UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

June 20, 2007

EPA-SAB-07-009

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

Subject: Consultation on EPA's Risk and Technology Review (RTR) Assessment Plan

Dear Administrator Johnson:

The EPA's Office of Air Quality Planning and Standards (OAQPS) requested that the Science Advisory Board (SAB) conduct a consultation to provide input on whether its proposed assessment plan is adequate to furnish the basis for regulatory decisions concerning specific source categories. A consultation is conducted under the normal requirements of the Federal Advisory Committee Act (FACA), as amended (5 U.S.C., App.), which include advance notice of the public meeting in the *Federal Register*. Although there will be no consensus report from the SAB as a result of this consultation, the panel would nonetheless like to underscore several key points that arose in the conduct of its consultation on the proposed Agency plan.

On December 7, 2006, the panel met via telephone conference where representatives of the OAQPS offered informative presentations to the members of the SAB Risk and Technology Review Consultative Panel. The focus of the presentations by EPA representatives for this consultation was on the emission data, dispersion and exposure modeling, dose-response assessment and risk characterization to be utilized in the proposed Agency plan. On December 19, 2006, the panel met again via telephone conference to discuss and deliberate on the charge questions. A copy of the overview of the plan and the charge questions to the panel are attached to this letter (Attachment 1). In brief, the SAB was asked to comment on the appropriateness and adequacy of using a new approach to perform an assessment with the goal of characterizing the exposures and risks associated with the emissions of hazardous air pollutants (HAPs) from 51 different industrial source categories. Feedback on the charge questions was provided by panel members and a compilation of their comments and recommendations is attached (Attachment 2) to this letter and appended to the minutes for this meeting.

On behalf of the panel members, we would like to express our sincerest gratitude to the presenters for their expertise, perspectives and insights. Their contributions greatly increased our

understanding of the Agency's current policies, methods, practices and future directions proposed for residual risk assessment. The panel members had several concerns about the RTR II process that are expressed in their individual comments. Highlighted in this letter are several key messages that emerged among the panel members as a result of the Agency presentations and discussions:

- There is a concern that the expedited review (RTR II) is focused on the most feasible sources to review, i.e., the easiest assessments to make, rather than on the sources with the greatest potential residual risk;

- The plan needs to be revised to make it clearer, and the process more transparent, by addition of flow charts that indicate the differences between the various RTR processes (RTR I, RTR II, RTR III, etc.); The current version does not clearly explain the basis for greater efficiency in the RTR II over RTR I;

- The plan should incorporate a framework for improving the National Emissions Inventory (NEI) as new/more accurate data become available. There is a concern that the use of the voluntary NEI data base for a regulatory purpose could induce changes in reporting that modify the data base;

- The uncertainty and/or bias of the model estimates (HEM-AERMOD) should be addressed; models should be improved to better reflect the atmospheric transformations of emissions that affect potential health/environmental impacts; and field measurements should be compared to modeled estimates to help determine how well models predict actual conditions for HAPs;

- If feasible, a probabilistic analysis of the exposure scenario and other factors should be developed in order to: (1) provide more transparency regarding variability and uncertainty in the exposure and risk estimates, and (2) allow a more informed judgment regarding the extent of conservatism included in the calculation/model;

- A transparent decision framework should be developed that: (1) identifies HAPs where ecological risks rather than human health endpoints are the basis for setting air emission limits, and (2) addresses the importance of facility emissions relative to background sources for naturally occurring HAPs;

- The plan should clearly and repeatedly state in any communication to the public that the assessment should be used only for the purpose intended by the Agency and that the estimates should not be construed as absolute estimates of residual risk for use in population-based studies;

- The existing toxicity databases should be evaluated to determine if they are complete enough to be used in the RTR process;

- A sensitivity analysis should be conducted to determine: (1) which inputs are the main drivers of the Residual Risk estimates, and (2) if differences in the levels of uncertainty for those inputs (for example, uncertainties in emissions from some sources compared to others) may potentially result in misclassification.

Finally, the SAB commends the Agency on seeking early advice to enhance their Risk and Technology Review Assessment Plan. We look forward to working with the Agency as they implement this plan and refine their assessments.

Sincerely,

/Signed/

Rogene Henderson, PhD
Chair, Risk and Technology Review (RTR) Consultative Panel

/Signed/

Granger Morgan, PhD
Chair, Science Advisory Board

Enclosures:

ATTACHMENT 1: Overview and Charge Questions
ATTACHMENT 2: Panel Member Comments

ATTACHMENT 1

Office of Air and Radiation
Risk and Technology Review (RTR) Assessment Plan
Peer Review Charge

Overview

As a part of the technical basis for rulemaking in the EPA's Risk and Technology Review (RTR) effort, EPA seeks input on whether its proposed RTR risk assessment methodology (emission data; dispersion and exposure modeling; risk characterization) is adequate to provide the basis for regulatory decisions concerning specific source categories. In sum, we are using a new approach to perform an assessment with the goal of characterizing the exposures and risks associated with the emissions of hazardous air pollutants (HAP) from approximately 30 different industrial source categories. These categories have previously been subjected to national emission standards, and the purpose of characterizing their risks now is to determine whether those emission standards (which were based on emission control technologies, work practices and other control measures available at the time they were promulgated) are adequate to protect public health with an ample margin of safety and prevent adverse environmental effects. The Residual Risk section (section 112(f)(2)) of the Clean Air Act (CAA) requires EPA to make this determination through notice-and-comment rulemaking. If EPA determines that the previous standards do not reduce lifetime excess cancer risks to the individual most exposed to emissions of less than one in one million, EPA must issue a standard that protects public health with an ample margin of safety unless EPA determines a more stringent standard is necessary to prevent an adverse environmental effect. In determining if a more stringent standard is necessary, EPA will take into account costs, energy, safety and other relevant factors.

The overall plan for the new approach is as follows: (1) conduct a risk assessment using currently-available source and emissions data; (2) share the source and emissions data and the results of the assessment with the public through an Advanced Notice of Proposed Rule Making (ANPRM), asking for public comments on the methods and the source and emissions data; (3) receive comments; (4) reconcile comments and correct the source and emissions data as appropriate, and; (5) re-assess the risks. The results of the revised risk assessment will be used to support proposals and promulgations of technology- and risk-based regulatory decisions for each of the categories through the regular notice-and-comment rulemaking process.

Since the planned risk assessment represents a significant activity of an influential nature (*i.e.*, it will be used for regulatory purposes), and since it departs from previous risk assessments for the residual risk program in several important ways, we are seeking a scientific peer review of the assessment through the Agency's Science Advisory Board.

A background on the residual risk program and the regulatory decision framework associated with it is provided below to provide reviewers with a context for the assessment.

Background -- Regulatory Context and Decision Framework

Section 112(f)(2)(A) of the CAA requires EPA to evaluate whether previously adopted standards applicable to emissions of HAP from source categories under the technology-based (Maximum Achievable Control Technology, or MACT) program provide an ample margin of safety to protect public health and prevent adverse environmental effects, taking into consideration costs, energy, safety and other relevant factors. Any standards set under this section are to ensure that the public health is protected to a level which provides an "ample margin of safety" unless we determine that a more stringent standard is necessary after considering costs and other relevant factors. (We also recognize that emissions of persistent, bioaccumulative HAPs (PB-HAPs) may have an adverse environmental effect and need further evaluation beyond a human inhalation risk assessment. Source categories that emit PB-HAPs will be addressed in the RTR III assessment.) The specific language of section 112(f)(2)(A) directs EPA to set additional standards through rulemaking if we determine that the MACT standards for the regulated source category do "not reduce the lifetime excess cancer risks to the individual most exposed to emissions from a source in the category or subcategory to less than one in one million..."

