levels of scientific development, and with differing levels of available data and confidence.

While we have attempted to provide specific information and recommendations to improve the 1996 and future NATA studies, we recognize that much of the need for improved information applies generally to the field of air toxics and health toxicology, and is not specific to the NATA. When uncertainties and concerns are apparent in the NATA methodology, we have attempted to distinguish between those endemic to NATA and those more broadly applicable across the field of environmental

health risk assessment. We also note that we have focused on the general methodology presented in the NATA document, and not the specific values of inputs and parameters used to implement it (though specific examples are identified to be illustrative of apparent problems and areas of concern). The absence of comment on specific emission, atmospheric transport, exposure and toxicity factors should not be construed to indicate Panel review and approval of these values. Separate peer review is required for the specific parameter values and factors used to implement the NATA.

3.2 Responses to Specific Charge Questions

3.2.1 Charge Question 1

Given the nature of the NTI and the methods by which it was developed and reviewed, have available emissions data been appropriately adapted for use in this assessment? Can you suggest improvements to EPA's application of the NTI for use in future initial national_scale assessments?

Given the enormity of this task, the Agency has made a valiant effort to compile a model-ready national air toxics inventory for the point, area, on-road and non-road source sectors for 1996. The NATA document appropriately acknowledges the limitations in the information and implications of this for the development of the 1996 NTI. The Emissions Modeling System for Hazardous Air Pollutants (EMS-HAP) which was developed to process the emissions inventory data for subsequent air quality modeling (see Appendix C of the NATA report) is impressive. However, there are a number of steps that should be taken to further improve the accuracy of the results of the assessment and reduce the uncertainties. Our comments address improvements that could be considered in future applications and iterations of the NTI and the National-Scale Air Toxics Assessment (NATA). They specifically address improvements for the collection of raw HAP emission inventories and the application of EMS-HAP for the various source sectors (i.e., point, non-point, on-road and non-road sources).

3.2.1.1 National Toxics Inventory (NTI)

Improvements in the development of the 1996 National Toxics Inventory (NTI) are evident when compared to the inventory that was prepared for the Cumulative Exposure Project (CEP). There are significant differences in the national emissions totals between the two studies presented in Table 4-4 of the NATA report. We believe that much of this difference is a result of improved data, progress made by the Agency in resolving the emissions inventory discrepancies, and the development of more advanced emission inventory methodologies. The emission inventory developed for the CEP relied heavily on VOC and PM emission estimates from an interim 1990 National Emissions Trends (NET) Inventory. The criteria pollutant emissions were converted to individual HAP emissions via speciation profiles which are now considered dated and are no longer used by the Agency to estimate HAP emissions. We are supportive of the iterative approach taken by the Agency to improve the emissions inventory and continue to view the development of future national air toxics inventories as a work in progress. The inclusion of emission and facility specific information collected by State and Local Air Pollution Control Programs for point sources represents a significant advancement in this effort.

The Table 4-5 Facility Count Summary by state provides the reader with some insight about the extent of the state point and area source inventories that were available to the Agency in developing the 1996 NTI. We understand that there could be some overlap between the NTI and the NET, so the word “unique” should be removed from the Table since it may suggest to the reader that the two inventories are mutually exclusive of one another. We agree that the NET provides a good resource for checking NTI’s completeness. A quick examination of the NTI/NET ratio indicates a range of 0.07 to 4.6. We are concerned that facilities may be missing from the 1996 NTI in states where this ratio is well below one. This would result in an underestimation of emissions for these states, directly impacting predicted ambient ASPEN concentrations and subsequent risk predictions.

In the next round of data collection for the 1999 NTI, the Agency should consider implementing some quality assurance/quality control measures to ensure that a satisfactory level of completeness and accuracy is achieved. This would include a careful review of the NET facility files for the states with extremely low ratios to determine how many HAP point and area sources are missing. Once these facilities are identified, an effort could be undertaken with the affected state or industry to review the

necessary raw HAP emissions information. The current emission inventory format developed by the Agency in the AIRS database, which lists the HAP emissions associated with each facility, provides an excellent way to efficiently review and verify the large amounts of emissions information. The identification of all missing point sources in the NTI will be a difficult task. The best future solution will be the development of a consistent national HAP emissions inventory data collection and reporting rule, with proper incentives for industry to participate and comply. This would help to eliminate the potential bias of missing facility emissions and the resulting underestimation of exposure and risk that currently exist for point sources in the 1996 NATA.

Recommendation for NATA 1999: Implement additional QA/QC measures to ensure that a satisfactory level of nationwide completeness and accuracy is achieved for the point and area source emission inventories.

In future NTI assessments of on-road emissions, the Agency should make an effort to incorporate State and Local Air Pollution Control Program data for on-road emissions. Some States have county specific (vehicle miles traveled) VMT and VOC data sets that are prepared as part of their State Implementation Plans (SIPs). The NTI uses HAP vehicular emission factors generated by MobTox5b and then multiplies them by county VMT estimates that are based on a population surrogate. An analysis comparing the VMT estimates for the New York Metropolitan Area prepared by the EPA and New York State indicated large differences in emission estimates (NESCAUM, 1999). The state VMT estimate in the NESCAUM report is based on actual vehicle count data from the Department of Transportation. The EPA VMT estimate is based on a population surrogate. The resulting county differences in VMT clearly indicate that the EPA method will result in underestimation of on-road emissions in more suburban counties, while largely overestimating on-road emissions in urban counties.

In addition, the NYS Department of Environmental Conservation (NYSDEC) attempted to verify the HAP emission factors generated by the MobTox 5b model (NESCAUM, 1999). To address this problem the MobTox input files were placed into the Mobile Model which generates

emission factors for total organic gases (TOG), but not air HAPs. These TOG factors were then compared to the VOC emission factors generated in the SIP demonstration for the New York Metropolitan Area (9 counties). The results of this analysis indicated that EPA's MobTox inputs tended to underestimate TOG emissions, which suggest that of HAP emissions are similarly underestimated. In future NATA assessments, on-road models that incorporate state or urban specific input variables (e.g., vehicle speeds, vehicle fleet type and age, etc.) should be developed to estimate on-road HAP emissions.

Recommendation for NATA 1999: Continue the development of the on-road model to accept input parameters developed by State and Local Air Pollution Control Agencies for the development of the 1999 on-road emission inventory.

The determination of the non-road emission inventory appears to be one of the weakest links in the NATA document. The NATA document does note the limitations associated with the development of the nonroad emissions inventory and acknowledges the recent 202(l)(2) rulemaking which outlines a research strategy to improve the non-road emissions inventory for future NATA studies. We reviewed Appendix C and the paper on the Geographic Allocation of State Level Non-Road Engine Population Data to the County Level (9/16/98) to take a more in-depth look at the factors used in NATA 1996 for determining and allocating non-road emissions. The document indicates that non-road construction equipment emissions were estimated by assuming there was a proportional relationship between the dollar value of construction and the amount of construction in a given area. This is not a good surrogate to use when estimating these emissions for urban counties in the northeast and perhaps some other areas across the country where housing and commercial building prices are extremely high. For example, the relative contributions of non-road diesel PM contributions are unrealistically high for the NYC Metropolitan counties. While the dollar value of construction is high in these counties, less of this construction is at new sites where non-road diesel is used extensively for earth moving. Rather, construction occurs more at existing sites, where the ground is already level (and, for example, much of the work is done by in-place cranes). A similar over-estimation of non-road diesel emissions is likely to occur in other urban areas that are already highly developed, given that these emissions are based

primarily on the dollar value of construction.

The relationship between the cost of construction expenditures and non-road diesel emissions varies across the country and the potential impact of the use of this emissions surrogate needs to be evaluated in future NATA assessments. This factor may also be impacting emission estimates for other HAPs (besides diesel) associated with nonroad construction (e.g., formaldehyde, benzene, acrolein, and acetaldehyde) in these urban areas.

Recommendation for NATA 1999: Critically reevaluate surrogates used to estimate the non-road emissions inventory and make adjustments were necessary. Continue the development and verification of the non-road emission inventory and non-road model for future iterations of NATA by expanding the research agenda to fill known important data gaps. These data gaps should be prioritized to reduce the most significant uncertainties associated with the non-road emission inventory and model predictions.

3.2.1.2 Reactivity Class Decay Rates

We have some questions about the reactivity classes and constants used in the EMS-HAP. The reactivity categories class values and decay rates should be identified for each HAP modeled in NATA. We are specifically concerned about how EMS-HAP handles emissions of 1,3-butadiene, a chemical that undergoes rapid decay in the daylight (estimated half-life = 1.6 hours), but slower decay at night (estimated half-life = 9 hours) (CARB, 1992; Harley and Cass, 1994). Our discussions with the Agency indicated that the reactivity constants for 1,3- butadiene and other photochemically reactive organic HAPs are assumed to remain constant for the entire day. We believe that EMS-HAP processing should account for diurnal and seasonal variations in decay rates, or order to be scientifically defendable. Furthermore, critical assumptions and uncertainties associated with the assignments of reactivity classifications for HAPs, and decay rates for various stability categories for modeling should be discussed in more detail.

Recommendation for 1996 NATA: Reactivity categories and decay rates should be identified for each HAP modeled in ASPEN. Critical assumptions and uncertainties associated with the assignment of reactivity classifications for HAPs should be discussed.

Recommendation for 1999 NATA: Update reactivity categories assignments and decay rates by incorporating HAP specific information when available. For HAPs identified as important risk drivers or regional contributors evaluate the impact of the assumption that each pollutant within a specific reactivity class is assumed to decay at the same rate.

3.2.1.3 Temporal Allocations

The use of the eight 3-hour blocks to calculate annual ambient concentrations for each time block in each census tract is a strong feature for anticipated downstream uses. It allows HAPEM to account for daily variations in HAP exposure by using the activity patterns for the point, area, onroad and off-road source sectors as presented in Appendix D of the EMS-HAP Users Guide. The emissions Equation 5-1 in Appendix C provides an excellent example of how emissions are divided to provide a grams/second emission rate for each three-hour period during day. For example, emissions rates for mobile source HAPs are higher during the 3-hour blocks which contain rush hours. Therefore, the potential HAP exposure while driving or walking during these time periods would be higher and can be accounted for by activity patterns contained in HAPEM. Figure 3-3 provides an excellent example of the daily fluctuations of a HAP concentration overlying the daily activity scenario of a cohort. This appears to be a very good approach for capturing daily variability in ambient exposure concentrations in relation to activity patterns.

It would be interesting to see the range of predicted daily values for some of the HAPs identified as risk drivers in future assessments. While the approach for diurnal dissaggregation of emissions is appropriate, we do note in the following section that, in its coupling with HAPEM, ignoring seasonal variation and using a sequence of independently sampled person-days to represent annual exposure does lead to a misrepresentation of long-term individual to individual variations in exposure,

and that the result may only be appropriate for estimating the median (rather than the full distribution of) exposures in a census block or county.

3.2.1.4 Quality Analysis and Quality Control (QA/QC) and the Reduction of Uncertainties

Under Section 3.5.2.6 of the Agency's Guidelines for Exposure Assessment, it is stated: "Any data developed through previous studies should be validated with respect to both quality and exptrapolation to current use. One should consider how long ago the data were collected and whether they are still representative." Although the Agency stated in the report that it went through three rounds of review with state and local agencies, this review process was apparently not stringent enough to be considered as a QA/QC evaluation. This is pointed out in the NATA document, when it states that, "EPA has not undertaken a full QA/QC evaluation of the NTI," (page 56) and "EPA did not attempt to verify the methods by which emissions were estimated or undertake a full quality control evaluation of the NTI" (page 104). The results of any assessment conducted by using models can only be as good as the quality of the input data used for the analysis. The importance of QA/QC process is obvious and the needs for further reduction of the uncertainties stated in subsequent discussion in this review report should also be clear.

Recommendation for NATA 1999: Implement additional QA/QC measures to ensure that a satisfactory level of completeness and accuracy is reached for all emission inventories.

a) Can you suggest improvements to the treatment of compound classes (e.g., chromium and compounds), given the nature of the information available in the inventory?

While in some instances ignoring speciation effects for an element or grouping compounds with similar behavior can lead to beneficial simplifications for analysis, this must be done with great care. The grouping of chromium compounds to improve modeling efficiency creates downstream problems

for the proper risk characterization of these compounds and introduces more uncertainty than necessary. The issue of how much hexavalent chromium (Cr^{6+}) is present in total chromium stack and ambient measurements has been investigated by numerous researchers over the past decade (Bell and Hipfner, 1997; Grohse, et al, 1998; Scott et al., 1997). The use of the assumption that 34% of the total ambient chromium is present in the carcinogenic hexavalent form clearly results in regional over and underestimations of risk. Chromium compounds should not be grouped and should be segregated based on valence state using the SIC codes when the inventory is developed. For example, census tracts which contain chromium electroplaters or chromate production facilities will have a much higher proportion of ambient Cr^{6+} than census tracts impacted by municipal waste combustion facilities. The Agency should apportion Cr^{6+} for each source category in the EMS-HAP stage and have two separate inputs into the model as chromium and Cr^{6+} using the available literature on this subject.

The use of the assumption that 65% of the predicted total ambient nickel is insoluble and in the crystalline form is a conservative assumption for assessing cancer risk. It is more conservative than the 50% assumption used in the Utility Study (EPA, 1998a). The Agency should investigate if the available literature on this issue would support an approach as described above for Cr^{6+}.

Given the available emissions information for polycyclic organic matter (POM), the grouping of POM species into two groups is appropriate. The inclusion of the toxicity equivalency factors (TEF) approach for dioxin compounds in EMS-HAP is also appropriate.

Recommendation for 1999 NATA: The Agency should apportion Cr^{6+} for each source category in the EMS-HAP stage and have two separate inputs into the model as chromium and Cr^{6+} using the available literature on this subject. In addition, a reactivity decay rate will have to be developed and incorporated into EMS-HAP for Cr^{6+}.

b) Can you suggest improvements to the methods used to spatially distribute area and mobile source emissions?

The Agency recognizes the uncertainty associated with estimates for area and mobile emissions sources that are compiled on a county-wide basis, and then allocated using spatial allocation factors (SAFs) to census tracts within the county. While it is difficult with current information to estimate emissions from these sources and to allocate the emissions in a more refined manner than is currently done in the NATA, suggestions are provide for future NATAs.

EMS-HAP handles point source location defaulting within census tracts by eliminating census tracts with a radius less than or equal to 0.5 km, because the ASPEN model would calculate excessively high concentrations for these small areas. A default consolidation mechanism should also be developed for area, on-road and non-road emission census tract spatial allocations in these small census tracts. A possible spatial allocation method for future iterations of NATA is discussed below .

The initial screening assessment may result in the generation of false positive results since the presentation of annual average concentrations by county and state are greatly influenced by area (e.g. square miles) and population densities. Therefore, in future iterations of NATA the Agency should consider an alternative approach before there is any attempt to characterize potential public health risk due to the inhalation of air toxics. This step would involve the isolation of counties with major metropolitan areas and the mapping of all averages in these locations using a uniform grid approach. To calculate the average concentration of pollutant X for each grid, the individual concentrations could be weighted by the area of the assigned census tracts. The average values for each grid can then be reported for each county. This would remove the dilution effect of using large areas and would limit the influence of small census tracts, since the size of a census tract is based on population density, not source activity. Source activity should determine the magnitude of predicted concentrations. This type of analysis would provide results that are directly comparable from one metropolitan area of the country to another.

Our concern is illustrated by the following brief discussion. Those counties in highly populated areas are predicted to have higher average concentrations while those in the lower population areashave lower predicted concentrations. While this is in part due to the presence of some air toxics

sources (particularly area and mobile sources) that do properly correlate (to some extent) with population, it also occurs because census tracts are not uniform in size: some may be as small as 0.03 km^2 while others are as big as 3084.2 km^2. Thus for the smaller census tracts, concentrations are calculated much closer to the source and therefore tend to be much higher on average. However, for larger census tracts, the average tract impacts may be more representative of average concentrations as much as 31 km away from the source. Both of these concentrations may be fairly representative of the average concentration for each tract however the larger tract covers a much larger area. The results indicate that the distributions in the larger tracts represent the averages of the averages. Therefore, when you look at predominantly rural States you observe very narrow bands of concentrations. In contrast, there is a wider distribution of concentrations in more highly populated States. Many of these smaller distribution bands may be valid, while others may not. As a result, small urban areas which may be of public health concern could be missed or overlooked. The approach taken to properly identify and characterize locations with high air toxics exposure will be critical in developing future risk management strategies.

Recommendation for future iterations of NATA: Consider an alternative modeling approach for counties with major metropolitan areas and small census tracts which would involve the mapping of all averages using an uniform grid approach. This type of analysis would provide results which are directly comparable from one metropolitan area of the country to another.

c) Can you suggest improvements to the methods used to specify default point source emission characteristics in lieu of missing emissions data?

The point source defaults used in the NATA for location and stack parameters are conservative approaches and appropriate. While it is reasonable to enter some default stack data for modeling purposes, it is not reasonable to use these values to create default emission data for facilities where all aspects of the needed data are missing. EPA must work with the facilities themselves and State and

Local government agencies to gather realistic information. In most cases, it is better to enter no information at all than to create surrogate data.

Some suggestions for removing stack parameter defaults for facilities that have not provided actual stack information would be to request information from the states for stack testing information which should be available for NET sources in many states, and ask the states or industry if they could summarize any stack parameter information contained in the test reports. This would entail a large effort, but it would help to avoid the use of default parameters and refine the results and contributions to exposure and risk from the point source inventory.

Recommendation for 1999 NATA: To avoid the use of default stack parameters, request that State and Local Air Pollution Agencies or industry summarize any stack parameter information contained in stack test reports if available for facilities that have been assigned default stack parameters.

3.2.2 Charge Question 2

Is the approach taken for the geographic aggregation of ambient and exposure concentrations generated by the ASPEN and HAPEM4 models appropriate in light of the limitations of the models, the available emissions data, and the results of the comparisons of ambient predictions with ambient monitoring data?

3.2.2.1 General Comments

The NATA efforts at modeling HAP airborne fate, transport and exposure represent a serious, diligent effort; and the USEPA NATA team should be commended for this work. A substantial effort has been made explaining and explicitly documenting caveats and limitations of the individual components and steps of the NATA approach. The choice of the census tract as a statistical receptor/exposure unit is a good starting compromise that allows for future coupling with

multimedia/multipathway assessments. The choice of county-level aggregation for the presentation of results is generally appropriate (for most of the air toxics considered) as long as limitations and caveats are clearly identified.

The local (rather than national-scale or even long-range) character of ASPEN calculations offers the practical advantage that it allows for independent local evaluation and refinement of estimates by State and local agencies. Since ASPEN incorporates well-established practices and techniques that local agency personnel should be quite familiar with, it should be expected that such local evaluations would be straightforward and productive. Clearly, the NATA effort represents work in progress; it should be expected that refinements and changes in the NATA approach will take place in both the present and future phases. In particular, HAPEM4 is an essentially new (for the field of air toxics), and potentially valuable element that has been added in this phase. This is the new component, that, from a methodological point of view, takes us from the ambient concentration-based approach of the CEP, to an actual population exposure assessment process. It is important, however, in order for a local application, evaluation, and refinement process to be successful – in fact, in order for such a process to start in the first place – that sufficient guidance and support be provided by USEPA to the State and local agencies regarding the use of new tools, such as HAPEM4. The Agency should provide the necessary resources so that, at a minimum, detailed and thoroughly tested user guides, that fully explain the methods and rationale behind the HAPEM4 approach, combined with demonstration case studies, are developed and provided to the State and local agencies.

As with every new effort, there are problems with data gaps, etc., nevertheless, the incorporation of HAPEM4 into the NATA process is a step in the right direction. It is important that the NATA team distinguish between successes and failures, and identify causes for both. In fact, it is important to ask not only why a model fails in a model-observations comparison, but, also, if a model performs well, if it does so for the right reasons.

3.2.2.2 Specific Concerns and Recommendations

The following is a list of major concerns and areas for possible improvement regarding the specific application of ASPEN in NATA. It should be noted that ASPEN relies on a standard Gaussian plume model formulation (specifically the Industrial Source Complex [ISC] model) and therefore has the well-known inherent limitations of Gaussian models, such as the inability to handle nonlinear chemical transformations or dispersion of contaminants in complex atmospheric flow fields (e.g. sea and lake breezes, etc.). In fact, some of the concerns discussed below arise precisely from the attempt to apply ASPEN, in the NATA approach, to situations that are beyond the range of applicability of its underlying classical Gaussian plume model formulation.

1. The study adopts a "one size fits all" approach to HAP fate and transport modeling, using a single model formulation for different air toxics and regions of the US. The Agency should identify where the model is applicable and works well, and where it does not, and correct and refine the modeling approach for these applications.

2. There is limited quality assurance of available input data (especially emission inventories). The Agency should adopt the use of visual GIS-based tools for inventory development/testing and for emissions preprocessing.

3. There is no consideration of regional/seasonal variability of background (in fact, no clear definition of what is meant by background is given). The NATA report should define what background is; perform refined statistical analysis to identify trends and clustering in background concentrations; and consider in the future simplified seasonal grid-based modeling for the prediction of background.

4. The ASPEN model assessment provides no consideration of long-range transport (LRT). The study should identify specific toxics with LRT concerns and perform grid based modeling (as e.g. in the CMAQ [Community Multiscale Air Quality] Hg modeling project).

5. There is no consideration of seasonal patterns in the local ASPEN calculations (in addition to

diurnal variation). In reality, both meteorology and emissions (as well as chemical transformations) can exhibit strong seasonal patterns and dependencies. For example, there is often a significant temperature dependence for fugitive emissions that occur via volatilization, so that emission rates will differ as a function of season. Similarly, certain activities that generate HAP emissions (such as lawn mower use in northern states) have a strong seasonal component; distributing these emissions uniformly over the year is inappropriate. Seasonal emissions preprocessing and seasonal evaluations of NATA should be used in the next iteration of NATA (i.e., for the 1999 assessment).

6. The ASPEN model is restricted to an overly simplified, inappropriate treatment of secondary air toxics (such as formaldehyde, acetaldehyde, and acrolein) that exhibit nonlinear chemistry. This problem is emphasized by the inconsistency of the ASPEN estimates with the OZIP (OZone Isopleth Plotting program) predictions for the percentage formed versus emitted, and the known dependencies of photochemical transformations on the variability of ambient conditions. The NATA 1996 study should specifically state the uncertainties and limitations associated with the treatment of secondary species involved in complex (nonlinear) photochemistry (as discussed further in the following), and the Agency should plan for development of a more appropriate approach for the next phase.

7. The NATA study provides for no consideration of regional limitations in the ASPEN model applicability, and the corresponding increase in model structure-related uncertainty in areas with complex terrain, sea/lake breeze effects, or other conditions not addressed by the ASPEN model. The NATA report should incorporate regional limitations in uncertainty characterizations by defining topographical/climatological regimes associated with ASPEN applicability (i.e. regimes with different structural uncertainty ranges).

8. There is no consideration of how representative (in addition to complete) the meteorological data are, in particular, with regard to where stations are located relative to emissions and exposed populations. Maps should be provided indicating the locations of the meteorological

stations versus the above topographical/climatological regimes and the distribution of census tract centroids.

9. The Agency has conducted very little diagnostic evaluation of ASPEN. The limited available HAP monitoring data from across the US should be used in an informal, case-by-case, diagnostic analysis, to answer questions such as: Does the model perform better in cases where parametric/input uncertainties are lower? Does the model perform better where model structural uncertainty is lower (i.e. where confidence and applicability are expected to be higher)?

10. The report utilizes inconsistent or ad hoc terminology for terms such as 'national-scale' (rather than "national level" or "nationwide"), 'background', 'cumulative/aggregate', 'grid model' (a term used the for OZIP – which is based on a single box formulation), and 'exposure-related' (rather than demographics-related). There should be an attempt to streamline the terminology and semantics conventions used in the report.

To address these and other uncertainties, we recommend that for the 1996 NATA, the air toxics considered be classified in terms of where ASPEN is expected to provide reasonable results. We recommend three categories: confident; in need of improvement/refinement; and uncertain. Secondary compounds, such as formaldehyde, that are formed in the atmosphere through nonlinear chemical reactions, should be placed in the uncertain category, as should compounds for which background concentrations were found to dominate. The secondary formation of formaldehyde, acetaldehyde, and acrolein in the ASPEN model is calculated by estimating the amount of known precursors which would react in the atmosphere (based on annual average decay rates) and then estimating the amount of product which that amount would form (based on average stoichiometric coefficients). More specifically, ASPEN tries to account for secondary species by adding a surrogate "precursor" species that can then be transported like any other species in the dispersion model. Emissions of the precursor species are calculated as a weighted sum of the emissions of some of the species whose reactions lead to formation of the compound. For formaldehyde, for example, emissions

of 23 compounds are included in the precursor sum.

The first problem with this approach is that formaldehyde is a product of many more than 23 primary organic compounds. It is also a product of many secondary compounds (e.g., higher aldehydes and ketones) that a weighted emissions scheme cannot capture. Another problem is that the extent of reaction of the primary species (and hence the amount of secondary species production) depends on relative humidity, sunlight intensity, temperature and the amount of other organic compounds and nitrogen oxides present in the atmosphere. The reaction systems are very nonlinear, in that formaldehyde and acetaldehyde themselves react to produce radicals that speed the production of secondary species from other organic compounds. These new secondary species include more formaldehyde and acetaldehyde. Lacking a detailed treatment of the coupled chemical reactions of many compounds, the ASPEN model cannot properly account for these nonlinear interactions.

All the (known or suspected) reasons for assigning an air toxic to one of the three categories (confident, applicable but in need of improvement, and uncertain) should be listed and clearly explained in the report. For example, it has been pointed out that potential causes for ASPEN underpredicting monitoring values for metals involve both (a) inadequacies of emission inventories; and (b) the fact that the metal monitors are generally located next to sources (i.e., in a "hotspot"), and the ASPEN modeling approach is not finely resolved enough to capture these hotspots.

The report should also classify geographic regions in terms of ASPEN's expected performance. Areas with complex terrain or meteorology should be distinguished from areas where Gaussian-type models are most applicable. Furthermore, in future assessments, the air quality modeling should be improved by capturing seasonal variations in emissions and fate and transport for all of the toxics. Priority should also be given to the adaptation and application of developing models such as CMAQ (Community Multiscale Air Quality model, a component of USEPA's Models-3 system) that are capable of treating secondary compounds and long-range transport of toxic air pollutants.

In contrast to the well known methods (and of their limitations) incorporated in ASPEN,

HAPEM4 represents application of relatively new, and therefore not as well-developed or –tested, methods for assessing personal exposure to air toxics. Most applications of exposure assessment of this type have been limited to the criteria pollutants (CO, O3, PM). The limitations of this first use of HAPEM4 for NATA have been presented in considerable detail in the NATA document and indeed, they are not trivial. Of particular concern are:

a) The use of single, best value estimates rather than statistical distributions for microenvironmental parameters;

b) No consideration of geographic or seasonal variability in microenvironmental parameters; and

c) Indoor sources are not considered in this phase. While not a scientific/technical limitation per se, this could present some problems when comparing predicted exposures to monitored personal exposures, and in communicating the relevant results in an effective manner.

Another serious issue is the artificially low variability in exposure calculated by HAPEM4 within each census tract. It is understood that this occurs since, (a) the current variability predicted by the model reflects only demographic variability, since ASPEN does not consider air quality gradients within a tract; and (b) the demographic variability is not adequately represented because the current treatment fails to incorporate day-to-day correlations in activity patterns for individuals. Due to these limitations, the 1996 NATA should be restricted to reporting median estimates from HAPEM, not distributions. Figures such as 4-16 should not be included because the percentiles of the distributions shown only represent a small component of the overall variability in exposure. Nevertheless, the benefits of including the HAPEM4 calculations in the overall NATA process are still significant. In particular, the incorporation of HAPEM4 sets a framework in place for the future – allowing iterative improvements in exposure assessments and allowing correction for the fact that the tract population is not concentrated at the tract centroid (even if this is the only concentration calculated for the tract). The commuting feature allows cohorts to move from tract to tract: this can be very important in urban areas with large concentration gradients from tract to tract.

To illustrate these benefits and demonstrate the significance of indoor sources, we recommend that the agency consider including a full-fledged HAPEM calculation for benzene. This example should

account for exposure to indoor as well as outdoor sources and correctly treat day-to-day correlations in activity patterns for individuals. The output from this particular example could be useful for the toxics portion of the 812 benefit/cost analysis. This example also should be helpful in guiding future efforts to characterize exposure for the full set of air toxics. Furthermore, there should be a coordinated effort for future iterations of NATA to utilize and test the new tools and methods currently under development at USEPA (such as the neighborhood scale version of Models-3, the various outcomes of the Human Exposure and Dose Simulation program, etc.) in addition to any refinements that are expected to be incorporated in the approaches currently used (ASPEN and HAPEM). Future efforts should also focus on the incorporation of other important pathways of exposure for multi-media pollutants, such as the fish ingestion route for methyl mercury, drinking water ingestion for arsenic, and soil ingestion for lead.