Risk assessments performed in the residual risk program are designed to generate answers to the questions posed by section 112(f)(2). Thus, the initial goal is to determine if previous standards do not reduce lifetime excess cancer risks to the individual most exposed to HAP emissions from a source in the category or subcategory to less than 1 in 1 million. In this context, the "individual most exposed " is an individual who has been determined to "live" in a census block that currently shows at least one resident (EPA is using the 2000 census data). If a source presents lifetime excess cancer risk levels to this individual of at least 1 in 1 million, then EPA must conduct rulemaking to ensure protection of public health with an ample margin of safety.

While the process for determining what constitutes an "ample margin of safety" was not explicitly specified in the 1990 CAA, section 112(f)(2)(B) expressly preserves the interpretation of "ample margin of safety" as it was used in the pre-1990 version of section 112 and as reflected in the 1989 rulemaking promulgating National Emissions Standards for Hazardous Air Pollutants for sources of benzene (the Benzene NESHAP). (54 FR 38044, Sept. 14, 1989). In that rule, EPA explained that, "in protecting public health with an ample margin of safety ....., (we) strive to provide maximum feasible protection against cancer risks from hazardous air pollutants by (1) protecting the greatest number of persons possible to an individual lifetime risk level no higher than approximately 1 in 1 million and (2) limiting to no higher than approximately 1 in 10 thousand the estimated risk that a person living near a plant would have if they were exposed to the maximum pollutant concentrations for 70 years." (FR 38044, Sept. 14).

In the approach used in the Benzene NESHAP rulemaking, the first step of the two-step ample margin of safety framework is the determination of acceptability, i.e., the level of cancer risks which can be considered "acceptable" based on health considerations only (costs, technical feasibility and other non-health-related factors are not considered at this stage). The determination of what represents an acceptable risk level is to be made in the context of "the world in which we live," that is, recognizing that our world is not risk-free. In the Benzene NESHAP, EPA determined that the cancer risk to the individual most exposed to emissions from

the sources addressed in that rulemaking could not be considered acceptable unless it was at or below approximately 100 in a million (or 1 in 10,000). This determination established a "presumptive" acceptable level of 100 in a million cancer risk, thereby providing a benchmark for judging the acceptability of maximum individual risk for future risk-based emission standards, but not constituting a rigid line for making those judgments. FR 38045, Sept. 14).

The second step of the ample margin of safety framework is the determination of what level of standard actually provides an ample margin of safety to protect public health. The maximum level of excess lifetime cancer risk associated with this standard can be no higher than what EPA determines to be the "acceptable" level for the particular category or subcategory, and it can be as low as or lower than approximately 1 in a million, but the ample margin of safety must be determined by balancing the costs associated with further reducing emissions against the health risk reductions achieved. To inform this judgment, residual risk assessments are designed to provide multiple metrics of risk (e.g., maximum individual risk, distribution of risks across the exposed population, total cancer incidence, noncancer hazard indices), as well as an indication of the limitations and uncertainties associated with the assessment.

Finally, to fulfill EPA's obligation under section 112(d)(6), every eight years we must review and revise as necessary (taking into account developments in practices, processes and control technologies) emission standards adopted under section 112. As recently explained in the notice of proposed residual risk rulemaking for the Hazardous Organic NESHAP (HON), we view costs and risk as relevant factors in determining whether it is necessary to revise standards under section 112(d)(6). 71 (FR 34422, 34437; June 14, 2006).

Previous Relevant Peer Reviews

Several previous peer reviews have covered elements associated with this assessment or assessments with similar scopes or contexts. A peer review of this assessment is not intended to duplicate these previous efforts. A brief summary of each peer review is provided:

- The *Residual Risk Report to Congress*, a document describing the Agency's overall analytical and policy approach to setting residual risk standards, was issued to Congress in 1999, following a peer review by the Agency's Science Advisory Board. Many of the design features of the RTR assessment were described in this report, although individual elements have generally been improved over the techniques described in that document.

- Individual residual risk assessments – several internal peer reviews and one external peer review were conducted on risk assessments for individual source categories, including Coke Ovens, Perchloroethylene Dry Cleaning, and Halogenated Solvent Cleaners. Each of these assessments used emission estimates from the National Emission Inventory, human exposure modeling at the census block level, dose-response methodologies, and risk characterization which are similar to those for the planned RTR assessment.

- The National Air Toxics Assessment, or NATA, for 1996, was peer-reviewed by an SAB panel in 2002. While this assessment was a comprehensive and cumulative risk assessment (it was designed to include all mobile sources, small industrial sources, and large industrial sources, as well as background contributions of air toxics), because of the

large amount of associated uncertainty it was deemed to be not appropriate for regulatory purposes, and it did not carry a census block-level resolution (it was performed at the census tract level). For this reason, on EPA's NATA web site the estimated risks are characterized as "starting points" for developing refined assessments.

- AERMOD, a recently-developed source-to-receptor air quality dispersion model, was the subject of significant interagency cooperation and peer review. It is now EPA's preferred local-scale air dispersion model for industrial sources of air pollution.

- Toxicity assessments – the individual dose-response metrics used in the RTR assessment have themselves been the subject of peer reviews through the agencies who developed them (including EPA, through its Integrated Risk Information System, or IRIS; the California Environmental Protection Agency, the Agency for Toxic Substances and Disease Registry, and Environment Canada). EPA proposes to select dose-response values from these sources in the same priority order it used for NATA (i.e., IRIS, then ATSDR, then CA).

- It is not the intention of this consultation to duplicate or comment on these previous peer reviews, but rather to acknowledge that we are using the most recent scientifically-credible dose-response approaches (as determined through peer review) and that, since dose-response science is continually-evolving, it is a source of significant, usually unquantifiable uncertainty.

Charge Questions for Peer Consultation

The peer consultation on the plan for the assessment will focus on: (1) the development of the source data and emissions inventory for the 30 source categories; (2) the analytical approach for quantifying and characterizing human exposures and risks. In particular, we would like reviewers to consider the following questions as a means of focusing their consultation:

1. Scope: Is the scope of the assessment appropriate for the stated purpose? Is the overall approach clearly and adequately explained for review?

2. Emissions and Source Data: The National Emissions Inventory for hazardous air pollutants represents an ongoing voluntary national effort whose creation results from the collaborative efforts of State, local, and tribal air agencies with EPA Regional and Headquarters staff.
    a. Short of creating a federal mandate for reporting emissions to the EPA, do the methods by which the NEI was developed, reviewed, and compiled result in a technically-credible database that can support regulatory assessment and action? If not, can you suggest ways to improve it?
    b. Do the plans for conducting an engineering review and incorporating currently-available refined emissions and source data into the inventory add value to the assessment? Does the plan for soliciting public comment through an advanced notice of rulemaking add scientific credibility to the inventory? Is the plan for

reconciling comments on the inventory adequate? If not, can you suggest other approaches for reconciling such comments?

3. Identifying Source Categories with Significant Non-inhalation Risk Potential: This assessment is only designed to include source categories whose risks are dominated by the inhalation pathway. Are the methods planned for selecting source categories with potentially-significant ecological risks or multi-pathway human health risks for a separate, more refined ecological and multipathway assessment sufficiently health-protective? Are there ways that you might suggest for improving such screening techniques that can make them less conservative and still scientifically-defensible?

4. Dispersion Modeling: Does the coupling of the AERMOD dispersion model with the census block human exposure modeling (HEM) approach to estimating individual and population exposures represent a credible approach for this goal? Are there other more credible approaches available for the estimation of inhalation risks from the types of source categories being examined? Is the level of accuracy of this approach acceptable for the purposes of residual risk decision-making? Are there any specific source categories for which this approach might be considered inadequate?

5. Acute Exposure Screening: The plan describes a screening methodology for identifying potentially-significant acute exposures from routine emissions. Is this method appropriately protective? If the potential for acute exposures of concern is identified, is the plan for refining the assessments appropriate?