3.2.2.3 Summary Recommendations for Charge Question 2

For the 1996 NATA:

1) The NATA document should be modified as per the specific recommendations of the previous section, i.e. to:

a) Explicitly identify the level of confidence/uncertainty associated with ASPEN predictions for the specific contaminants considered (using the three group classification recommended in this review), for particular geographical regions and locales;

b) Explain and discuss the fact that only a single component (county to county differences in the median) of exposure variability is characterized in the current application; and

c) Discuss explicitly the limitations of the 1996 NATA approach (i.e. those associated with the treatment of long range transport and characterization of background, nonlinear chemistry of secondary air toxic formation, seasonal variability in emissions and climatology, etc.)

2. A "full-fledged HAPEM" calculation for benzene should be performed and included in the 1996 NATA report as a prototype example for future applications to other toxics: this application should account for exposure to indoor as well as outdoor sources and correctly treat day-to-day correlations in activity patterns for individuals in order to properly address exposure variability.

For future NATA applications:

1) Future NATA applications should address the limitations identified in this review and, for example, consider the effects of factors such as seasonal variability in emissions and climatology, improve the treatment of variability in outdoor air quality within a census track, consider the contribution of indoor sources of air toxics to total exposure, and account properly for inter- and intra-individual variability of exposure.

2) Future NATA applications should test, adapt, and employ (a) more comprehensive, multiscale, air quality models, such as Models-3, that can account for both local and long range transport and for nonlinear chemical transformations, as well as (b) evolving modeling tools for exposure analysis that are currently under development by USEPA and other organizations, and

3) Future applications should also focus on the development and application of a consistent, integrated, framework that incorporates multiple routes and pathways of exposure for multi-media pollutants.

3.2.3 Charge Question 3

Has available dose-response information (e.g., different sources of information, a different prioritization scheme) been appropriately used in this assessment? Can you suggest

methods that could improve upon the use of available dose-response information?

The NATA report does a generally good job of evaluating and using available dose response information for the assessment. The approach used to determine the dose-response based on the level of confidence in the quantitative information from secondary data sources parallels that used by state and federal health agencies when setting guidelines and standards for air toxics. The preferences implemented in the current assessment proceed from using IRIS values of RfCs and UREs to the use of ATSDR MRLs (noncancer), and finally to the use of California EPA RELs and UREs. This order of preferences is reasonable and recognizes that the RfCs, MRLs, and RELs are measures of similar, but not exactly the same human health endpoints. Of the 30 UREs reported in Appendix G, 21 derive from IRIS, four are from Cal EPA data, and one derives from EPA NCEA.

The quantification of the cancer risk is stronger than that of the non_cancer end points. With three or four exceptions there is little actual "new" research that indicates that the UREs or RFDs need to be revised. Still, it is the practice of state health assessors to review the most current data even when using federal or other secondary databases such as IRIS to assess the impact of new information. It is possible that constraints in the federal risk assessment process limit the ability of federal scientists for a similar timely implementation of such a validation process. Federal reviews of IRIS can take several months or years which would preclude a timely "validation review".

Recommendation: For the 1996 NATA verify the accuracy of the IRIS and other tables. Indicate the reference sources for the IRIS data in the document. For chemicals that do not use the NATA protocol such as butadiene and nickel show the rationale for the RFC in detail. The NATA should expand the information of the dose response tables to show: source of values, date of assessment, peer-review, EPA review status and qualitative statement relative to more current studies. For the 1999 NATA EPA is encouraged to update all IRIS RFC values.

3.2.3.1 Ongoing Revisions in Toxicity Factors

The toxicity values reported in Appendix G and used in the NATA study were not examined in detail by the Panel to ascertain whether they are the most recently reported values. The EPA is re-examining the carcinogenic potency of 19 of the assessed HAPs. Presuming that these re-evaluations are ongoing, how will the NATA assessment process incorporate new or revised estimates of cancer and noncancer dose-response information in its periodic reappraisal of risks posed by toxic HAPs? Will any revisions to UREs as a result of this activity be incorporated into a revised 1996 air quality assessment or future assessments? The dose-response information summarized in Tables 3_5 and 3_6 should include some characterization of how recent are the IRIS (and other sources of) estimates of cancer and non-cancer data. In addition, if UREs or RfCs are undergoing re-evaluation, this should be indicated in the same tables. A related question is whether dioxins will be added back into future assessments?

Recommendation: For the 1999 NATA include dioxins. Also consider setting a calendar date that will be used for selection of reference information from secondary sources and a schedule for periodic update for NATA risk estimates.

3.2.3.2 Degree of Conservatism in Health

The NATA report uses UREs (unit risk estimates) developed by the USEPA and the California EPA to determine plausible upper bound estimates according to the priority system present in Appendix G of the NATA report. In so doing, it is clear that these estimates are designed to provide a degree of conservatism in health estimates. In places in the report it is noted that actual HAP risks "are likely to be lower, but may be greater (than those reported in the document)." While true, the conservative nature of health factor estimates are widely recognized, so that repeated use of such statements is not necessary.

For some chemicals in the NATA, toxicity factors based on MLEs are available and utilized, while for others, upper bound estimates based on upper confidence limits (UCLs) are used. Since UCLs, generally used when fewer data are available, are more conservative than MLEs, it is not clear whether these choices affect the relative likelihood of different compounds being included among the list of risk

driving HAPs. Furthermore, as noted in response to Charge Question 4, summing cancer risks based on UCL's can lead to an even greater (though unspecified) level of conservatism in the estimate of the aggregate risk from multiple compounds. Importantly, acrolein was determined to be one of the HAPs of great concern according to the results of the risk characterization. Yet, the reported physical effect used in derivation of the RfC was mucous membrane irritation (Appendix G). This RfC also included an uncertainty factor of 1000. What is the utility (and how much confidence is there) in the resulting risk estimate?

Recommendation: Indicate in the document the differences in relative risk expected if MLEs were to be used instead of upper bound estimates of cancer potency. Provide comment on the effect of different safety factors on the selection of acrolein as a risk driver.

3.2.3.3 Validating Risk Predictions

For the CEP analysis, the uncertainties in the dose response data were marginal compared to the difference in the relative exposure estimates, based both on the ASPEN estimates and the state monitoring data which validated the ASPEN estimates. Thus, in most cases the dose-response data available are thought to be adequate as used in the NATA process. Nevertheless, it may be desirable to "ground truth" the risk estimates through comparison with Health Based Guidelines and standards determined by Public Health scientists in the states to support state air toxics regulations.

Recommendations: For 1999, request that States provide reference concentrations as part of inventory or state review of NATA. The State estimates could be provided in an appendix table for comparison purposes.

3.2.3.4 Use of Oral vs. Inhalation Data

Two unit risk estimates were extrapolated from oral exposure data. The process used is scientifically consistent with the process used by states when faced with similar needs. In most cases

the extrapolation is best based on estimates of blood levels, either measured or calculated through use of a pharmacokinetic methodology, rather than based solely on an overall body weight comparisons. However, there is concern in the 1996 NATA report that one of the highest UREs , that for quinoline (3.4e_03), is based on an inhalation potency derived from oral exposure values.

Recommendation: For 1996, provide an estimate of the potential variability of the oral to inhalation extrapolation.

3.2.3.5 Specific points on Butadiene and Nickel

The Agency should clearly explain in the document the scientific reasoning used, when it deviates from the written policy in NATA. The toxicity RFC s for butadiene and nickel used in the NATA report are not in IRIS, or have been removed from IRIS. There are apparently widely differing values proposed for the RFCs. A calculation comparing the range of values under consideration should be made and discussed in the document.. The current IRIS number for butadiene is more consistent with the number from CalEPA (CalEPA number 1.7e_4, current IRIS value 2.8 e_4, NATA value 2e_5). A clearly defined selection rationale is needed for both nickel and butadiene if the IRIS RFC is not used.

Recommendation: Show the risk rationale for butadiene and nickel and the range of possible estimates using different dose response values.

3.2.3.6 Aggregation of Risks

Limitations in dose response information play an important role in the aggregation of risk estimates. When dose response data changes, the implications are different for cancer and non-cancer risk calculations, as shown below.

3.2.3.6.1 Aggregation of the Cancer Risk and Dose Response

The precision in the dose response data is important when aggregating cancer risk. The approach to aggregate the cancer risk used in NATA Section 3.4.2 is acceptable. Because only the values that exceed a specified risk level are used when aggregating risk the risk values are more likely to be in a linear portion of the dose response curve. It would be of interest to assess whether there is a substantive difference in the aggregate or apportioned risk when risk levels over 1xe-6 are compared to risks over 1xe-5. The use of the independence formula rather than the non-additive aggregation is prudent, but the implications of the use of either approach on the overall estimates of risk from the mixture are not apparent. The approach to PAHs and POM mentioned in bullet 2 is reasonable. A relative ranking of the chemicals that "drive" each aggregate risk is partly determined by the dose response selection. It could be placed in a footnote.

3.2.3.6.2 Risks Other Than Cancer and Dose Response

In the case of risks other than cancer, the risk is not linear throughout all possible range of exposures. However, the risk probably approaches linearity over the range of ambient air exposures. The probability that an exposure exceeds a reference value needs to be first established and followed by assessment of the dose-response relationships. This process must show the severity of the outcome. If the dose-response data are based on different outcomes with some very severe compared to others, the short-term reversible effects could be ranked incorrectly. The selection of endpoint could also alter the dose-response reference values.

None of the 33 compounds in the 1996 NATA are likely to exhibit linearity throughout the entire range of dose-response. It is important to keep in mind that these compounds were selected from 188 HAP chemicals based on their higher toxic risk and potential exposure in urban areas . When NATA is extended to less potent compounds, deviations from linearity in the dose-response relationship could be of greater importance.

3.2.3.7 Other Issues With Respect to Dose Response

Some members of the Panel cautioned against using the available dose-response RFCs in combining risk estimates. The aggregation of risks and grouping by target organ is an undefined approximation and for some members of the Panel that is a concern.

The grouping of hazards by endpoint or by target organ is helpful for planning of interventions to reduce risk. Interventions usually consider route of exposures. It is important to determine whether the reference risk value is valid across target organs when a compound has toxicity in different organ systems. Since NATA is a screening rather than a regulatory process, the errors in including compounds with a common target organ and different mode of action are less important. Combining different modes of action should be less of a problem in assigning risk drivers.

Recommendation: The approach for aggregation of the "non-cancer" risks will underestimate the portion of the population at risk, due to the use of median exposure values. This should be noted clearly in a footnote. However the approach should be adequate for 1996 NATA objectives of obtaining initial risk rankings and prioritizations.

Recommendation: For the 1999 NATA it would be of interest to determine the aggregate risk values using percentiles. Comparison of approaches could be used to provide an approximate sensitivity analysis. Noncancer dose response must be completely addressed in the 1999 NATA.

3.2.3.7.1 Indirect exposures

The omission of indirect routes of exposure is a serious public health limitation in the NATA risk estimates that must be addressed in future assessments. The persistent bioaccumulating toxics (PBT's) should at least be assessed for food and water impact as they represent a major potential health concern at the state regulatory level.

Recommendation: The 1999 NATA should include the effects of indirect (non-

inhalation) exposures for PBTs.

3.2.3.8 Uncertainties in the Dose Response

Uncertainties listed in Section 3.4.4.1 (of the NATA document , pp. 49-51) are included in the URE but these are the standard uncertainties and are not unique to the NATA process. The report fully emphasizes these risks but in doing so tends to overstate the uncertainty in the NATA process. In fact, every risk assessment has these uncertainties. The problem is a more general own owing to the lack of scientific study and data. That uncertainty should be clearly conveyed to the public as what it is, a failure in toxicological research agendas. The reference concentration (Table 3-7) uncertainties in the NATA document, in which UF and MF are combined, are inappropriate, confusing the NATA and dose-response uncertainties. The discussion in the current NATA draft seems to indicate that the NATA process increases the dose-response uncertainty found in population risk calculations. It does not do so.

Recommendation: For the 1996 NATA more clearly indicate which of the uncertainties are due to the ASPEN/HAPEM process and which are due to the more general risk assessment process. It would be helpful to compare the process used in NATA with that employed by the states that have established standards based on acute and sub-acute effects and cancer endpoints.

3.2.3.9 Other General Comments

a) The assumption of additive effects among chemicals is a component in the 1996 NATA assessment. What are the procedures, if any, whereby this assumption will be re-visited in future assessments as data become available that may challenge this simple model?

b) The dose-response values cannot be viewed entirely independently from the exposure assessments. Given that the air exposure calculations appear to underestimate exposure, it

seems prudent to use upper-bound estimates of the cancer and noncancer dose-response factors. Yet, the net result is difficult to understand (and use) in relation to either (1) accurately assessing the absolute risks posed by the individual HAPs, or (2) evaluating the relative risks posed by these chemicals.

c) It is important to the NATA process that the IRIS values be current and that timely update of the database be given a higher a priority. Page 92 of the NATA document, item 6, states, "EPA is currently reassessing the carcinogenic effects of 19…[of the air toxics]...unit risk estimates could change substantially…" That would change the risk screening numbers and potentially the drivers identified. Change in IRIS risk estimates is not an uncommon occurrence. States that have set HAP standards always include a date at which the process of setting standards is started and a future date for the revision. States often do not *(or do?? – please clarify)* change standards or HAP guidelines as reference values are revised. Given the dynamic nature of the science EPA could consider a similar procedure.

d) The tables of toxicity factors need to be checked for accuracy and clearly referenced. For example, the table has an IRIS RfC for TCE, but IRIS text says "not available at this time." In one case the IRIS value was removed from the database. It is occasionally the case that states find, when setting HAP levels, that IRIS data has been removed from the IRIS data base. When this occurs the states refer back to the most recently published value in IRIS and verify the RFC with current peer reviewed data. They do not stop the regulatory process. When data are needed for screening, EPA should apply a similar process.

e) Page 47 footnote. This says that the HEAST tables were used, but it is not clear for which values. When HEAST tables (which may contain values that are NOT verified by the Agency) are used, the validation procedure used by states *(what is this procedure?? Citation??)* should be followed.

f) Page 90. Nine lines from bottom. It says that the RfC "or similar value" was used. Please

expand on this comment with an example in a footnote.

Recommendation: The Panel recommends that certain changes are needed for clarity. For RfC's add a column giving the IRIS statement about the overall uncertainty. The "citation" (e.g., IRIS, CalEPA) should enable the reader to easily find a complete source document for the value used. If the authors have performed additional calculations, this process should be clearly identified and a reference provided to that additional information.

3.2.3.10 Micro Environments and Dose Response

Variations in HAPEM chronic risk estimates are more dependent on the exposure estimates than on the dose-response used to calculate risk, but dose response factors will be more important when addressing acute risks. The exposure levels determine the health risk in longer-term studies. In the case of risk estimates from shorter term exposures, which are not addressed in the 1996 NATA, the acute non-cancer dose response data will be important.

Recent changes in HAPEM have improved the exposure modeling and the potential ability to obtain short-term risk estimates. The use of 3-hour time blocks of exposures and stochastic match up of the exposures is very important for the acute risk estimates. Once such an approach is properly implemented (and the accuracy of the local inventory verified through comparisons with the local, county and state exposures), acute risks can be included as part of the NATA. Stronger dose response rationale will be needed at that time to avoid underestimation of the actual short-term risks.

A sensitivity evaluation is needed with respect to the 37 micro environmental parameters to show the quantitative relationships between risk and the dose response. This should be considered for the next assessment. It is not a basis to delay the 1996 NATA release. PEN, PROX and ADD *(can you elaborate here??)* need to be sufficiently conservative to assure that the exposures reflect sources actually impacting the communities.

There is an ongoing issue with background levels that is most important in the non-cancer area. EPA needs to provide a discussion of the possible magnitude of the background effect. Because the acute dose-response data are based on cumulative thresholds, all exposure sources need to be considered including the added risk over background. The backgrounds from long-range transport and natural sources could raise the exposures toward the threshold, thereby increasing the risk contributions from other sources.

Recommendation: For the 1999 NATA investigate the implications of limitations in target organ acute toxicity data on the NATA estimates of short term risk and risk ranking by compound and by geographic regions.

3.2.4 Charge Question 4

What are the strengths and the weaknesses of the overall conceptual approach to risk characterization used in this assessment? Given the underlying science and the intended purposes of the assessment, can the Panel suggest ways in which the risk characterization could be improved?

a) Is the method used to aggregate cancer risks appropriate? The aggregation of carcinogenic risk within two categories, based on weight-of-evidence classifications, is of particular interest.

b) Is the method used to aggregate non-cancer hazards appropriate? The summation of hazard quotients within target organs, the categorization of sums by ranges of uncertainty factors, and the inclusion of all target organs (as opposed to only the organs associated with the critical effect) are of particular interest.

3.2.4.1 Strengths of the Overall Conceptual Approach

The overall conceptual approach to the risk characterization is reasonable. It generally follows the guidelines and procedures of risk assessment (with exceptions noted later for mixtures). Pollutant-specific risks to populations are generated and pollutants are grouped into national and regional risk

drivers as well as important national and regional contributors. Risks of multiple pollutants are aggregated to generate national cancer and non-cancer hazards by sources (major, area, on-road mobile, non-road mobile, and background). However, as detailed subsequently, some of the key specific elements in implementation of the conceptual approach are not consistent with assessment guidelines or current best practices.

OAQPS faces two challenges in characterizing risks from this analysis. First, it must find a technically valid way to aggregate predictions and summarize findings for a very large set of individual estimates for individual chemicals at numerous locations. It is a very difficult task to summarize information in a way that does not bury some of the important fine points. Second, it must develop a lucid presentation for consumption by both a sophisticated technical or policy analysis audience as well as the general public. In many areas OAQPS has done a good job and met these challenges. However, there are also a number of key areas where decisions to summarize and generalize findings are questionable.

This charge deals with the integration of the dose-response assessment and the exposure assessment. Thus, it encompasses the strengths and weaknesses of these risk components.

3.2.4.2 Weaknesses of the Overall Conceptual Approach

Some fundamental issues are raised, but not fully discussed about the scope of the NATA, namely issues about effects from less-than-lifetime exposures and total exposure to air toxics. The assessment includes only chronic health effects and not acute or subchronic health effects. In actual environmental health assessments, acute health effects are very important for the evaluation of mortality and morbidity from outdoor air pollutants. By not including acute or subchronic health effects in this assessment, it is not possible to evaluate critical short-term health effects of outdoor air pollutants.

Recommendation: For 1996 NATA, include more discussion of the implications of considering only chronic health effects. For 1999 NATA, include less-than-lifetime exposure health assessments, exposure assessments, and risk assessments. Some of these actions will

require the development of standard assessment guidelines and new evaluations and entries into IRIS.

The NATA focuses on inhalation risks from outdoor sources of air toxics, including exposures that occur outdoors and indoors as related to penetration of outdoor air. If exposures from indoor sources of air toxics are not included, the potential risk to the public from total exposure to these chemicals cannot be understood, given that some air toxics have substantial and others insignificant indoor sources. Additional pathways (e.g., some air toxics deposited on the ground or bodies of water can enter the food chain) are not considered. Basically, even if the NATA findings on inhalation risks from outdoor sources of air toxics were perfect, important elements of risk from these chemicals are being ignored, rendering the entire assessment more limited than portrayed. Such "missing" information will, in some cases, have a significant impact on total risk. Air toxics regulatory authority covers outdoor sources, including all pathways, making this important for NATA. However, including risks from indoor sources is important to the "total" risk issue and provides guidance to risk managers and the public on all of the potentially most effective approaches to reducing risks from these chemicals. It is also essential when computing health effects when the dose-response function is nonlinear or has a non-zero threshold, since outdoor sources may not be sufficient to cause thresholds to be exceeded (or steeper portions of the nonlinear dose-response function to be reached), however, such thresholds may be exceeded when other sources of exposure are included.

Recommendation: For 1996 NATA, increase discussion of potential impacts of total exposure, including the indoor source issue. For 1999 NATA, include other sources of exposure in the risk analysis.

OAQPS states that this assessment was undertaken to: help identify pollutants of greatest potential concern, prioritize efforts to reduce emissions, provide a baseline for measuring future trends, and help set research priorities. The document appears to discourage applications on a local or regional level, yet it provides information at the county level. Clarification of the appropriate scale for application of the information would be useful.

There is much discussion of how the NATA results could be over estimating risk, but not much in terms of how the results might be underestimating risk. For example, the demographics of available time-activity databases and the population therefore simulated by HAPEM is skewed towards middle class workers, and has been criticized for failing to take into account less fortunate populations and their lifestyle and workplaces. The factors in HAPEM are generalized factors, which do not account for variability in exposure to outdoor air across the country, e.g. areas of the country where windows are left open for more days of the years than others. Day-to-day correlations in activities are not preserved in the activity pattern sequences, which means, for example, that a day 1 activity pattern may specify a house with an attached garage, and in day 2 a house with no attached garage. Furthermore, the exposure estimates represent midrange estimates, and results from the high end of exposure are not provided. On the hazard number side, OAQPS relies in many instances on MLEs, which are based on "best estimates" rather than high-end estimates for some chemicals (Table 3-5). The total risk estimates also do not include the estimated risks from diesel, though elsewhere in the NATA document diesel is indicated to be a significant source of hazardous air pollutants. Finally, as is discussed elsewhere, a check between HAPEM model predictions and monitoring data show that the model often (indeed, in most cases examined) underestimates observed ambient concentrations.

Recommendation: For 1996 NATA, provide a more balanced discussion of the possible sources of under- versus over-estimations of HAP exposures and risks.

3.2.4.3 Aggregate and Cumulative Risk Issues: General Issues

The NATA evaluates the relative importance of various source sectors (major, area, mobile on-road, mobile non-road, and background) by aggregating health risks (cancer and noncancer) across pollutants to estimate populations affected by different source sectors. The procedure used for aggregating cancer risks is based on three underlying assumptions. They are linearity, additive effects, and comparable units. To derive UREs, a linear dose-response model is used to extrapolate risks from high to low doses. To estimate population risks, linear extrapolation is again applied to the range of population exposures based on UREs. The assumption of linearity will not be violated if dose-response

curves used in the procedures are linear. Even if some of the dose–response curves are not linear; it is assumed that they are approximately linear around UREs. It is also assumed that they are approximately linear from UREs to population exposure levels.

The assumption of additive effects is used for estimating cumulative risks resulting from multiple pollutants. Since there is no good information on the interactive or synergistic effects among multiple pollutants, it is logical to assume that all pollutants act independently and additively. This assumption allows the risks of multiple pollutants to be computed by simply adding up all individual risks. However, due to the lack of related studies, the validity of this assumption is difficult to test.

The third assumption follows the second assumption, in that summation of risks is only meaningful if the risk units to be added up are equal or at least comparable. Population risks are determined by population exposures and UREs. To aggregate population risk across pollutants appropriately, population exposures should be unbiased and UREs should be comparable. An ideal URE should have the property of reflecting the severity of cancer risk with minimum uncertainties. An URE is actually an estimate of cancer potency with uncertainties. There are two kinds of uncertainties associated with UREs. One is the weight-of-evidence (i.e., the classification of known, probable, and possible human carcinogens). Another uncertainty involves the actual value of the UREs (i.e., upper bound estimates). The aggregation of cancer risks based on weight-of-evidence has the advantage of increasing comparability. Determining UREs using the same method, such as MLE, for all pollutants, is another way to increase the comparability of risk units.

The same underlying assumptions can also be used to judge if the method used to aggregate non-cancer hazards is appropriate. Risk characterization is based on exposure and dose-response curves, regardless of whether it is a cancer or non-cancer risk. However, the nature of the RfC is more complicated than the URE. To generate a risk unit for non-cancer hazard, the NOAEL or LOAEL is divided by an uncertainty factor and a modification factor (UF X MF) is used to determine the RfC. For the air toxics in NATA, the values of UF X MF range from 1 to 1,000. This uncertainty factor moves the RfC away from its original dose-response curve. Therefore, unlike the URE of cancer risk,

it is not possible to apply linear extrapolation to population exposure levels from RfC's. To evaluate risks at population exposure levels, the HQ is generated as a function of exposure by dividing by the RfC. The HQ cannot be interpreted as a probability of non-cancer risk. HQ is a measure of potential health risk, but lacks a clearly defined meaning of risk.

To add up hazard quotients across pollutants within target organs, the assumption of additive effects is needed. This assumptions is often invoked even though, within the same target organ, different pollutants have different modes of action. For many such effects, additivity is a simple and logical assumption, but it lacks the support of empirical data. Regarding comparability, the RfC is far less adequate than URE. The UCL used for URE is a conservative measure with statistical reference, while the UF is a measure of uncertainty without theoretical (statistical or biological) justification. Because of the large size of the uncertainty factor in certain cases, the UF's used could be a key factor driving the estimated population risk. Take the example of acrolein; the UF of 1,000 is assigned to its RfC due to interspecies extrapolation (a UF of 10), lack of chronic studies (another UF of 10), and accounting for sensitive human populations (an additional UF of 10). Because of the above uncertainties, the RfC (2.0E-05) of acrolein becomes 1,000 (10 X 10 X 10) times lower than the LOAEL (2.0-E-02) estimated from animal studies. The resulting high computed values of HQ for acrolein contribute to estimated risk across a large affected population, however, this is mainly due to the large uncertainty factor and not due to the high potency (or low threshold) of non-cancer health effects. As a leading national hazard driver, the estimated population risk of acrolein can certainly be attributable partially, maybe even largely, to the UF of 1,000.

For noncancer hazards, efforts were also made to increase comparability for aggregated risks. To increase comparability of HQs across different pollutants, TOSHIs were developed grouping noncancer risks by target organs. The categorization of sums by ranges of uncertainty factors (UF>100 and UF 1-100) is another way to increase the comparability of risk aggregation.

The issue of background exposure to some of the air toxics was raised in the text. However, it is difficult to discern its chemical-specific impact. For example, figures 5-3 and 5-4 include background as a source, suggesting a cancer risk in excess of 1 in a million. This is a significant statement, making

more discussion useful. For example, it would be useful for the reader to know which compounds in Figure 5-6 had a significant background component to the risk (note- this is a figure of exceedances of HQ levels based on all source sectors). A simple indicator (e.g., use of an asterisk for those chemicals having significant background contributions) would be helpful. As noted below, a general recommendation is made for greater explication of the reasons why different compounds are predicted to be risk drivers.

Recommendation: For the 1996 NATA expand the discussion of the rationale for the approaches used to aggregate cancer and noncancer risks and the impacts of these approaches on uncertainty. Also, expand the discussion on the possible extent of the influence of background concentrations and other model assumptions on the risk outcomes.

3.2.4.4 Cancer Risk Characterization

Known human carcinogens are summed separately from probable human carcinogens in the NATA document. Probable human carcinogens are lumped with possible carcinogens. This is not conventional, nor is it appropriate. The only difference between known and probable classes of carcinogens is data from human studies, and human studies of these compounds are relatively rare. Thus, it seems more appropriate and certainly more precautionary for OAQPS to combine the knowns and probables separately from the possibles. Also, OAQPS should provide an estimate for all types of cancers summed together and then break the results out by group.

Recommendation: For the 1996 NATA, evaluate the impacts of combining the A and the B1 carcinogens, leaving the B2 carcinogens as a separate entity and see whether this changes the conclusions about risk drivers or the risk characterization. If this evaluation has significant impact, decide on the optimal approach for the main presentations and provide an appendix with the alternate approach, along with an evaluation that integrates Class A, B1, and B2 carcinogens. When deciding on one approach over another, document the rationale for the selection and any history of use of a particular approach.

Uneven and unsystematic biases may amplify or cancel each other following the steps of modeling process, and thus, the end results might change the actual rank order of risks in an undesirable manner. For example, all Unit Risk Estimates (UREs) used in this assessment are based on linear extrapolation. For some pollutants, which are less than linear, this process may overestimate the risk. In contrast, most UREs used in this assessment are based on upper confidence limit (UCL), but a few are based on maximum likelihood estimates (MLEs). Estimates based on the MLE are less conservative than tjpse based on UCL.

Recommendation. For 1999 NATA, EPA should complete the revision of the cancer assessments for the air toxics of interest, using the most up-to-date data and assessment approaches and document them in IRIS.

It is very helpful that OAQPS identifies those chemicals disproportionately responsible for the risks in the study; again, using the analysis to identify a priority list of HAPs is a useful and practical application for the study. However, this section does not discuss or take into account some contaminants previously identified in the report as particularly underestimated in the model. In particular, chromium, cadmium, and lead are underestimated in the model.

Recommendation. For the 1996 NATA, the section that discusses which HAPs are important risk drivers should take note of the possibility that other compounds underestimated by the model could also be risk drivers.