6. Exposure Assessment: Beyond the use of AERMOD-HEM, is the methodology planned for characterizing exposure commensurate with the needs of residual risk assessments? Specifically, does the underlying theory and data used to account for the effect of population migration on exposure make our lifetime population risk assessment more or less defensible than assuming that the exposed population lives in the same location for a lifetime of 70 years? Is omitting the attenuation of exposure concentrations associated with building penetration justifiable when estimating lifetime risks for these chemicals and these types of sources? Is omitting the impact of short-term human activity patterns on exposures acceptable for these purposes?

7. Dose-Response Values: Is the plan for using available dose-response information (e.g., sources of information, prioritization scheme) appropriate for the purposes of this assessment? If not, can you suggest ways to improve it?

8. Risk Characterization: What are the strengths and the weaknesses of the overall conceptual approach to risk characterization planned for this assessment? Does the characterization plan adequately cover sensitive subpopulations and early-life exposures? Does the risk characterization plan appropriately aggregate cancer risks? Does the risk characterization plan appropriately aggregate noncancer risks? What are the strengths and weaknesses of the planned approach for characterizing important uncertainties, variabilities, and limitations? Given the underlying science and the intended purposes of the assessment, 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?

9. Overall: Has any important scientific information been omitted from this assessment plan that could impact a subsequent regulatory decision? In your opinion, will the overall approach for the 30 source categories provide results that will be sufficient to support regulatory decision-making in the context of EPA's residual risk program?

ATTACHMENT 2

Comments from Dr. Timothy Buckley

1. Scope: Is the scope of the assessment appropriate for the stated purpose? Is the overall approach clearly and adequately explained for review?

The Risk and Technology Review (RTR) Assessment Plan (Draft 11/20/2006) is a well written report. The methods and plans are well organized and clearly described. The proposed risk assessment approach is appropriate for meeting the regulatory needs under 112(f)(2) for assessing source category residual risks.

7. Dose-Response Values: Is the plan for using available dose-response information (e.g., sources of information, prioritization scheme) appropriate for the purposes of this assessment? If not, can you suggest ways to improve it?

The tiered approach for selection of hazard ID and dose-response values for use in the RTR is appropriate and reasonable. The development of dose-response values is a complex and challenging process that requires distillation and synthesis of an always evolving primary literature in a systematic and transparent way. IRIS provides an appropriate 1st priority database because it contains a large number of relevant chemicals and is scientifically credible, in recent years including a process of peer review. The use of ATSDR and CalEPA dose response findings as secondary and tertiary sources appears reasonable in evolving from federal to state agencies. This section of the RTR Assessment Plan can be strengthened by more completely describing the differences between the three databases (e.g. peer review, # of chemicals, staff and resource allocation) that support priority selection. It would also be of interest to know what the overlap is across the three data sets and a general assessment of how values differ (e.g. is one more conservative than another).

The hazard identification and dose-response approach for selecting acute effects is not as clear cut in that not all HAPs have values listed and selection is individual chemical specific based on professional judgement. The criteria for selection including conceptual consistency and level of peer review are appropriate.

On pg 16 it is stated that dose-response values for chronic oral exposure were obtained from OAQPS rather than IRIS but this selection is not justified. Why wouldn't IRIS be the appropriate selection as is the case for inhalation exposure?

9. Overall: Has any important scientific information been omitted from this assessment plan that could impact a subsequent regulatory decision? In your opinion, will the overall approach for the 51 source categories provide results that will be sufficient to support regulatory decision-making in the context of EPA's residual risk program?

The report is comprehensive in its consideration of the science that underlies the RTR Assessment Plan.

The report's scientific credibility can be strengthened by including the extensive peer review literature that relates to NATA risk and ambient concentration estimates. Studies comparing/validating ambient concentration measurements to modeled estimates are particularly relevant and provide assurance of the model estimates.

The scope of the RTR II is stated to be based on practical considerations including the availability of emissions, complexity of the assessment, etc (pg 4). Although it is understandable that feasibility is considered, a stronger rationale for inclusion under RTRII would be an assessment of what source categories are believed to pose the greatest risk and therefore where will the greatest public health benefit be gained.

Section 1.2 RTR II Process could benefit from a flow chart that shows the process elements and how they relate to one another.

Comments on RTR-II Charge Questions 3&8

A.R. Schnatter
December 18, 2006

Charge question 3
Are the methods planned for selecting source categories with potentially significant ecological risks or multi-pathway human health risks for a separate, more refined ecological and multi-pathway assessment sufficiently health-protective? Are there ways that you might suggest for improving such screening techniques that can make them less conservative and still scientifically defensible?

The screening risk assessments for these compounds appear, in most cases, to be sufficiently health protective, but the selection of the compounds should be scrutinized more closely to assure that compounds and sources that may have such effects are not missed. Comments consider both the selection methods and the procedures used in the screening risk assessments.

Selection of PB compounds
The proposed method uses existing Agency tools/policies that are not intended to prioritize ecological risks associated with HAPs.

The proposed method assumes that the 14 HAPs designated as "persistent and bioaccumulative" (Appendix 5) may pose an adverse environmental effect and are thus the focus of the screening procedure proposed. This list of priority HAPs were selected based on previous EPA priority lists/policies:

- PBT profiler
- Great Waters pollutants of concern
- TRI PBT rule

Based on information provided on EPA's website, the PBT Profiler is a screening tool, and PBT estimations rendered by it are not sufficient for definitive PBT determinations. The profiler is a research rather than a regulatory tool, and is used to identify chemicals that may need further evaluation for potential Persistence, Bioaccumulation and Toxicity characteristics. The use of the PBT profile for prioritizing HAPs that pose ecological risks is can be questioned.

Pollutants of concern for the Great Waters policy were selected on the basis of available data on effects and deposition. These pollutants were known to be persistent and/or bioaccumulative and cause adverse effects in humans and the environment. All pollutants were known to occur specifically in Great Waters with atmospheric deposition being one potential source. It should be pointed out that not all the compounds listed in this policy were considered persistent and bioaccumulative (e.g. cadmium, BAP). For example, BAP, and PAHs (POMs) in general, were designated as chemicals that were not bioaccumulative chemicals of concern in EPAs Great Lakes Water Quality Initative.

The TRI PBT rule required reporting thresholds for substances designated as persistent bioaccumulative toxic (PBT) chemicals. These chemicals were judged to be of particular concern

not only because they are toxic but also because they remain in the environment for long periods of time, are not readily destroyed, and build up or accumulate in body tissue. While substances that are persistent and/or bioaccumulative will have higher exposure potential, HAPs that do not possess these attributes may still pose ecological risks.

The approach proposed by EPA also does not differentiate differences in effect endpoints that are relevant to human health versus ecological risk assessment. For example, in developing initial PB-HAP emissions thresholds for POM, facility specific emissions were converted to toxic equivalents of BAP using inhalation cancer risk estimates of 8 categories. However, cancer is not the relevant endpoint for POM's ecological effects. Rather, POM ecologic risks are based on survival, growth and reproduction due to narcosis.

An alternative approach is to develop a prioritization strategy that focuses on a relative human vs. ecological risk assessment process. For most HAPs, protection of human health will likely be protective of the environment. A transparent decision framework should be developed that (1) identifies HAPs where this is not the case (ecological risks rather than human health endpoints drive derivation of air emission thresholds) and (2) addresses relative importance of facility emission relative to background sources for naturally occurring HAPs in prioritization.

Multi-pathway screening risks

Initial review of the preliminary assessment of non-inhalation human exposure to HAPs using a multi-pathway modeling approach raises two general concerns. Firstly, there is a lack of transparency in the model algorithms used to calculate exposure via non-inhalation routes. The selection and use of particular model input parameters is unclear and undefended. Without greater transparency, it is difficult to offer a comprehensive review of the modeling aspects. Second, a simpler modeling approach may help reduce conservatism, while maintaining a high degree of health protection and scientific integrity, by focusing attention on the most important processes driving chemical fate and human exposure relative to the non-inhalation pathway.