There is a concern with the "addition" of upper bound cancer estimates to estimate the overall aggregate risk. The sum of multiple 95th percentile yields a value that is generally much further out on the tail (i.e., much more conservative) than the 95th percentile value for the sum. That concern is especially valid when the slope functions differ significantly from chemical to chemical or if an exact risk for a specific population is desired. In the case of former, comparison with the MLE estimates should be used to reveal any discrepancies in estimates that might occur due to adding multiple upper 95th

percentile values that differ significantly from their respective MLE estimates. That should be noted in a footnote of the report. In the case of the latter, it can be noted that NATA is not attempting to determine the exact aggregate cancer risk for any area, but to determine relationships between regional risks and risk drivers. Thus while the use of MLE estimates would be more accurate, when summing cancer risks, the summation of upper bound estimates may in may cases be employed without altering the risk ranking of the compounds.

The precision of the dose-response data is important when aggregating cancer risk. The approach to aggregate the cancer risk used in section 3.4.2 is acceptable. Only values that exceed a specified risk level should be used when aggregating risk because the risk values are more likely to be in a linear portion of the dose response curve. It would be of interest to assess whether there is a substantive difference in the aggregate or apportioned risk when risk levels of 1xe-6 are compared to 1xe-5. The use of the independence formula rather than the non-additive aggregation is prudent, but the implications of the use of either of the approaches are not apparent. The approach to PAHs and POM mentioned in bullet 2 is reasonable. The ranking of the chemicals that "drive" each aggregate risk would be helpful and could be placed in a footnote.

Recommendation. For the 1996 NATA, please clarify this issue of the difference between seeking a relative ranking vs. an absolute risk and the differential influence that conservative assumptions employed when aggregating risk may have on these.

3.2.4.5 Non-Cancer Risk Characterization

A HQ and HI approach are common means of assessing non-cancer risks. As everyone agrees, there is a high degree of uncertainty in this approach. However, the means of doing this calculation in this draft NATA document do not follow EPA guidelines and are scientifically questionable and therefore need to be revisited.

The HI methodology is commonly accepted for chemicals having a common mode/mechanism of action. In the absence of data, some assessors default to using a common organ (in accordance with

EPA mixtures assessment guidelines). The key phrase is in the absence of data. In some cases, chemicals having known different modes/mechanisms were added (e.g., formaldehyde which produces nasal effects was added to cadmium which produces lung effects through different mechanisms).

Recommendation for the 1996 NATA. Either create the HI based on mode/mechanism of action or remove the HI, applying it properly in the 1999 NATA.

The calculation of greatest concern is the target-organ-specific-hazard index (TOSHI) (pp46ff; pp90ff). As described on p 46 and p 92, TOSHIs were developed by summing the HQs (the exposure divided by the RfC) for individual air toxics that affect the same organ or organ system. It was calculated by taking the RfC for a chemical based upon the critical effect and dose to one organ and transferring this RfC to all other organs affected by that chemical. The RfC methodology begins with the identification of the "critical effect" commonly defined as that endpoint having the lowest NOAEL (or LOAEL) (or the benchmark equivalent); it is a human equivalent concentration, including an estimate of dose to the target organ. Uncertainty factors and modifying factors are thenused, according to the guidelines. An RfC results. Often other organs are affected, but at higher NOAELs, so they are not the "critical effect". An RfC based on such a higher NOAEL would be higher. Dose calculations would also be different. Even more uncertainty can result. If EPA wishes to use a TOSHI approach, it is essential that EPA goes back to the database for each chemical and actually develops TOSHI's with a high level of scientific rigor. Without that effort, they should be eliminated from the document.

It is recognized that the IRIS database for many of these substances is out-of-date, but timing considerations for revision of this version of NATA may restrict the TOSHI reevaluation to this IRIS database. Although this would compound any errors due to the date of evaluation, it is preferable to the incorrect approach used. It is also troubling that this problem is not clearly identified as a significant limitation. For example, page 47, item 4 says that TOSHI will be used "(data permitting)", implying a degree of accuracy in the model that in most cases cannot be justified.

With respect to table 3-7 and the discussion on TOSHI in section 3.4.3, some chemicals appear in more than one group (e.g., Cr is listed for respiratory, liver/kidney, and immune system. Please clarify whether they are counted more than once. Are they counted in both or in only one? Is this double counting?

Recommendation for the 1996 NATA. Either reexamine the IRIS database and calculate target-organ specific "RfC's" based on NOAELs (or Benchmark dose equivalents) for these other organs or delete the TOSHI. If the TOSHI are deleted here, they should be developed (with up-to-date data) for the 1999 NATA.

In the case of risks other than cancer, it is true for many chemicals the risk is not linear throughout all possible exposures. But risk probably approaches linearity over the range of ambient air exposures (Dave et al, how certain are we of this, given how poor the dose-response data base is?). The probability that an exposure exceeds a referenced value needs to be first established by EPA, then the dose-response relationships with respect to risk and severity of outcome can be assessed. None of the 33 compounds are likely to linear throughout the entire range of dose response. It is also important to keep in mind that these compounds were selected from 188 HAP chemicals based on assumptions about their higher toxic risk. When NATA is extended to less potent compounds, knowledge of the shape of the dose-response could be important.

The endpoints in table 3-6 are highly variable but can fairly characterize a national or statewide risk when viewed from a national level. The potency differences are usually more reflective of differences in the pharmacodynamics of the exposure than differences in mode of action. The parmacodynamics differ by route of exposure, inhalation generally leads to the highest short-term blood levels. Therefore, it is important to show which HQ is based on an inhalation exposure and which is based on another pathway. The differences could be normalized using a physiological pharmacokinetic analysis.

3.2.4.6 Alternative Risk Evaluations

The integration of an exposure assessment with a health assessment is extremely difficult, even under data-rich circumstances. Because this luxury does not exist for air toxics, there will be considerable errors in unknown directions as data collected for one purpose are used for another purpose in unvalidated models. It therefore would be of value to know the relative influence of errors in exposure vs. errors in health factors. One issue of particular concern is the magnitude of the net uncertainty factors in the RfC's. It would be of interest to know the degree to which the uncertainty was driving the risk. For example, acrolein is identified as having a higher noncancer risk than other compounds. Is this due more to the uncertainties in the dose-response assessment or the exposure assessment?

Recommendation: For the 1999 NATA, consider running the risk analysis using alternative toxicity values for a few key chemicals to provide a scenario-based approach to the importance of these values in the overall assessment. This action should be taken in the near future to help inform priorities on research areas.

Many places in the text discuss the uncertainties and variabilities inherent in NATA and the current inability to quantify the impacts of these unknowns. However, many choices were made in the assessment, e.g., using modeled exposure estimates without estimates based on the measurements of exposure from various sources like NHEXAS, TEAM, or the other literature sources; using one health value rather than another (e.g., for butadiene), and it would be interesting to consider some selective groundtruthing for some selected air toxics. OAQPS should select the air toxics for such an analyses based on available databases. Benzene is one example where a groundtruthing exercise would be informative.

The NATA risk classification for all the air toxics is based upon a reasonable logic that the broader the risk distribution, the more likely the source was local. For some of the air toxics, the database should be rich enough to perform a source apportionment. For example, source apportionments of benzene have been published years ago and more recent ones may be available for use. For example, a review article by Wallace (Clinical and Experimental Allergy, 1995, 25:4-9) illustrates a source apportionment based on the TEAM studies. This analysis estimates that 82% of

benzene emissions are due to cars, 14% are due to industry, and the remaining 4% are due to cigarettes, personal, and home sources. However, this same analysis shows that 40% of benzene exposure is due to smoking cigarettes, 5% is due to environmental tobacco smoke, 18% is due to automobile exhaust, 18% is due to personal activities, 6% to home sources, and 3% to industry sources. When such information is available, it should be used.

Recommendation: For the 1996 NATA, select 1 or 2 air toxics having substantial databases and develop a risk assessment based on their data and compare it to the model results of the current draft. For the 1999 NATA, explicitly incorporate all the credible data in the assessments and incorporate the results of validation/evaluation research on models.

3.2.4.7 On the Issue of Children

On page 99, under 5.5.3, paragraph 1, the NATA document states, "it is necessary to consider adults and children separately." On page 100 in top paragraph discussion on children; line 4, the text states, "dose-response assessments for non-cancer effects developed by EPA… do not currently include separate reference concentrations…for adults and children." These comments are misleading. Indeed, there are not separate RfC's. As stated in several places in the document, the definition of the RfC includes the coverage of "sensitive sub-groups." This part of the definition is derived from the use of an uncertainty factor of up to 10 for intraspecies extrapolation (i.e., from average to sensitive sub-groups). There has been much debate engendered by the Food Quality Protection Act and it's requirement for an additional factor of 10 to ensure protection of children from pesticides. Is EPA implying that additional protection (beyond the standard uncertainty factor) is required for children exposed to air toxics? If so, what is the scientific basis for this? As mentioned above, the RfC, being based on lifetime exposure, is not an appropriate index for children who have not lived for 70 years. Where children are a special concern, the data need to be evaluated and assessed appropriately. The paragraph ends with a comment about higher TOSHI's for adults than for children. This only compounds errors. The entire discussion in this section needs to be revisited.

Recommendation. For the 1996 NATA, the discussion of children should be clarified to indicate that they are an important life stage to be considered and therefore are already incorporated in the chronic assessments. However, the exact degree to which these assessments either under- or over-estimate risks to children is unknown. For the 1999 NATA, the incorporation of less-than-lifetime health assessments should permit more focus on children.

3.2.4.8 Additional Clarification Issues

For the most part, the document is internally consistent, except for a few instances.

a) Page 18, L 4 says that "current Agency risk assessment…guidelines" were used. As described elsewhere, in some cases the assessment practices of others (e.g., CALEPA) were used and procedures can be different;

b) Page 35, Microenvironmental data, para 1, last line. This says that an ADD factor was used "that accounts for …i.e., indoor emission sources." However, in many other places the document said that indoor sources were not considered. Page 37 says that the ADD factor was set to zero;

c) Page 84 discusses the interpretation of census track and higher order aggregations. As mentioned elsewhere, the census-level is too uncertain to be used. Then the next paragraph says that 'The results of the exposure assessment are only meaningful when examined at the individual county level or above." Is this "meaningful" comment really true, given the caveats?

d) Page 91 line 2. This sentence says that the "risk characterization focused on results at the national level, which is the level at which EPA believes the results are most meaningful." If this is correct, why provide county-level data?

e) Page 41, Section 3.4 The risk characterization section is a mixture of dose-response assessments and risk characterizations. They should be separated for more clarity;

f) Page 42 line 11 from bottom. Clarify terminology: why is cancer a risk and non-cancer a hazard;

g) These analyses were "based on the median exposure within each of the approximately 61,000 census tracts nationwide."(Page 93 and elsewhere in this area.). In many earlier sections, the document states that the variability of the data at the census tract level causes the authors to only show the information at the county level. Other places say that the exposure assessment is "only meaningful when examined at the individual county level or above." (Page 84). It would be useful to further justify the quality of using such aggregations of information;

h) The document should be slightly reorganized. Chapter 4 is the exposure assessment, but Chapter 5 jumps right into the risk characterization. A new Chapter 5 should be constructed to contain the hazard identification and dose-response information for the health assessment. The next chapter would be the integration—the risk characterization; and

i) Page 99, under 5.3.3, paragraph 1: This section on aggregate TOSHI implies that non-cancer aggregate risk is more complex than cancer risk because for non-cancer, "it is necessary to consider different toxic effects and mechanisms…" However, cancer mechanisms also differ, so this should be reworded.

Recommendation: For the 1996 NATA, consider clarifications of the above points.

3.2.5 Charge Question 5

Although EPA has concluded that available data are not sufficient to develop a reliable quantitative estimate of cancer unit risk for diesel emissions, it is clear that this pollutant class

may be of significant concern in a number of urban settings. The risk characterization in this report includes a discussion of diesel particulate matter to help states and local areas frame the importance of this pollutant compared to the other air toxics. In the context of this assessment, is the discussion in this report regarding making risk comparisons among other air toxics appropriate? Can you provide any suggestions that would improve upon this approach to comparing the toxic health effects of diesel particulate matter with other pollutants?

The inclusion of diesel exhaust particles (DEP) as an air toxic in the context of this Assessment is arguable. It can be argued on the basis of: a) the lack of a unit risk estimate (URE); and b) the complex nature of DEP; that the material should not be included at all. It is the view of the Panel, however, that it is appropriate for DEP to be included in some manner in this assessment. There is a widespread and longstanding concern for the health impacts of DEP, and the public and other users of the NATA would expect it to be included. The exposure to DEP is ubiquitous, and the exposure assessment included in this document provides useful perspectives. Although the level of risk is not known and continues to be debated strongly, some level of risk is plausible.

The Agency was interested in whether or not the caveats they included in the NATA document are consistent with the recommendations of the Clean Air Scientific Advisory Committee concerning the diesel Hazard Assessment Document (HAD) (not yet published). In general, the caveats concerning the uncertainty of the level of risk and the decision not to use a specific URE for lung cancer were appropriately stated, with the exception of perhaps two issues. First, the wording suggests that CASAC endorsed the range of probable cancer risk portrayed in the document. Although CASAC agreed to close on the diesel HAD with the range included, there was not consensus regarding the appropriateness of its inclusion or the validity of the values bounding the range. Opinion was divided, thus, although CASAC agreed that inclusion of the range would not prevent closure, there was not a consensus to endorse the range and there were members who were opposed to its inclusion. Second, the explanation provided in the NATA document was not sufficient to give an uniformed reader a good sense of why the Agency did not adopt a URE for DEP cancer risk, or why it did not adopt the California URE as a backup (as it did for some of the other air toxics).

The attempt to treat the risk from DEP in parallel with the risks from other species results in an obviously awkward construction. Given that there is no acceptable URE for DEP cancer risk for this exercise, the insertion of repeated statements that the Agency believes that DEP is one of the most important of the air toxics appears incongruous, and a circumvention of the process used for the other species considered. Without better explanation, the reader perceives that if the Agency decides an air toxic is important, it can state this as a belief without the rigor of establishing a URE. The present explanation does not give the reader a very solid understanding of why this conclusion was reached for DEP. It is understandable how exposures, or at least regional concentrations, of DEP are estimated, but it is not very understandable from the present treatment what the situation is with respect to risk.