The text and attachments of Appendix 5 do not clearly identify the model equations used to estimate human exposure by multimedia pathways. While there are a number of model parameters listed in Attachment A, a description of how these inputs are used is not included. It is also important to note that there is no explanation as to why the listed values have been selected. For example, on page 5-16, a soil mixing zone depth of 20 cm is selected, while a value of 2 cm is used in the HHRAP default. There also appears to be two separate sets of food consumption rate data (page 5-17), however, it is not indicated which (or if these) apply to adults versus children. Moreover, a listed value of 0.074 kg kg^{-1} day^{-1} for milk consumption seems unlikely (5.2 kg day^{-1} for a 70 kg adult or 1.1 kg day^{-1} for a 15 kg child).

An ambient air temperature of 11 ^{o}C is used in this analysis while the HHRAP default is 25 ^{o}C. It is not clear why this value has been used and more importantly, whether or not physical-chemical properties have been adjusted accordingly (i.e. physical-chemical properties are commonly measured at 25 ^{o}C and need to be adjusted for application at different temperatures). There is also concern regarding the selected environmental parameters representing the worst

case scenario. For example, the water flow rate from the lake has been set to zero, which is probably not truly realistic.

The current approach for assessing multipathway exposure to HAPs requires a large number of model input parameters. At a screening level, the goal should be to reduce these inputs to the least number possible. One method may include the use of sensitivity analyses; which help to highlight the parameters that drive chemical exposure. By indicating the most sensitive parameters, efforts can be focused on obtaining the best estimates or measurements for these inputs. This approach can be used to reduce model conservatism. Furthermore, and maybe most useful, would be to use a multimedia fate model to help identify the environmental and physical-chemical properties for which non-inhalation exposure pathways become most important (i.e. the ratio of the intake fraction in water, food or both to the intake fraction by inhalation is greater than 1). Not only are models, such as EQC or RAIDAR more transparent and widely used and accepted, they offer simpler approaches to determine important environmental and physical-chemical parameters.

Charge question 8

8a*) What are the strengths and weaknesses of the overall conceptual approach to risk characterization planned for this assessment?*
Strengths:

1. attention is given to using validated emissions data as the basis for estimating exposure.
2. an attempt is made to provide a consistent approach for assessing risks for different sources.
3. an attempt is made at estimating the effect of uncertainties in risk characterization

Weaknesses:

1. use of a single value (e.g. slope factor or reference concentration) to characterize potency. The uncertainty in these values is not made transparent in final decision making
2. use of potency values from different agencies
3. non-transparent methods to incorporate uncertainties or variabilities into final steps of the process.

Overall, many of the techniques proposed seem most appropriate for a screening risk assessment – i.e. to produce a ranking of source categories that would require more in-depth risk assessments that account more directly for many of the scientific uncertainties inherent in use of slope factors, population estimates, exposure modelling, emission estimates, etc.

8b) Does the characterization plan adequately cover sensitive populations and early life exposures?

It is difficult to answer this question generically. For some HAP's (and therefore some source categories), the cancer slope factors or reference concentrations may have uncertainty factors or other health protective assumptions that would be expected to adequately cover sensitive sub-

populations. Indeed, in some cases, the dose-response information could have been derived from these sensitive sub-populations, and therefore would cover them. There is always the possibility that a more sensitive, unknown subgroup exists. However, in the absence of plausible scientific evidence that these subgroups exist, it would be difficult to account for their "possible" existence.

When known sensitive subgroups exist of a reasonable size, the RTR should account for this sensitivity. On page 24, the RTR-II document states that risk characterizations will assess physiologically susceptible demographic groups or life stages, if known. This is certainly appropriate and should be encouraged.

Page 24 also states that for HAP's acting by a mutagenic mode of action, extra factors of 10 (for children aged 0-1), 3 (for children aged 2-15), or 1.6 (for 70 years beginning at birth) will be applied. This is likely to be overly conservative, since cancer is primarily a disease of the elderly and the longest latent periods (e.g. for asbestos and mesothelioma) are on the order of 40 years. The extra factors proposed for "early life sensitivity" are likely to result in higher risk estimates than are truly present. They may be justified in a screening exercise, (to perhaps make them somewhat comparable to the conservative uncertainty factors used to derive RfC's), but are probably not justified in final risk assessments of source categories.

8c) Does the risk characterization plan appropriately aggregate (cancer and non-cancer) risks?

Page 25 correctly notes that non-cancer HQ's that act by similar modes of action can be aggregated. It's also noted that when this information is absent, HAP's that affect the same target organ will be aggregated to form a target organ specific hazard index. However, some HAP's are likely to act on the same target organ through different modes of action. This situation (same target organ/different MOA) is not a possibility for all unknown MOA's. Since a large number of TOSHI's could be made up of unknown, aggregated MOA's, EPA should consider balancing this conservative assumption with relaxed judgments on the evidence needed to apply a known MOA. The strategy of developing target organ specific hazard indices when an MOA is not known is better justified in the screening assessments rather than the final risk characterizations.

8d) What are the strengths and weaknesses of the planned approach for characterizing important uncertainties, variabilities and limitations? *Given the underlying science and the intended purposes of the assessment, 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?*

This is perhaps the most important, but least clear aspect of the RTR-II assessment plan. Table 2 begins to provide a good framework for assessing various uncertainties (and variabilities) in the RTR-II plan. The outputs listed in section 4.5 include a table of "generic sources of uncertainty and variability for all source categories and for each specific source category. Both of these are strengths. While this is good practice, it is still unclear how uncertainties are factored into final decisions such as whether a new control technology is indicated for a given source, or for a given set of HAP's. This is the primary weakness.

The characterization of uncertainty and variability could be improved and made more transparent by including estimates, where possibly quantifiable in the outputs mentioned in section 4.5. EPA should consider regarding the 'first pass' of the RTR-II plan a screening assessment to rank high priority sources and/or HAP's. Here, uncertainty could be characterized as proposed by the agency. Then, a second step could be made for higher priority sources and/or HAP's identified through the first pass. In this step, a more complete and transparent assessment of variability and uncertainty could be conducted for the sources or HAP's that are ranked high by the screening assessment.

A couple of examples of what this more complete assessment of uncertainty and variability might entail follow. In the exposure assessment area, centroid locations for receptors should be supplemented minimal and maximal distances within blocks. Exposure estimates with and without plume depletion assumptions should be made.

Another example involves dose response values. Rather than point estimates of dose response values, ranges for these should be given. For example, the IRIS URE for benzene is 2.2 to 7.8 x 10^{-6}, and within that range any calculated URE has equal scientific validity. Yet only the 7.8 value is used in risk characterization. Uncertainty in reference concentrations are thought to cover one order of magnitude. Thus a reference concentration that shows this uncertainty explicitly should be provided.

While this process may ultimately be more resource intensive, it involves a better scientific assessment, which should focus attention on the sources and/or HAP's that may be truly affecting population health.

There is one issue in the Table 2, where the magnitude and direction of uncertainties are mentioned. It is suggested that the largest uncertainty is that resulting from not considering background exposures. Yet, as the agency has stated in other places, the goal of the RTR-II plan is to estimate the additional risk from specific sources. Thus, background exposures are appropriately not included. While this could result in underestimation of total exposure, it does not result in underestimation of exposure or risk from the sources included in RTR-II. Since Table 2 includes the influence on risk estimates from sources or HAP's, "background risk" should not be included, or if it is, the influence on risk estimates (from sources) should be "mixed".