The Panel suggests that the Agency develop a more thorough explanation of the current status of knowledge concerning DEP health risks, and place it in one section devoted to that purpose. The section need not be a separate chapter, nor need it be very long. Perhaps a few pages would suffice. The Panel also recommends that the section include a summary of non-cancer as well as cancer risks. It is plausible that the non-cancer health burden from environmental diesel emissions may exceed the health burden from cancer. It would also be useful for this section to mention links between health issues associated with DEP and those associated more generally with ambient fine particulate matter (PMfine). Because DEP comprises a minor, but significant portion of PMfine in urban inventories, and a major portion in certain microenvironments, the health effects of DEP must be integral to those attributed to PMfine, including possible mortality and morbidity effects associated with cardiopulmonary disease, influenza and asthma. Mentions of DEP at other steps of the Assessment can be referenced to this section. As a result: 1) the reader will have a better understanding of the Agency's views and the reasons for them; and 2) the construction will appear less awkward and will give less impression of a circumvention of the process established and used consistently for the other air toxics.

The current wording that the Agency believes DEP to be one of the most important air toxics should be tempered to more closely reflect the degree of uncertainty that presently exists regarding DEP risk. It is reasonable that as part of the expanded explanation, the Agency summarizes their view that the evidence suggests that DEP may be very important. Under the circumstances however, it does

not seem reasonable for the Agency to conclude that DEP is one of the most important air toxics on the basis of belief. In fact, the Agency may be correct in its belief, but it may also be incorrect. If we knew with acceptable certainty, we would have an acceptable URE.

Recommendation: Diesel emissions should be included in the NATA. A specific section should be devoted to a clear, succinct explanation of the basis for the Agency's conclusions regarding health risks from DEP. The section should touch on both cancer and non-cancer risks, and links to risks attributed to ambient particulate matter. The wording should be moderated to more accurately reflect the uncertainty of the health risks and CASAC's position regarding the cancer risk range in the Diesel HAD.

3.2.6 Charge Question 6

Given the limitations inherent in this preliminary assessment, have uncertainty and variability been appropriately characterized?

Can you suggest ways that the characterization of uncertainty and variability could be improved, made more transparent, or integrated more effectively into the risk characterization? b) Can you suggest methods for quantifying individual as well as composite uncertainties associated with the emissions inventory, dispersion modeling, exposure modeling, dose-response assessment, quantitative risk estimates, and accumulation of risk across air toxics?

The NATA 1996 document provided to the SAB presents a variety of qualitative discussions of sources of uncertainty in the risk assessment and a top-down effort to characterize the overall uncertainty in the analysis. We support the overall approach of estimating the top-down uncertainty factors based on the multiplicative elements of the assessment. A top-down approach is well suited to the preliminary nature of the overall assessment. In contrast, a more detailed effort to propagate uncertainties from the bottom up would not be viable in the current assessment, given the limitations of the baseline analysis.

Although the NATA review panel generally supports the use of a top-down approach, the current implementation requires significant additional work. In particular, the methods and supporting information used in the assessment are not yet adequate to allow the assignment and propagation of probability distribution functions for representing uncertainty in each of the NATA components (emissions, fate-and-transport, exposure and dose-response).

The top-down uncertainty estimates presented in Section 5.5 of the NATA report consider three factors: modeled ambient concentrations from ASPEN, the ratio of personal exposures to ambient concentrations, and dose-response factors. The monitor-to-model comparison used is a reasonable approach for estimating uncertainty in the ASPEN modeling results, and makes effective use of the limited monitoring data that are available. However, the use of measured correlations between personal exposure and ambient concentrations is not an appropriate means of estimating uncertainty in the exposure/concentration ratios used in NATA. Although the NATA deliberatively excluded exposures due to indoor sources and personal activities, these sources strongly influence and may even dominate measured exposures for certain chemicals. Moreover, the use of observed exposure/concentration ratios for fine particulate matter (PM) and ozone to gain insight into the exposure/concentration ratios expected for the air toxics addressed in NATA is inappropriate, since fine PM and ozone are not good surrogates for most of these compounds. In particular, the daily and seasonal time scales, and spatial distributions of fine PM and ozone are likely to differ significantly from those for air toxic compounds which are present predominantly as primary pollutants, and these spatial and temporal patterns can have a significant impact on personal exposure. Furthermore, the uniform distributions used in the illustrative calculations for PM and ozone exposure variability are completely arbitrary, and the uniform distribution used to represent uncertainty in the dose-response factors also appears to be arbitrary.

Since current data are not available to support development of probability distribution functions, a scenario-based approach for representing uncertainty should be used instead. Scenario analysis also has the advantage that it would emphasize data gaps and assumptions that might contribute to inaccuracies in the assessment. At this stage, highlighting possible inaccuracies is more important than

the focus on imprecision implied by the use of continuous probability distribution functions in Section 5.5. The approach proposed in section 5.5 may suggest that the estimated central tendency of a predicted quantity has a misleadingly high degree of reliability.

For each of the components of the NATA, summary tables should first be developed identifying alternative assumptions or data sources along with the amount of available versus missing data for the assessment. The "scenario" analysis would then combine high and low estimates of each factor, or estimates based on the major alternative sources of data or methods for calculation, rather than requiring distributions. For example, results straight out of ASPEN could provide the "low" value of metals concentrations, while the factor of five that reflects the model's underestimation compared to measurements could be incorporated to provide the "high" estimates. Similarly, in cases in which UREs are being or have been re-evaluated, risks calculated using previous versus current or proposed values could be compared to demonstrate the range of uncertainty in the estimates. An event, or "scenario tree" could be used to represent the adoption of each of the major conceptual or data-source assumptions in the combined assessment, and indicate the implications of each. The scenario tree would provide insight into which combinations of assumptions lead to the most important differences in predicted exposure and risk and consequently in prioritization of air toxics, and which assessment components warrant highest priority for further research or data collection.

An important use of the recommended scenario analysis is to guide the collection of new information to refine the study. For example, if the uncertainty associated with an estimated risk for a given compound is dominated by the uncertainty factors used in the derivation of the dose-response relations, investments in refined exposure modeling will not payoff proportionally in improving the risk estimate. Under such circumstances, there should be some mechanism for the NATA to communicate to the appropriate group (within, or outside of the Agency) the need for more accurate and precise dose-response information. At a minimum, the NATA process should clearly indicate which risk estimates are dominated by uncertainties in exposure estimates and which are determined by uncertain dose-response information as part of the risk characterization.

Recommendation. For the 1996 NATA, use the scenario-based approach described above to represent the uncertainty in the analysis, placing the emphasis on inaccuracies, rather than imprecision.

3.2.6.1 Specific Comments

The qualitative discussions of uncertainty sources given throughout the current report are valuable. However, the report should more carefully distinguish between sources of uncertainty that are specific to the NATA and sources of uncertainty that are common to all health risk characterization efforts. Where possible, greater delineation of major versus relatively minor sources of uncertainty would also be valuable.

Recommendation. For the 1996 NATA, differentiate between NATA-specific and universal sources of uncertainty, and between major and minor sources of uncertainty.

In Section 3.4.4, more consideration needs to be given to interpretation of the NATA results in view of the fact that the UREs and RfCs are thought to be "conservative" but the exposures are likely to be underestimated. The report generally implies that the assessment results are more likely to err on the side of overestimating risks than underestimating them. However, it is not clear that this is the case, since emissions and ambient concentrations appear to be underestimated, indoor sources are neglected, only median populations are considered, and dose-response estimates do not differentiate between healthy adults, children and other sensitive populations.

Recommendation. Use the scenario analysis to help bound the NATA risk estimates and avoid oversimplified characterization of the "nominal" results as conservative.

Section 4.2.2 of the NATA document should clarify the uncertainties associated with the various aspects of the emissions inventory to create more transparency about potential over and under estimations for each source sector. Tables 4-3 and 4-5 provide a good overview of the uncertainty

associated with the major point source inventory. However, it is difficult to draw clear inferences from comparisons of some of the emission estimates, since these comparisons mix differences due to methodology, time period and the set of sources that are addressed. Moreover, the uncertainty associated with area source, on-road mobile source and non-road mobile sources needs to be presented in greater detail in the current version of NATA.

A table should be included which provides the reader with an estimate of the confidence (high, medium or low) for each EPA-generated emission factor and the activity data used to generate the NTI for all non-point stationary sources (area sources). This is extremely important since these factors account for 70% of all of the non-point emissions. The Agency should make an effort to the make the non-point emissions inventory more transparent in the main document. Readers should not have to probe through layer upon layer of references in order to understand how this part of the NTI was developed. These same transparency concerns exist for the on-road and off-road mobile source emissions inventory. In order to improve future NATA assessments and spur future research, some degree of confidence needs to be included in the current NATA assessment for each individual component of the NTI. We recommend that the limitations of the NTI at least be ranked in order of importance for each general source sector (e.g. major, area/other, on-road mobile, and non-road mobile).

Recommendation. Provide more detail in the main NATA documentation on uncertainties associated with emissions from area, on-road mobile and non-road mobile sources.

In Section 4.3.4.2, characterizing the difference in results obtained using 1990 versus 1996 meteorological data as uncertainty is misleading. The differences reflect both uncertainty and variability.

Recommendation. Distinguish between reducible uncertainty (due to lack of information) and irreducible variability.

In Section 5.5.7, the discussion of uncertainties in risks aggregated across pollutants rests on the unlikely assumption that the uncertainties associated with each pollutant are independent. Some discussion should be added of how uncertainties in aggregate risks might behave if the assessment uncertainties are correlated across pollutants, as is likely in some cases. For example, uncertainties in motor vehicle activity factors simultaneously affect benzene, 1,3-butadiene and other air toxics associated with this source.

Recommendation. If uncertainty estimates are to be extended to aggregate risks, careful consideration needs to be given to which sources of uncertainty act independently across pollutants versus those uncertainties that simultaneously affect multiple pollutants.

A major output of the NATA may involve lists of counties estimated to be among the top X (or top Y%) of counties in terms of computed exposure and risk for all compounds, or selected HAPs. Should such lists be developed as part of NATA, it will be very important to identify the sensitivity of the results to differences in assumptions, using the scenario tree approach described above. Readers should be able to identify the specific reasons why a county is included in any list, for example, due to high estimated emissions of a particular type (facility, area, mobile on-road or off-road) for particular sets of compounds; low ambient dilution and dispersion (due either to local meteorology or the presence of small census tracks with high emissions); or specific demographic or time-activity factors. The presentation should also indicate the plausible scenarios under which the county is *not* included in the list.

Recommendation. Should lists of high-exposure/high-risk counties be developed as part of the NATA results, information should be provided on the key factors that determine whether or not a county is included on the list, and the sensitivity of the list to alternative scenarios considered in the scenario-tree evaluations.

3.2.7 Charge Question 7

Have the results of the assessment been appropriately and clearly presented? Can you

suggest alternative methods or formats that could improve the presentation and communication of these results?

The NATA assessment is complex and presents a challenge for compilation into a single document that flows well and leads the reader through the processes that are used. The current document is intended for use by technical experts. It will be critical to develop the summary documents to accurately communicate with non-technical audiences. The WEB page is apt to be the primary tool for communicating with such non-technical readers.

The draft is organized logically along the risk assessment paradigm and transparently takes the reader through the steps of the assessment. The steps are clearly described as well as the results. However, the detail necessary to make the assessment fully transparent also makes the document very long. It would be most useful if there were an executive summary that would summarize the key findings and conclusions. The introduction clearly describes the goals of the assessment and could form the outline for an executive summary. These distilled conclusions could then become the answers for a "Frequently asked questions" section on the public Web page. The assessment document and appendices do address each of the stated goals of the NATA study, but often it is difficult to find them. Thus an executive summary could for example, include statements such as in 6.3.1 which succinctly addresses Goal 1 - Identifying air toxics of greatest potential concern. If the readers can start with the core of the results, they will then have the context to critically follow the supporting materials to see that the results are appropriate.

The limitations at each step are clearly described and, if anything, are too comprehensive, giving the reader the impression that there is little confidence in the results. In some instances there is considerable confidence and others the model results are more speculative. While all the caveats are important for transparency, it would also be helpful in the beginning to have the authors describe the top 5 or 6 limitations that they believe have the greatest impact on the results and conclusions. In some of the chapters this is done very nicely and a qualitative as well as quantitative description is provided. If the limitations are agent specific, then that also needs to be described as is done with diesel particulate.

The maps and graphical displays of results are very helpful and compactly present the complexity of the project components and results.

The Web page will likely be the prime method for communicating with the general public. The current page is a good start for distilling the assessment down to manageable materials without losing critical information. This will be a critical communication tool to reach the majority of the public. Again the key will be to choose and display those aspects and results that the Agency finds most important and in which it has the greatest degree of confidence.

A challenge presented by the complexity of the report is to find a means to clearly communicate to the lay public which pieces of the assessment are understood and characterized with a relatively high degree of confidence, and which require further data gathering and model improvement before reliable estimates can be assured. Given the importance of environmental pollution information such as this (e.g., the widespread use of the TRI and the current NTI data by business, environmental groups and citizens), we recommend that the Agency, especially in materials intended for non-technical individuals, clearly distinguish between those parts of the NATA that are well established, vs. those which are in an earlier, developmental stage. In developing the web page for communicating results, the EPA should consider use of a hierarchical set of pages to differentiate between:

a) Information that is based solely on data or data reports, e.g., emissions datasets and ambient concentration and personal monitoring datasets for different compounds in different locations;

b) Information that is based on relatively simple or highly confident model calculations, such as ambient air concentration values computed by ASPEN for well-characterized air toxics that are not affected by secondary pollutant formation processes, in areas (terrain and meteorology) where ASPEN can provide reliable prediction, or total exposures to ambient pollutants computed assuming a simple indoor-outdoor penetrating factor; and

c) Information based on new model developments, where research is ongoing to improved the basis for prediction.

These pages could be color coded and titled to indicate: a) existing NATA data (using, for example, a blue background); b) existing NATA models (pale green background); and c) models undergoing research and development (yellow for caution).

For the lay public it will be important to place the consequences of exposure into a public heath context. A "thermometer" type graph could be used to display the levels at which different effects are seen, or to present different cancer risk levels. Examples of the types of displays that might be used can be seen in the Agency for Toxic Substances and Disease Registry (ATSDR) toxicological profiles as well as in materials developed by the State of New York. See, for example,

http://www.health.state.ny.us/nysdoh/environ/btsa.htm and

http://www.health.state.ny.us/nysdoh/environ/btsa/figure1.pdf

The public will be very interested to learn which counties in the United States rank highest for exposures and cumulative risks. In earlier sections of this report we have identified the significant uncertainty that we believe to be present in the quantitative scores derived for each county and that such rankings pose significant concern, given the limitations in the data used. However, despite any recommendations and cautions to avoid comparative ranking, the data in the report will allow others to do such comparisons if EPA does not provide such descriptive summary information.