Final Comments from Dr. Jeff Fisher

Charge Question 5. Plan describes screening methods for identifying important short term exposures. Is the method protective? Can the method be refined to support acute exposure assessments?

ACUTE EXPOSURE APPROACH
Page 7, fourth paragraph, **'It will be assumed that the maximum one-hour emission rate from a source is ten times the average annual emission rate for that source.'** The factor of 10 is based on a paper by Allen et al. (2004) using VOC data. More specific information may be considered during the ANRPM comment period, especially if the screening identifies acute exposures of concern.

ACUTE EXPOSURE DATA BASE FOR TOXICITY SCREENS
Appendix 6 contains a list of AEGL-1 and AEGL-2 values, ERPG-1 and ERPG-2, MRL, and REL values.

Comments
It is unclear to me what the relationship between increasing the emission rate by a factor of ten and the resultant model predicted exposures or the health risks. The authors need to edit their text to include more information about the nature of the data they receive for use in dispersion/risk modeling and explain how they perform exposure simulations after using the 10x factor. The authors do more computational work than is described in the text of this document. My questions below reflect some of my uncertainty about what is actually carried out with the calculations.

The Allen et al. (2004) paper, as cited by the authors of this Plan, suggest that short term emissions are greater than annualized emissions by a factor of 2-9 fold. It appears that the resulting exposures can vary many fold, depending on assumptions such as weather conditions. I think a better description of the consequences of the 10x factor is needed. If this assumption holds for many source categories, this appears to be a public health protective measure. It would be better to obtain stronger justification for this phenomenon. Are there monitoring data that can support this? Is the under-prediction of short term exposures by the modeling programs a function of the assumptions or model parameters? It appears that the data used in the model is in units that do not lend itself to short term analysis? Please explain this for people who do not do this type of computational work. Why do you need to stay with hourly intervals and annualized information? Can the models be changed to predict an 8 hr exposure with 5 min increments?

The implementation of the acute exposure toxicity data base is not well defined. I think some effort is needed to determine best how to use existing data bases (Appendix 6), a priori. I have copied and pasted definitions of the toxicity values for the various data bases (see below). There are several considerations for using these data bases, such as how current is the chemical value in the data base, does the toxicity value makes sense relative to what you are doing (professional judgment) and how is the toxicity value interpreted relative to how it will be used in the Risk and Technology Review Assessment Plan. I am on the NAS AEGL subcommittee and most familiar with this program. I think the REL values, by definition, may be most useful for this situation.

These toxicity values represent hot spots from which the public may have one hour intermittent exposures, where AEGL and ERPG values represent emergency situations and one time exposure (in theory). MRL values may be useful if you are interested in longer term exposures. I am unsure if much is known about temporal aspects of atmospheric exposures. AEGL values for carcinogens contain a cancer risk calculation to ensure that the theoretical risks of cancer are not exceeded for a short term high exposure. I believe this rarely happens.

AEGLs are developed for *emergency situations and **one time exposure.*** Thus, when these values are finally approved by the NAS/AEGL subcommittee, single exposure data is much preferred over repeated exposure data.

> AEGLs represent threshold exposure limits for the general public and are applicable to emergency exposure periods ranging from 10 minutes to 8 hours. Three levels — AEGL-1, AEGL-2 and AEGL-3 — are developed for each of five exposure periods (10 and 30 minutes, 1 hour, 4 hours, and 8 hours) and are distinguished by varying degrees of severity of toxic effects. The three AEGLs are defined as follows:
>
> AEGL-1 is the airborne concentration (expressed as parts per million or milligrams per cubic meter [ppm or mg/m^3]) of a substance above which it is predicted that the general population, including susceptible individuals, could experience notable discomfort, irritation, or certain asymptomatic, non-sensory effects. However, the effects are not disabling and are transient and reversible upon cessation of exposure.
>
> AEGL-2 is the airborne concentration (expressed as ppm or mg/m^3) of a substance above which it is predicted that the general population, including susceptible individuals, could experience irreversible or other serious, long-lasting adverse health effects or an impaired ability to escape.
>
> AEGL-3 is the airborne concentration (expressed as ppm or mg/m^3) of a substance above which it is predicted that the general population, including susceptible individuals, could experience life-threatening health effects or death.
>
> Airborne concentrations below the AEGL-1 represent exposure levels that could produce mild and progressively increasing but transient and nondisabling odor, taste, and sensory irritation or certain asymptomatic, non-sensory effects. With increasing airborne concentrations above each AEGL, there is a progressive increase in the likelihood of occurrence and the severity of effects described for each corresponding AEGL. Although the AEGL values represent threshold levels for the general public, including susceptible subpopulations, such as infants, children, the elderly, persons with asthma, and those with other illnesses, it is recognized that individuals, subject to unique or idiosyncratic responses, could experience the effects described at concentrations below the corresponding AEGL.

California Air Toxic Hot Spots Program (REL)

The objective is to present a method for deriving acute (one-hour) inhalation Reference Exposure Levels (RELs) for hazardous airborne substances. The acute REL is an exposure that is not likely to cause adverse effects in a human population, including sensitive subgroups, exposed to that concentration for one hour **on an intermittent basis**. These health-based acute RELs are applicable to risk characterization of air releases, defined in Health and Safety Code Section 44303, as: "including actual or potential spilling, leaking, pumping, pouring, emitting, emptying, discharging, injecting, escaping, leaching, dumping, or disposing of a substance into the ambient

air and that results from routine operation of a facility or that is predictable, including, but not limited to continuous and intermittent releases and predictable process upsets or leaks."

1.1.1 Definition of Reference Exposure Level (REL)
The concentration level at or below which no adverse health effects are anticipated for a specified exposure duration is termed the reference exposure level (REL). RELs are based on the most sensitive, relevant, adverse health effect reported in the medical and toxicological literature. RELs are designed to protect the most sensitive individuals in the population by the inclusion of margins of safety. Since margins of safety are incorporated to address data gaps and uncertainties, exceeding the REL does not automatically indicate an adverse health impact.

ERPG

The Emergency Response Planning Guidelines (ERPGs) were developed by the ERPG committee of the American Industrial Hygiene Association. The ERPGs were developed as planning guidelines, to anticipate human adverse health effects caused by exposure to toxic chemicals. The ERPGs are three-tiered guidelines with one common denominator: a 1-hour contact duration (Figure 1). Each guideline identifies the substance, its chemical and structural properties, animal toxicology data, human experience, existing exposure guidelines, the rationale behind the selected value, and a list of references.

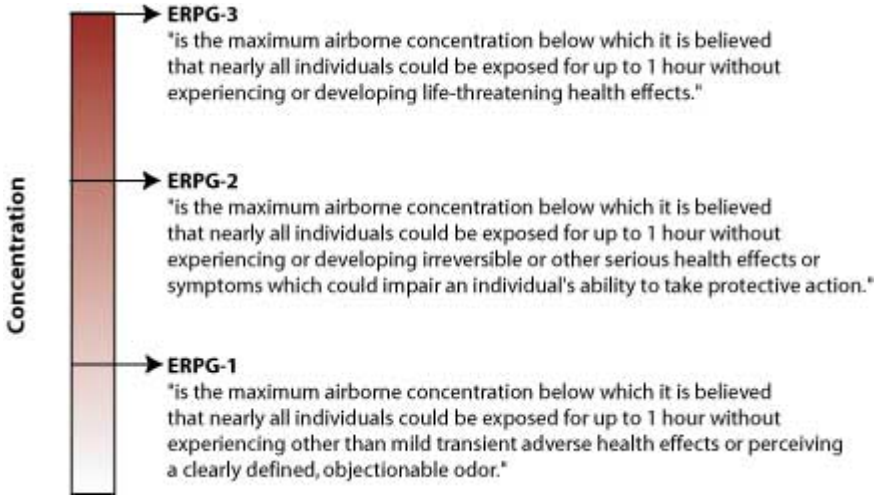

FIGURE 1. The three-tiered ERPG public exposure guidelines. The definitions and format are from the ERPG publication.

The ERPG guidelines do not protect everyone. Hypersensitive individuals would suffer adverse reactions to concentrations far below those suggested in the guidelines. In addition, ERPGs, like other exposure guidelines, are based mostly on animal studies, thus raising the question of applicability to humans. The guidelines are focused on one period of time: 1 hour. Exposure in the field may be longer or shorter. However, the ERPG committee strongly advises against trying to extrapolate ERPG values to longer periods of time.

The most important point to remember about the ERPGs is that they do not contain safety factors usually incorporated into exposure guidelines such as the TLV. Rather, they estimate how the general public would react to chemical exposure. Just below the ERPG-1, for example, most people would detect the chemical and may experience temporary mild effects. Just below the ERPG-3, on the other hand, it is estimated that the effects would be severe, although not life-threatening. The TLV, on the other hand, incorporate a safety factor into their guidelines, to prevent ill effects. **The ERPG should serve as a planning tool, not a standard to protect the public.**

MRL

MRLs are derived when ATSDR determines that reliable and sufficient data exist to identify the target organ(s) of effect or the most sensitive health effect(s) for a specific duration for a given route of exposure to the substance. MRLs are based on noncancer health effects only and are not based on a consideration of cancer effects. Inhalation MRLs are exposure concentrations expressed in units of p ATSDR uses the no observed adverse effect level/uncertainty factor (NOAEL/UF) approach to derive MRLs for hazardous substances. They are set below levels that, based on current information, might cause adverse health effects in the people most sensitive to such substance induced effects. MRLs are derived for **acute (1- 14 days)**, intermediate (>14 364 days), and chronic (365 days and longer) exposure durations, and for the oral and **inhalation** routes of exposure. ….Parts per million (ppm) for gases and volatiles, or milligrams per cubic meter (mg/m3) for particles. ATSDR does not use serious health effects (such as irreparable damage to the liver or kidneys, or birth defects) as a basis for establishing MRLs. Exposure to a level above the MRL does not mean that adverse health effects will occur.

Charge Question 6. I think the idea of using dispersion models and risk assessment procedures for technology control is an important step toward national methods and procedures. Keep up the good work! The question of providing sufficient details to refine the risk estimates by accounting for human daily activity/terrain, may go beyond the screening level. I would prefer to say, for a given location, that the emission restrictions are sufficiently protective to protect some who resides in the area of concern without leaving the region. If this is the 'worst case exposure' calculation, then this becomes a bench mark, which is probably not realistic, but provides a worst case scenario for the protection of public health. The use of population movement patterns is desirable. The text should articulate that the calculations do not account for risks achieved from living in other regions of the US or world and vise versa, when someone leave the region of concern.

Response to SAB Charge Question Two
Responses Provided by Dr. Mark J. Rood, Primary Reviewer
December 19, 2006

Overall, the "Risk and Technology Review Assessment Plan" (RTR II) is well written when discussing the emissions and source data analyses.

2. Emissions and Source Data: The NEI for hazardous air pollutants represents an ongoing voluntary national effort whose creation results from the collaborative efforts of State, local, and tribal air agencies with EPA Regional and Headquarters staff.

a. Short of creating a federal mandate for reporting emissions to the EPA, do the methods by which the NEI was developed, reviewed, and compiled result in a technically-credible database that can support regulatory assessment and action? If not, can you suggest ways to improve it?

Yes, the methods by which the NEI were developed, compiled, and reviewed will result in a technically-credible database that can support regulatory assessment and action

The report could be strengthened by describing the sources of "voluntary" data and what is done when the data are not provided for relevant sources. There is also a reference to "EPA's National Emissions Inventory (NEI) contain[ing] 2002 emissions data and source characterization information for sources of HAP emissions[3]" in the RTRII on page 4. However, reference 3 refers to a report about "industrial VOC emissions and their impact on ozone formation." It would be useful to include a brief description of the source(s) of the emission data that were used for the draft 2002 national inventory and a list of the relevant references that describe those sources, instead of referring to the 1999 NEI and the TRI as mentioned in the RTR II (page 9).

b. Do the plans for conducting an engineering review and incorporating currently-available refined emissions and source data into the inventory add value to the assessment?

Completion of a careful engineering review of the parameters used as inputs to dispersion models (e.g. existence of the sources, emission inventories, physical stack parameters, stack gas properties, temporal variability of the source strength, and location of stack with respect to plant boundary) is an important part of RTR II. The reported intent of the review is to screen the data to readily identify short-comings and problems. However, the example RTR questionnaire and summary spreadsheet, that is located in Appendix 1, refers to "looks accurate," "looks correct," and "looks reasonable." There is no guidance as to how to interpret those qualitative assessments for a quantitative emission inventory. There is also reference to significant discrepancies between two inventories (i.e. > 50%) that will initiate additional actions. It would be constructive to justify the magnitude of such difference. Inclusion of select sensitivity analyses will strengthen the report when describing criteria used to make decisions about the database.

The maximum one hour emission rate is assumed to be ten times the annual average emission rate, and such assumption is based on reference 3 (Allen et al., 2004). Short-term emissions of pollutants related to ozone formation (e.g. NOx, VOC, HRVOC, 1,3-butadiene, butene, propene, and ethene) were compared to their annual average emissions in that report. It was unusual (e.g. a few events per year) for the short-term emission rate to be greater than 10 times the annual average emission rate for the four county region. However, the actual short-term emissions could have local impact on air quality over limited durations, but overall it appears that such short-term emissions do not add

significantly impact to the *annual based* inventories *for the studied region* (four counties around Galveston TX).

The draft RTR II reports that the highest peak emission event was $\leq$ 8.5 times the annual average. However, Figures 9, 10 (i.e. 1,500 lb 1,3-butadiene/event compared to 97 lb/hr annual average), 12, and 13 appears to report events where short term emission rates were more than 10 times the annual averaged emission rate. Overall, the assumption of using 10 times the annual average for short term events appears reasonable, assuming that the acute risk assessments are applicable for a region, not a specific source.

Does the plan for soliciting public comment through an advanced notice of rulemaking add scientific credibility to the inventory?

Soliciting public comment through the advanced notice of proposed rulemaking (ANPRM) is an important aspect of the draft RTR II, especially due to the voluntary nature of the emission inventory that is used as an input to the dispersion models. Potential oversights that could occur during the development and review of the emission inventories by USEPA could be overcome by taking advantage of information provided during the public comment period, and such effort is expected to improve the scientific credibility of the emission inventory. Facilities and States that have not provided emission data could be more receptive to providing such information during the public comment period.

Is the plan for reconciling comments on the inventory adequate?

The methodology used to reconcile differences between comments provided during the ANPRN and the draft 2002 national emission inventory are described in detail in Section 2.1.3, on p. 8 of the RTR II Plan. The Plan includes descriptions of how USEPA will review the comments about the emission inventory, update the database, and provide responses from the ANPRN to the original?/new? data providers. There is reference to a detailed Quality Assurance document that describes quality assurance issues for the development of the draft 2002 national emission inventory. Such report complements the discussion about how to resolve differences between the comments obtained during the ANPRN and the national emission inventory, but it is not apparent how the quality assurance issues will be considered when updating the emission inventory based on comments from the ANPRN.

If not, can you suggest other approaches for reconciling such comments?

The approach proposed by USEPA appears to be reasonable.

Response to SAB Charge Question Four
Responses Provided by Dr. Mark J. Rood, Secondary Reviewer
December 19, 2006

4. Dispersion Modeling: Does the coupling of the AERMOD dispersion model with the census block human exposure modeling (HEM) approach to estimating individual and population exposures represent a credible approach for this goal?