The Panel is divided concerning the wisdom of presenting results of any type that identify specific counties as "hot-spot", high-exposure/ high-risk locations. Some members of the Panel believe strongly that states, citizens and other stakeholders will greatly benefit from this information and that, since other organizations will be able to access and manipulate the NATA results to produce it, it is

better to have the Agency perform this service. Others feel just as strongly that the uncertainty in NATA estimates is too great to justify identification of specific "hot-spot", high-risk counties, and that even if others could generate such a list, this was preferable to the EPA itself producing it (with the implied "official support" that this would entail). We note this disagreement within the Panel and hope that we have clarified the advantages and disadvantages to the Agency of producing a list of counties with high estimated NATA exposures and risks.

Should the Agency elect to produce a list of high exposure/high risk counties as part of the NATA, we recommend that the Agency do this by developing a qualitative ranking with perhaps an alphabetic listing in a table of the counties that score in the top Y (e.g., 1 to 5)% of exposure and risk, along with an indication of each variable that contributes to this high ranking (emissions by source type, local meteorological conditions, demographic or time-activity factors, or particular compound classes or toxicity assumptions associated with those compounds). Across the table could be listed the factors that contribute to the ranking and an "X" could be placed in the table when a listed county is in the top percentage group for that variable. This would allow the reader to identify which counties were in the top group as a result of the key contributing factor(s), rather than just their presence on the list as a result of the final, aggregated estimate of risk. While comparative ranking between individual counties within the top grouping (i.e. which is #1) would be highly problematic, it is likely that there is sufficient stability in the predictions to indicate that those in the top grouping as a result of factors known with a relatively high degree of confidence do deserve closer scrutiny.

Recommendation NATA 1996: It would be most useful if there were an executive summary that would summarize the key findings and conclusions.

Recommendation NATA 1996: At the start of each section, it would be helpful to have the authors describe the top 5 or 6 limitations that they believe have the greatest impact on the results/conclusions.

Recommendation NATA 1996: The Agency, especially in materials intended for non-

technical individuals, should clearly distinguish between those parts of NATA that are well established, vs. those which are in an earlier, developmental stage.

Recommendation NATA 1996: For the lay public it will be important to place the consequences of exposure into public heath context. A graphic representation such as a "thermometer" type graph could be used to display the levels at which different health effects are seen, or to present different cancer risk levels.

Recommendation NATA 1996 and NATA 1999: We recommend that the Agency consider developing a qualitative ranking with perhaps an alphabetic listing in a table of the counties that score in the top grouping in terms of exposure and risk, but that this table be accompanied by an indication of the factors that contribute to each county being among the high exposure/ high risk grouping, and the degree of confidence that can be placed in these factors.

3.2.8 Charge Question 8

The exposure methodology in NATA is being considered as one candidate for providing the basis for a national scale benefits analysis (as required in section 812 CAA). Please comment on the strengths and weaknesses of this approach, recognizing the limitations outlined in the NATA report?

Section 812 of the Clean Air Act Amendments of 1990 requires the EPA to periodically assess the effects of the Act on the public heath, environment and the economy. These assessments seek to compare benefits (e.g., health expressed in various monetary terms) and costs (e.g., costs of emission management options). Air toxics represent one aspect of the assessment that has not yet been quantified. The NATA exposure methodology is being considered as one viable approach to quantifying the relationships between emissions, concentrations, exposures and risks. In the 812 studies, the risks are then translated into monetary values to be compared to emission management

option costs.

Given the needs of the 812 study for an approach that can provide a sound basis for estimating benefits, the Panel must conclude at this point that the current exposure methodology and results in NATA are not yet ready for use in a national scale benefits analysis. This review has already noted the limitations of the models and data bases being used in NATA. Use of the current approach in the 812 studies would be subject to the same critiques.

Once the needed improvements noted above are implemented, application to benefits assessment can be considered. The particular improvements that have been listed as essential deal with the shortcomings of the models and the fact that a meaningful benefits assessment must consider the full distribution of exposure and risk (not just median values). It should also address acute health effects. Once exposure predictions are improved as noted and then validated, the cost-effectiveness of alternative toxics management strategies (for emissions and exposure reductions) could be compared, stopping short of a full benefits assessment, A full benefits assessment would need to consider health risks, mortality and morbidity avoided.

In our response to question 2 regarding the applicability of HAPEM, we recommended that a full distribution analysis of exposures to a HAP that has adequate data available across the US be conducted. One candidate HAP is benzene since adequate information is available for benzene to be able to do the analysis. If this recommended analysis is conducted, then it would be possible to conduct an initial benefits assessment for that HAP, to illustrate the type of analysis that is envisioned for a broader assessment of multiple toxics in the future.

Recommendation for the 1996 NATA. Results from the proposed assessment, for an information-rich HAP such as benzene, would be appropriate for the 812 study and should be considered. Descriptions of the limitations of the NATA for the 812 national benefits assessment need to be clearly articulated in both the NATA and the 812 studies.

3.2.9 Charge Question 9

Do you have suggestions for research priorities that would improve such air toxics assessments in the future?

An extensive research effort should be mounted to address the wide array of the data and model development needs to significantly improve the scientific foundation for future NATA studies as well as regulations based on the health risks of air toxics. The needs include both fundamental and chemical-specific research and span the whole of the risk paradigm (i.e., emissions, ambient concentrations, exposures, effects, and risks). The NATA document (pp. 126-127) does a good job of outlining the variety of research needs. Because air toxics research has been under-funded by the Agency for so long, considerable new resources are needed to address these needs. Fortunately, the NATA allows identification of the uncertainties that are inhibiting the development of reliable quantitative assessments, so that new resources could be well-focused. Prioritization is always difficult when there are so many needs, but perhaps this effort could be assisted by some sensitivity analyses based on the NATA.

Using the information developed in research programs is just as important as generating the information. Thus, no air research program can be useful until it is incorporated in Agency models for assessments. In the case of new research on health effects and dose-response factors, such information must be entered into IRIS. In numerous sections of this document, the importance of having an up-to-date, current IRIS database has been discussed. Support of IRIS also needs appropriate resources.

We understand that the EPA ORD is completing a research strategy for air toxics, so there in no need for SAB to duplicate this effort. We recommend that this plan be developed in concert with external experts on the related topics and that the subsequent draft be reviewed by this or a similar Panel. The Health Effects Institute is also preparing an Air Toxics Strategy, so ORD might also derive benefit from their activity. In addition, research needs on diesel particulate matter can be gleaned from the recent diesel assessment (Health Assessment Document for Diesel Exhaust, EPA/600/8-90/057E,

July 2000, SAB Review Draft). All of this must happen rapidly if new research is to be completed in time to impact the next NATA (and imminent air toxics regulatory assessments).

The issue of near-term and long-term research needs to be explicitly addressed. It will likely take EPA some time to complete the Air Toxics Research Strategy, and then implementation will require lead times consistent with future budget development. In the meantime, the knowledge base and dose-response assessment base for the 1999 NATA must be improved. In Appendix B we describe specific areas of focus that the Panel has identified as important for such a research effort. A more rigorous delineation of the Agency's research plan, for air toxics in general and NATA in particular, should be made considering this and other inputs and information, and subject to SAB review.

Recommendation: EPA should rapidly develop a research plan to identify the work (information collection, research, and assessments) it will perform with existing resources over the next few years that will directly improve the 1999 NATA. It should also proceed to complete the Air Toxics Research Strategy and have it reviewed by the SAB.

3.3 Summary of Recommendations

(I suggest that we simply list the recommendations (in bold italics) that are spread throughout Section 3.2, in the order that they appear, first for the 1996 NATA, then for future NATA's - - - MS)

The following recommendations are provided by the panel for the 1996 NATA:

Reactivity categories and decay rates should be identified for each HAP modeled in ASPEN. Critical assumptions and uncertainties associated with the assignment of reactivity classifications for HAPs should be discussed.

Etc., etc. . . .

The following recommendations are provided by the panel for the 1999 and future NATAs:

Implement additional QA/QC measures to ensure that a satisfactory level of nationwide completeness and accuracy is achieved for the point and area source emission inventories.

Etc., etc. . . . .

4. REFERENCES

[Please keep in mind that your text and statements need supporting reference materials cited in the report. I have listed the reports referred in the text or provided at the NATA March review meeting, or provided in subsequent editing sessions . We need your help here! - KJK]

Bell RW and JC Hipfner. 1997. Airborne Hexvalent Chromium in Southwestern Ontario. J Air Waste Management Assoc. 47: (8) p905 - 910.

California Air Resources Board (CARB) 1992. Technical Support Document: Proposed Identification of 1,3 - Butudiene as a Toxic Air Contaminant, Part A Exposure Asessment. State of California Air Resources Board.

Cumulative Exposure Project

Grohse PM, WF Gutknecht, L Hodson and BM Wilson. 1988. The Fate of Hexavalent Chromium in the Atmosphere. CARB Contract No. A6-096-32. Research Triangle Institute. Report No. RTI/3798/00-01F.

Harley, RA and GR Cass. 1994. Modeling the Concentrations of Gas-Phase Toxic Organic Air Pollutants: Direct Emissions and Atmospheric Formation. Environ. Sci.Technol. 28: 88 - 98.

National Research Council. 1983. Risk Assessment in the federal government. Managing the process. National Academy Press, Washington, DC

http://www.health.state.ny.us/nysdoh/environ/btsa.htm

Residual Risk Report to Congress

Northeast States for Coordinated Air Use Management (NESCAUM). 1999. NESCAUM Peer Review of USEPA's Draft Estimation of Motor Vehicle Toxic Emissions and Exposure in Selected Urban Areas. Boston, MA. (http://www.nescaum.org/committees/MSPeerReview/MSPeer.shtml)

Scott PK, BL Finley, MA Harris and DE Rabbe. 1997. Background Air Concentrations of Cr (VI) in Hudson County, New Jersey: Implications for Setting Health-Based Standards for Cr(VI) in Soil. J Air Waste Management Assoc. 47:(5) p592-600.

U.S. EPA. 1997. Cumulative Risk Assessment Guidance - Phase I: Planning and Scoping Memo from Administrator Browner, July 3, 1997. Available on-line at http://www.epa.gov/ordntrnt/ORD/spc/prhandbk.pdf)

U.S. EPA. 1998. Draft Integrated Urban Air Toxics Strategy to Comply with Section 112(k), 112(c)(3) and Section 202(1) of the Clean Air Act. 63FR 49240

U.S. EPA. 1999. Integrated Urban Air Toxics Strategy (as documented in 64 FR 38705 and available on-line at http://www.epa.gov/ttn/uatw/urban/urbanpg.html)

U.S. EPA/OSP/ORD. 2000. Science Policy Council Handbook on Risk Characterization., EPA 100-B-00-002, Office of Science Policy, Office of Research and Development, Washington, DC, December 2000 (Available on-line at http://www.epa.gov/ordntrnt/ORG/spc/prhandbk.pdf)

U.S. EPA/SAB. 2000. Review of EPA's Health Assessment Document for Diesel Exhaust, Science Advisory Board., Clean Air Scientific Advisory Committee, EPA 600/8-90/057E, December, 2000

U.S. EPA/SAB. 1989. Resolution on Use of Mathematical Modeling for Regulatory Assessment and Decision-Making, Environmental Engineering Committee/Modeling Resolution Subcommittee, EPA-SAB0EEC-89-012, January 13, 1989

U.S. EPA/SAB. 1996. The Cumulative Exposure Project: Science Advisory Board's Review of the Office of Policy, Planning, and Evaluation's (OPPE) Cumulative Exposure Project (Phase 1), Integrated Human Exposure Committee, EPA-SAB-IHEC-ADV-96-004, September 30, 1996

United States Environmental Protection Agency (EPA). 1998a. Study of Hazardous Air Pollutant Emissions from Electric Utility Steam Generating Units - Final Report to Congress. Vol. 1 Office of Air Quality Standards and Planning. Research Triangle Park, North Carolina 43/R-98-004a.

U.S. EPAS/SAB. 1998. Review of the US EPA's Report to Congress on Residual Risk, Prepared by the Residual Risk Subcommittee of the SAB, EPA-SAB-EC-98-013, September 30, 1998

U.S. EPA/SAB. 1999a. Advisory of the Charter for the Council on Regulatory Environmental Modeling (CREM), Science Advisory Board Executive Committee, Environmental Models Subcommittee, EPA-SAB-EC-ADV-99-009, June 1, 1999

U.S. EPA/SAB. 1999b. An SAB Advisory on the "White Paper" on the Nature and Scope of Issues on Adoption of Model Use Acceptability Criteria, Science Advisory Board Executive Committee, Environmental Models Subcommittee, EPA-SAB-EC-ADV-99-011, July 30, 1999

U.S. EPA/SAB. 1999c. The Clean Air Act Amendments (CAAA) Section 812 Prospective Study of Costs and Benefits (1999): Advisory by the Air Quality Models Subcommittee on Modeling and Emissions, Advisory Council on Clean Air Compliance Analysis, Air Quality Models

Subcommittee, EPA-SAB-COUNCIL-ADV-99-013, August 12, 1999

U.S. EPA/SAB, 1999d. The Clean Air Act Amendments (CAAA) Section 812 Prospective Study of Costs and Benefits (1999): Advisory by the Health and Ecological Effects Subcommittee on Initial Assessments of Health and Ecological Effects; Part 2, Advisory Council on Clean Air Compliance Analysis, Health and Ecological Effects Subcommittee, EPA-SAB-COUNCIL-ADV-00-001, October 29, 1999

U.S. EPA/SAB. 1999e. The Clean Air Act Amendments (CAAA) Section 812 Prospective Study of Costs and Benefits (1999): Advisory by the Council on Clean Air Compliance Analysis: Costs and Benefits of the CAAA, Advisory
Council on Clean Air Compliance Analysis, EPA-SAB-COUNCIL-ADV-00-002, October 29, 1999

U.S. EPA/SAB. 1999f. Final Advisory by the Advisory Council on Clean Air Compliance Analysis on the 1999 Prospective Study of Costs and Benefits (1999) of Implementation of the Clean Air Act Amendments (CAAA), Advisory Council on Clean Air Compliance Analysis, EPA-SAB-COUNCIL-ADV-00-003, November 19, 1999

U.S. EPA/SAB. 2000. Review of Draft Air Toxics Monitoring Strategy Concept Paper, Review by the Science Advisory Board Executive Committee, Air Toxics Monitoring Subcommittee, EPA-SAB-EC-00-015, August 18, 2000

U.S. EPA. 2001. "National-Scale Air Toxics Assessment for 1996," EPA-453/R-01-003, Office of Air Quality Planning and Standards, Research Triangle Park, NC, January 2001

The website address for charge 7 NYSDOH Toxicity trees is as follows:
http://www.health.state.ny.us/nysdoh/environ/btsa.htm

APPENDIX A - A MORE DETAILED DESCRIPTION OF THE SAB PROCESS

The SAB Staff recruited Dr. Mitchell Small, Chair of the Executive Committee's Environmental Models Subcommittee and Professor in the Departments of Civil & Environmental Engineering and Engineering & Public Policy at Carnegie Mellon University, to serve as Chair of the Subcommittee. Working with the Chair, other SAB Members and Consultants, Agency Staff, and suggestions from the public, the SAB Staff compiled a list of over 50 scientists and engineers ("Wide Cast") whose expertise appeared to be relevant to answering the questions in the Charge. Subsequently, the Chair, the Staff Director and the DFO reviewed the list in some detail and identified 22 individuals (Narrow Cast") to contact regarding their interest and availability to participate on Panel. Based on this information and the importance of having a balanced range of views on the technical issues represented on the Panel, the Chair and the DFO made recommendations for membership to the Staff Director, who made the final decision on the composition of the Panel. This process included assigning different members Lead and Associate responsibilities for each of the Charge questions.