AERMOD is a Gaussian based dispersion model that is very useful, but also has its limitations (e.g. estimation of plume dispersion coefficients, treatment of chemically reactive species during transport, extent of modeling domain (50 km radius), surface roughness, and plume reflection). However, I am unaware of the current status of regional or global models that could be used for the same purposes as AERMOD. Use of modeled short-term (e.g. 1 hr) exposures to estimate actual short-term exposures could provide a wide range of differences in those values. Text in the RTR II indicates that pollutant concentrations will be over-estimated for materials that exhibit transformations during transport, but such transformations should be very small during transport in AERMOD's modeling domain. Such approach does not consider reactions such as the formation of pollutants (e.g. ozone, gas to particle conversion, and oxygenated organic species). These pathways will result in elevated concentrations of pollutants when compared to the results provided by AERMOD, especially outside of AERMOD's modeling domain, or within AERMOD's modeling domain when stagnant meteorological conditions exist. Comparison of measured field results to modeled results will help to better understand how well the models predict actual conditions for HAPs.

Effort to couple AERMOD with census block human exposure modeling (HEM) appears reasonable, with sufficient precision (finest resolution is comprised of about 40 people or 10 households). However, the input meteorology is based on 1991 calendar year data and the census data are from 2000. Averaging of multiple years of meteorological data before during and after 2000 appears to be a more representative approach. The report would be strengthened to provide a few sentences for the justification of the time period used for the meteorological data compared to the census data.

Overall, the description of using AERMOD, AERMET, and HEM is well written to describe their trade-offs. The approach described in the RTR II appears credible, but the errors caused by invoking the assumptions in the models are difficult to quantify.

Are there other more credible approaches available for the estimation of inhalation risks from the types of source categories being examined?

Not to my knowledge.

Is the level of accuracy of this approach acceptable for the purposes of residual risk decision making?

I am not able to locate quantitative analyses about the accuracy of using AERMOD with AERMET, and HEM. However, there are useful discussions describing qualitative issues

that will influence the accuracy of this approach (p. 12-13). Hence, I am not able to comment on the level of accuracy of this approach as it pertains to residual risk decision making.

Are there any specific source categories, sources, or pollutants for which this approach might be considered inadequate?

Pollutants that experience chemical reactions (e.g. SO_2, photochemical reactants (e.g. NO_x, VOCs)), exhibit gas to particle conversion (e.g. H_2SO_4, NH_3), hygroscopic materials that are removed from the gas phase by clouds or precipitation (e.g. alcohols, organic acids, SO_2 and HNO_3), and are lost from the atmosphere due to dry deposition (e.g. particulate material) are examples of materials that appear to be inadequately considered by AERMOD. There does not appear to be adequate treatment of materials that experience transformations once they are deposited from the atmosphere (e.g. elemental and ionic Hg), but such multi-pathway considerations are expected to be treated in a follow-up RTR.

Comments from Dr. M. Morandi. Charge Question 1.

Charge Question 1. Scope: Is the scope of the assessment appropriate for the stated purpose? Is the overall approach clearly and adequately explained for review?

The scope of the assessment appears appropriate for the purpose of estimating source category-specific residual risks as a tool for prioritizing the Agency's rule making activities. While there are considerable uncertainties inherent in the proposed approach - some of them resulting from limitations of the underlying data (e.g., source emissions) but others deriving from the state of scientific knowledge (e.g., health impact from pollutant mixes) - unless the uncertainties impact differentially the estimates of residual risk for some source categories compared to others, the proposed assessment method will not result in misclassification of the source-specific residual risks in terms of priorities (i.e., misclassifying a source as not having a significant impact on residual risk compared to others, when it fact it does). Given the process proposed, which is conservative, it is unlikely that misclassification of this type will occur, but the Agency should undertake a sensitivity analysis to determine which inputs are the main drivers of the RR estimates and if differences in level of uncertainties for those inputs (for example, uncertainties in emissions from some sources compared to others) could potentially result in such misclassification.

There is some concern that this type of effort, as it has happened with NATA to some extent, acquires a life of its own once "number" are available to the general public and even the broader scientific community, so that they may be misinterpreted and applied for inappropriate purposes. While this is not the fault of the Agency, it is important that the Agency clearly and repeatedly states in any communication to the public that the assessment is intended only for the purpose intended by the Agency and that the estimates should not be construed as absolute estimates of residual risk for use in population-based studies. The current documentation contains caveats, but these should be listed up-front, and perhaps with some further elaboration on the limitations of the approach for uses other than the intended purpose.

The narrative of the approach is presented in a clear manner but it may be useful to add one or more diagrams with decision-making nodes clearly indicated so that it is easier to follow the text. It would also be advisable to add a statement regarding what the Agency proposes to do if and when input data may be insufficient in quantity or quality to make the estimates sufficiently reliable for a specific source type (see the comments about the potential for misclassification.) Performing the sensitivity analysis suggested above may also help in the prioritization in cases where rankings are very close.

Comments from Dr. M. Morandi: Charge question 9.

9. Overall: Has any important scientific information been omitted from this assessment plan that could impact a subsequent regulatory decision? In your opinion, will the overall approach for the 51 source categories provide results that will be sufficient to support regulatory decision-making in the context of EPA's residual risk program?

In general, the methodology incorporates current scientific understanding. One area that needs further discussion is how fugitive emissions, especially those from large facilities, will be treated and what the impact of uncertainties in the emission estimates or the identification/location of a source of fugitive emissions might have in potentially misclassifying sources or strongly underestimating exposures for some categories. It may be possible to identify some of these situations and determine if this is the case. Overall, the methodology will be useful for supporting decision making, but it is not clear at this time if that will be the case for all 51 source categories.

Comments from Dr. Randy Maddalena

Charge question #1:

Scope: Is the scope of the assessment appropriate for the stated purpose?

The scope of the assessment is stated in Section 1.1 of the RTR Assessment Plan in terms of the choice of source categories. Essentially, the scope of the RTR II will be limited to source categories that already have readily available post-MACT emission data in the most recent National Emissions Inventory (NEI) where the primary/dominant route of exposure is inhalation.

The purpose of the assessment is not stated very clearly but the last paragraph of the introduction implies that the purpose of this "new approach" is to save time and money while improving the consistency of the residual risk process and to ultimately satisfy the regulatory requirements of the CAA. This will be done by grouping together source categories and assessing the residual risks concurrently using available emissions data. If this is the "stated purpose" then the scope of the study as indicated in Section 1.1 makes a lot of sense.

Beyond the need for streamlining the residual risk process, the RTR assessment plan (and the residual risk process in general) seems to have two more overarching purposes. The first is to determine whether additional risk reductions are necessary to protect public health and the environment from industrial source categories after they have come into compliance with Maximum Achievable Control Technology (MACT). The second, if necessary, is for the RTR assessment to provide the basis for regulatory decisions concerning specific source categories. It may be that these two overarching purposes are implied by stating that the approach will satisfy the regulatory requirements of the CAA but it would be helpful if the purpose of the assessment were stated clearly stated in the introduction.

Is the overall approach clearly and adequately explained for review?

The report is well written and reasonably concise.

As indicated above, I think it would be helpful if there was a section before the current section 1.1 that clearly stated the purpose of the new approach in the context of the overall residual risk process. This is currently in the introduction but it does not come across very clearly.

Section 1.2, describes the RTR II process but it is difficult to pull out the specific steps and chronology of events/tasks. For example, the section talks a lot about the collection and

evaluation of emissions data but does not talk about the screen for PBTs and multipathway HAPS, which seems to be an important part of the assessment. It would be helpful if this section were written to present each step or task in the process and identify key decision points along the way but without necessarily focusing on the details. A time-line or flow chart might be useful here.

I would like to see a list of acronyms.