The Agency transmitted review materials to the Subcommittee members in late January, 2001. On February 21 the SAB Staff convened a publicly-accessible, Federal Register-noticed conference call meeting between Panel members and Agency staff. The goal of this information-gathering meeting was to clarify any questions that Panel Members might have, to identify any gaps in the information sent to the Panel, and to identify areas that the Agency should be prepared to clarify at the face-to-face meeting. Minutes of the meeting were posted on the SAB Website: www.epa.gov/sab.

In addition, public comments were received and distributed to the Panel Members at the February 21, 2001 informational conference call meeting from many of the groups that attended and spoke at the March 20 & 21, 2001 meeting. On March 20-21, 2001 the Panel convened in the ballroom of the Raddison Governor’s Inn Hotel, Research Triangle Park, NC. Those groups

providing formal written public comments are listed below. All parties spoke during the public comments session on March 20th, except for the latter two groups, which transmitted written public comments without attending the meeting. The groups and presenters are listed as follows:

a . The Acrylonitrile Group, Mr. Chuck Elkins,

b. The Residual Risk Coalition, Mr. Chuck Elkins,

c. The Colorado Air Pollution Control Division, Ms. Lisa J. Silva,

d. The Ethylene Oxide Council, Dr Jane Teta,

e. The Engine Manufacturers Association, Mr. Timothy French

f. The Halogenated Solvents Industry Alliance, mr. Stephen P. Risotto,

g. The Hydrazene Panel of the American Chemistry Council, Ms. Claudia O'Brien of Latham and Watkins,

h. The International Truck and Engine Corporation, Ms. Claudia O'Brien of Latham & Watkins,

i. Dr. Robert J. Carton, Chief of Environmental Protection, U.S. Army Medical Research & Materiel Command, Fort Dietrick, MD (written comments submitted, but not in attendance at meeting), and

j. Dr. Amy D. Kyle, Univ of Calif, Berkeley, CA (written comments submitted, but not in attendance at meeting),,

During the March 20 & 21, 2001 public meeting, the NATA Review Panel heard presentations from the Agency staff on the first day, as well as public comments. This was followed by detailed discussion by the NATA Panelists on the nine charge questions. The second day saw the discussion being completed by the NATA Panel on the Charge questions in the morning, followed by preparation for a poster session on key points within each charge question, as well as re-writing of the pre-meeting written comments by the NATA Panelists to their assigned charge questions, and teaming in groups by the NATA Panelists to develop merged language edits.

By the end of the first day, the individual comments and merged edits were incorporated into a template for a first draft, which was given to the Chair to synthesize into a second draft. Dr. Small

emailed the second draft to the NATA Panel on April 6th. There was a contingency provision in announced in the Federal Register Vol. 66, No. 29, February 12, 2001, pages 9846-9847, to hold a public conference call on April 24th, should it be needed. The NATA Panel decided to exercise this option, and planned to conduct a technical editing public conference call in which the public can follow the NATA Review Panel's discussions on their working draft, which is not yet a public consensus report. The NATA Panel anticipated that a public consensus draft would be completed around May 1st, and planned to hold a public conference call to reach closure on edits to that draft report on May 14th in order to give the NATA Panelists and the public adequate reading on the draft report. The draft took longer to develop, and a technical editing work session was also held also on May 25th. This "working" public draft was developed on June 6th and posted onto the SAB website on June 7th for discussions on June 13th.

NOTE: We have posted notices, agendas, and the publically-available draft reports onto the SAB website (www.epa.gov/sab), along with related efforts to reach out to all potentially affected and interested parties. This may also include discussion of the utility of a conference call meeting one month prior to the March meeting to discuss and negotiate the charge, determine if the review materials are adequate, and begin the pre-meeting review and writing process. We may also touch on the use of a URL site for all Agency review materials, appendices, background briefings, etc. - K. Jack Kooyoomjian 6/6/01)

APPENDIX B - AREAS OF FOCUS IDENTIFIED BY PANEL MEMBERS FOR RESEARCH TO IMPROVE FUTURE NATA STUDIES

The NATA Review Panel recognizes that evaluation of the NATA national-scale results is an iterative process and wishes to go on record in support of the research priorities already recognized by the Agency. In particular, the Agency's Office of Research and Development (ORD) has drafted an air toxics research strategy to assist the Agency's NATA effort in understanding the risk and assessment tools, as well as the proper uses and limitations these tools for future assessments. We support efforts to engage in the following activities which are discussed in the 1996 NATA document on pages 126-127:

a) Improve the quality of emission data,
b) Improve the support for urban-scale modeling,
c) Improve the characterization of background,
d) Provide support for future model-to-monitor comparisons for ambient air toxics concentrations,
e) Provide support for future model-to-monitor comparisons for exposure,
f) Improve dose-response information,
g) Extend EPA risk assessment guidelines to be more inclusive of children and other vulnerable subpopulations, and
h) Improve modeling to include multipathway exposures.

The Agency's list appropriately identifies important limitations of the current assessment that need to be addressed. In particular, the Panel underscores the importance of the acknowledged needs to improve the quality of the emissions data, improve the ambient modeling capability, obtain better data on background concentrations, and incorporate exposure pathways other than inhalation for persistent and/or bioaccumulating compounds.

As the research strategy is implemented, it will compete for other needs. We encourage the Agency to refine the research strategy to address fundamental scientific needs, specific data information and modeling needs, and chemical-specific needs. The following text offers additional thoughts in these areas.

A) **Fundamental Scientific Needs:** The discussion that follows focuses principally on the research needed to provide the capability to reliably estimate uncertainty and variability for population distributions of risk to the general populace and susceptible populations.

1) Improved multimedia, multipathway, multipollutant transport, fate, and transformation models that have been scientifically evaluated (e.g., validated) that estimate the relationship between sources and environmental levels,

2) Improved of multimedia, multipathway, multipollutant exposure and dose models (that have been scientifically evaluated/validated) to relate environmental concentrations to the population distribution of actual human exposure and dose,

3) Improved and harmonized cancer and noncancer assessment methods that can be applied to air toxics as multimedia, multipathway chemicals,

4) Improved methods to estimate distributions of cumulative risk,

5) A strategy is needed for health assessments of mixtures of particles and gases,

6) Fundamental improvements in modeling of secondary pollutant formation are needed in future NATAs. The NATA report recognizes the need for improvements to the ASPEN framework for modeling ambient, concentrations. However, as discussed in response to charge question 2, the ASPEN model is fundamentally unsuited for secondary air toxics and pollutants for which long-range transport is important. Development and application of MODELS3 or a similar model of atmospheric chemistry and transport over local to national

scales is required in order for future assessments to adequately treat secondary air toxics and compounds with significant long-range transport contributions, and

7) Improving background concentrations for air toxics: The Panel agrees with the Agency that improving the characterization of background concentrations for air toxics so that they can be treated as region and season-specific is also an important priority.

B) **Specific Data Information and Modeling Needs:**

1) Longitudnal activity patterns for different cohorts is necessary: At present, only daily-time activity information has been used in the NATA. Some improvement in longitudnal activity patterns for different cohorts is necessary. In future assessments, the implementation of the HAPEM model needs to be improved to adequately reflect the full range of interindividual variability in air toxics exposures. One research need for doing this correctly is to investigate and incorporate longitudnal activity pattern data for different cohorts. For the 1996 NATA report, the Panel recommends that the Agency attempt to illustrate the required approach by applying it to benzene, for which relatively extensive exposure data are already available. Although the panel recognizes that NATA is deliberately focused on the risks associated with "outdoor" air toxics , we recommend that the Agency develop the capability to account for indoor sources in the assessment. This is needed to be able to evaluate the HAPEM model, since exposure measurements naturally include exposure to air toxics from indoor as well as outdoor sources. It is also needed in order to adequately characterize the risk associated with air toxics for which both indoor and outdoor sources are important,

2) Improve the current "zero" value used for the ADD factor in HAPEM and

develop as necessary in a research approach for air toxics,

3) Review TEAM and NEXHAS data to determine relevance for incorporation to improve HAPEM,

4) Updating of IRIS Database: Updating the IRIS database is an important information need and support of IRIS also needs appropriate resources. Using the information developed in research programs is just as important as generating the information. Thus, no air research program can be useful until it is incorporated in Agency models for assessments and until the dose-response assessment is entered into IRIS. In numerous sections of this document, the importance of having an up-to-date, current IRIS database has been discussed,

5) Use of Geographic Information System (GIS) tools for displaying and communicating emissions estimates: The Agency should focus on developing improved methods for direct cross-validation of emission estimates. This might include use of Geographic Information System (GIS) tools for displaying and communicating emissions estimates to state and local agencies and stakeholder groups that are well-positioned to ground-truth the data,

6) Improve Estimates for Non-Road Mobile Source Emissions: Non-road mobile source emissions appear to be major contributors to risks associated with toxic air pollutants. However emissions models and inventory development methods for non-road mobile sources are not as well developed as those for on-raod vehicles. The efforts to improve methods for estimating emissions from non-road mobile sources that are underway at the Agency deserve priority, and should be followed closely by staff working on NATA, and

7) Seek information from other institutions as appropriate: The Health Effects Institute (HEI) is also preparing an Air Toxics Strategy, so EPA ORD might also derive benefit from their activity. However, all of this must happen rapidly if new research is to be completed in time to impact the next NATA (and imminent air toxics regulatory assessments),

C) **Chemical-Specific Needs:** Among the chemical specific information needs are the following:

1) Improved emissions inventories to obtain better environmental, exposure, and dose measurements to enable development, evaluation, and verification of models,

2) Dose-response and mechanistic studies are needed targeted to the specific uncertainties that drive the risk for the chemicals of higher concern, and

3) GIS tools to facilitate diagnostic study of relationships between economic activity for industrial sectors and the emissions estimated for those sectors.

APPENDIX C – GLOSSARY

ADD	Additive Factor (Used in the exposure model HAPEM4 to account for the contribution from indoor sources to personal exposures)
AIRS-A	Aerometric Information Retrieval System (Data base)
ASPEN	Assessment System for Population Exposure Nationwide (dispersion model)
ATSDR	Agency for Toxic Substances and Disease Registry
CAA	Clean Air Act
CAAA	Clean Air Act Amendments
CAQPS	Office of Air Quality Planning and Standards (U.S. EPA/OAQPS)
CASAC	Clean Air Scientific Advisory Committee (of the U.S. EPA/SAB)
CEP	Cumulative Exposure Project
CHAD	Consolidated Human Activity Database (an EPA database for 40 cohort groups)
CMAC	Community Multi-scale Air Concentration (? Any Suggestions here?-KJK)
CMAQ	Community Multi-scale Air Quality (model)
CO	Carbon Monoxide
Cr	Chromium and Isotopes (e.g., Cr+3 - Trivalent and Cr+6 - Hexavalent Chromium)
DEP	Diesel Exhaust Particulates
EMS	Emissions Modeling System
EPA	U.S. Environmental Protection Agency (U.S. EPA)
GIS	Geographic Information System
HAD	Hazard Assessment Document
HAP	Hazardous Air Pollutant
HAPEM	Hazardous Air Pollutant Exposure Model
Hg	Mercury

3187	HQs	Hazard Quotients
3188	IRIS	Integrated Risk Information System (data base)
3189	ISC	Indusatrial Source Complex (model)
3190	IUATA	Integrated Urban Air Toxics Strategy
3191	LRT	Long Range Transport
3192	MACT	Maximum Achievable Control Technology
3193	MLEs	Maximum Likelihood Estimates
3194	MobTox	Mobile Toxic Emission Model (for mobile sources, e.g., MobTox5b)
3195	MODELS3	A Comprehensive Modeling Framework Currently Under Development by
3196		U.S. EPA/ORD
3197	MRL	Minimum Risk Level
3198	MS	Mobile Sources
3199	MSAT	Mobile Source Air Toxics
3200	NATA	National-Scale Air Toxics Assessment (also National Air Toxics Assessment)
3201	NCEA	National Center for Environmental Assessment (U.S. EPA/ORD/NCEA)
3202	NET	National Emission Trends
3203	NHEXHAS	National Human Exposure Assessment Project
3204	NLEV	National Low Emission Vehicle
3205	NTI	National Toxics Inventory
3206	NYC	New York City
3207	NYS	New York State
3208	NYSDEC	New York State Department of Environmental Conservation
3209	O3	Ozone
3210	OAQPS	Office of Air Quality Planning and Standards (U.S. EPA/OAR/OAQPS)
3211	OAR	Office of Air and Radiation (U.S. EPA/OAR)
3212	ORD	Office of Research and Development (U.S. EPA/ORD)
3213	OTAQ	Office of Transportation and Air Quality (U.S. EPA/ORD)
3214	OZIP	OZone Isopleth Plotting model (for predicting ozone in urban areas)
3215	PAH	Polynuclear Aromatic Hydrocarbons (one type of POM)

PBTs	Persistent Bioaccumlative Toxins
PM	Particulate Matter
POM	Polycyclic Organic Matter
QA/QC	Quality Analysis and Quality Control
RELs	Reference Exposure Levels
RfCs	Reference Concentrations
RFG	Reformulated Gasoline
SAF	Spatial Allocation Factors
SIC	Standard Industrial Classification
TEAM	Total Exposure Assessment Methodology
TEF	Total Exposure Factor (?) Toxicity Equivalency Factor
TMDL	Total Maximum Daily Load
TOG	Total Organic Gasses
TOSHI	Target Organ-Specific Hazard Index
TRI	Toxics Release Inventory
UREs	Unit Risk Estimates
URF	Unit Risk Factor
U.S.	United States
VMT	Vehicle Miles Traveled
VOC	Volatile Organic Compounds

END OF TEXT