In addition to the model summaries provided in the text and appendices, it might be helpful for reviewers if each of the individual models used in the assessment were linked to the relevant page in the EPA's CREM Models Knowledge Base.

FINAL COMMENTS

Richard A. Fenske, Ph.D., MPH
University of Washington
U.S. EPA Science Advisory Board member

#3. Identifying Source Categories with Significant Non-inhalation Risk Potential: *This assessment is only designed to include source categories whose risks are dominated by the inhalation pathway. Are the methods planned for selecting source categories with potentially significant ecological risks or multi-pathway human health risks for a separate, more refined ecological and multi-pathway assessment sufficiently health-protective? Are there ways that you might suggest for improving such screening techniques that can make them less conservative and still scientifically defensible?*

The Agency indicated that it has already completed Part 1 of the screening process. This initial screening was based primarily on a "modeling assessment of emissions of 13 PB-HAPs (excluding dioxins/furans) from a hypothetical facility emitting into a domain that included a 'worst-case' subsistence receptor population, constituting a conservative exposure group." The screening process used an approach that back-calculated the emission rate of each PB-HAP that produced either a lifetime cancer risk of 1 in one million, or a hazard quotient (HQ) of 1 for non-cancer effects at the worst-case modeled receptor. The Agency has referred to this emission rate as the "threshold emission rate" for the PB-HAP.

For the estimation of ingestion exposures, in lieu of site-specific data, the Agency incorporated a health-protective farming/fishing scenario that was believed to represent individuals most exposed to HAPs emitted from the model facility. The estimated exposures were combined with dose-response values for chronic oral exposure recommended by QAQPS for screening-level risk assessments.

Since this was a screening process, the use of "worst-case" assumptions, standard factors, and scenarios would appear to provide an ample margin of safety to protect public health.

The Part 2 screening process will take place after the ANPRM, and will employ a screening version of the Total Risk Integrated Methodology (TRIM) model. The Agency indicates that this screening process will be similar to the Part 1 process, but will take advantage of an improved emission inventory, and will be using an "improved modeling platform". Presumably, this improved modeling platform is the TRIM model. The TRIM model appears to be a fate and transport model that will produce estimates of pollutant concentrations in soil, water, and biota that may be ingested.

The process described in the RTR-II document describes a later, more refined multi-pathway assessment for those source categories deemed to hold the potential for significant non-inhalation risks. However, this second-tier analysis appears to be focused primarily on refined fate and transport modeling (TRIM model). It is not clear how the Agency will refine the human exposure aspects of the estimates; e.g., the receptor population, intake rates.

The Agency should ensure that risk estimates would not be overly conservative in regard to assumptions and scenarios. It would be helpful for the Agency to provide a more detailed discussion as to whether/how it plans to modify the assumptions, standard factors, and scenarios in the second-tier analysis.

Appendix 5 refers to 51 source categories that were subjected to non-inhalation screening thresholds. However, Attachment B in the appendix lists 33 source categories, and 51 sub-categories. Presumably the authors have applied their analysis to sub-categories rather than categories. This point should be clarified.

The criteria used to separate the five source 'categories' with "a significant number of exceedances" from the 17 source 'categories' that indicated a "limited potential" for non-inhalation risks were not clear. And again, the terminology is confusing. For example, the authors list "pulp and paper" as one of the 5 source categories in the exceedance group, but the "pulp and paper MACT I and III" sub-category has only 1 of 127 facilities that exceed the threshold (<1%), whereas "pulp and paper MACT II has 13 of 134 facilities that exceed the threshold (9.7%), and the third pulp and paper sub-category has 8 of 348 facilities in exceedance (2.3%). Are distinctions across these sub-categories important in this screening process? If not, then why is the analysis conducted at the sub-category level? And what should we conclude about the source category "ferroalloys" that has an exceedance rate of 36% (4 of 11), or several other source categories that have double-digit exceedance rates? The Agency could improve their explanation of the Part I process by providing the criteria used to sort source categories into "significant" and "limited potential".

If feasible, a probabilistic analysis of the exposure scenario and other factors would provide more transparency regarding variability and uncertainty in the exposure and risk estimates, and would allow a more informed judgment regarding the extent of conservatism included in the calculation/model.

#5. Acute Exposure Screening: *The plan describes a screening methodology for identifying potentially significant acute exposures from routine emissions. Is this method appropriately protective? Can you suggest ways to refine the proposed acute exposure assessment process to enable it to support decision-making?*

Section 3.2 of plan (pp. 20-22) describes sources of acute dose-response information. It is not clear why, on pages 21 and 22, there are discussions of URE values for formaldehyde, nickel, 2-nitropropane, and POM. Perhaps this information belongs elsewhere.

The authors' judgment that the multiple sources of acute hazard information are probably not comparable seems correct. Each of the four sources have used different criteria for calculating hazardous air concentrations, and within each source values have probably been generated over time with different methodologies. All of these sources are reputable, and the information they provide is probably the "best available" science or professional judgment. The Agency has not delved into these sources to determine what value would be most appropriate for a specific pollutant. It is hard to judge whether these values are appropriately protective without a more careful examination of how the Appendix 6 acute exposure values were derived by the individual sources.

USEPA Science Advisory Board
Risk and Technology Review (RTR) Assessment Plan
Response to Charge Question 2
Bryan W. Shaw, Secondary Reviewer

2. Emissions and Source Data: The NEI for hazardous air pollutants represents an ongoing voluntary national effort whose creation results from the collaborative efforts of State, local, and tribal air agencies with EPA Regional and Headquarters staff.

a. Short of creating a federal mandate for reporting emissions to the EPA, do the methods by which the NEI was developed, reviewed, and compiled result in a technically-credible database that can support regulatory assessment and action? If not, can you suggest ways to improve it?

Development of accurate NEI estimates is always challenging. Incorporating a framework for improving the NEI as more accurate data becomes available is critical.

b. Do the plans for conducting an engineering review and incorporating currently-available refined emissions and source data into the inventory add value to the assessment? Does the plan for soliciting public comment through an advanced notice of rulemaking add scientific credibility to the inventory? Is the plan for reconciling comments on the inventory adequate? If not, can you suggest other approaches for reconciling such comments?

The plans for conducting an engineering review adds value. However, to maximize the benefit, the assessment must be conducted in a thoughtful manner. Furthermore, the subjective nature of the assessment may not provide the level of accuracy or precision needed to be useful.

Public Comment – Yes this is crucial and adequate.

Response to Charge Question 2
Bryan W. Shaw, Primary Reviewer

4. Dispersion Modeling: Does the coupling of the AERMOD dispersion model with the census block human exposure modeling (HEM) approach to estimating individual and population exposures represent a credible approach for this goal?

There are inherent shortcomings associated with any model. In this instance the model is a Gaussian based model. The challenges are primarily associated with dispersion parameters and chemical reactions/deposition. The chemical reactions can be critical as the concentration of pollutant of interest at the receptor can change dramatically due to chemical change or deposition (this can be an increase or decrease). As models are improved to better reflect the atmospheric chemistry, this will improve models usefulness for estimating potential health/environmental impacts associated with emissions of interest.

Attempt to couple AERMOD and HEM seems to be appropriate so long as effort is made to ensure compatibility of data.

Are there other more credible approaches available for the estimation of inhalation risks from the types of source categories being examined?

Not that I am aware of at this time.

Is the level of accuracy of this approach acceptable for the purposes of residual risk decision making?

I do not have quantitative assessment of the accuracy of this approach. However, this does seem to be a reasonable approach so long as the uncertainty and/or bias of the model and disparities between AERMOD and HEM data compatibility are addressed.

Are there any specific source categories, sources, or pollutants for which this approach might be considered inadequate?

Potential source categories that may not be well addressed by this approach include any source where significant chemical reaction is not well defined or where deposition is not well described by the model. I have not identified any such categories.