

August 6, 2012

Office of Environmental Information (OEI) Docket
Mail Code: 28221T
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue NW
Washington, DC 20460
Submitted to the Docket at <http://www.regulations.gov>

Re: Request for Public Comment on the EPA's Draft Toxicological Review of Ammonia: In Support of the Summary Information in the Integrated Risk Information System (IRIS). Docket # EPA-HQ-ORD-2012-0399; FRL-9683-8

Dear Sir or Madam:

The Center for Advancing Risk Assessment Science and Policy (ARASP),¹ which is managed by the American Chemistry Council (ACC), fosters activities to promote the adoption of policies and practices that assure the best available science and methodologies are the foundation for chemical assessments. ARASP is pleased to provide the following comments in response to the U.S. Environmental Protection Agency's (EPA) Federal Register notice announcing a 60-day public comment period and a public listening session for the external review draft human health assessment titled "Toxicological Review of Ammonia: In Support of Summary Information on the Integrated Risk Information System (IRIS)" (EPA/635/R-11/013A).²

In a June 2012 press release³ EPA announced the availability of its draft assessment for ammonia and noted that it represented major progress for EPA in implementing the April 2011 National Academy of Sciences (NAS) recommendations⁴ for improving IRIS assessments. EPA stated

¹ ARASP is a coalition of independent groups and associations that promotes the development and application of up-to-date, scientifically sound methods for conducting chemical assessments and is comprised of the following member organizations: Acrylonitrile Group, ACC Chlorine Chemistry Division, ACC Ethylene Oxide Panel, ACC Formaldehyde Panel, ACC Hexavalent Chromium Panel, ACC High Phthalates Panel, ACC Hydrocarbon Solvents Panel, ACC Oxo Process Panel, ACC Propylene Oxide/Propylene Glycol Panel, ACC Public Health and Science Policy Team, ACC Olefins Panel, American Cleaning Institute, American Petroleum Institute, CropLife America, Halogenated Solvents Industry Alliance, Silicones Environmental, Health and Safety Council of North America, and the Styrene Information and Research Center.

² 77 Fed. Reg. 34039 (Jun. 8, 2012)

³ EPA Draft Ammonia Assessment Available for Public Comment / Draft assessment continues agency's responsiveness to NAS recommendations. Release Date: 06/01/2012.
<http://yosemite.epa.gov/opa/admpress.nsf/d0cf6618525a9efb85257359003fb69d/a8178896daa4af9985257a10005dbf37!OpenDocument>

⁴ National Academy of Sciences (NAS). NRC (National Research Council). 2011. Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde. Committee to Review EPA's Draft IRIS Assessment of Formaldehyde. Board of Environmental Studies and Toxicology. Division of Earth and Life Sciences. Available at http://www.nap.edu/catalog.php?record_id=13142.

that *“The draft assessment uses a new streamlined document structure that is more transparent and clear; includes a template for describing the literature search approach; identifies the strengths and weaknesses of analyzed studies; and describes how EPA applied their guidance, methods, and criteria in developing the assessment.”* ARASP supports science based chemical assessments that utilize transparent and explicit data evaluation criteria and methodologies. ARASP has reviewed the draft Toxicological Review of Ammonia and found that it does not provide adequate detail on the data evaluation and synthesis practices used by the Agency, nor does it include transparent criteria for the implementation of the literature search strategy used to reach the conclusions. The draft also fails to effectively implement the recommendations of the NAS. We raised these issues during oral comments presented at the July 12th Listening Session (see appendices 1 and 2) and are providing more detailed comments in the attachment.

ARASP hopes the EPA will review the detailed comments provided in the attachment and strongly recommends that all future toxicological reviews provide: (1) detailed information about the specific frameworks EPA utilized for synthesizing scientific data and how those frameworks were employed in the toxicological review; (2) the specific criteria employed by the Agency for evaluating study quality; (3) a clear listing of all the exclusion and inclusion criteria used in the literature search along with information about which studies were used in the weight of evidence determination; and (4) data tables that array the scientific information using a mode of action framework. In addition, we do not believe that it is appropriate to use the preamble as a means to communicate new criteria, guidance and approaches, that have not been properly peer reviewed. The adoption of new Agency approaches should be done through an open and robust process that involves peer review and stakeholder participation before being implemented in an assessment. If you have any questions or require additional information please feel free to contact me by phone at 202-249-6707 or via email at Kimberly_Wise@americanchemistry.com.

Respectfully Submitted,

Kimberly Wise, Ph.D.
Senior Director
Chemical Products & Technology Division
American Chemistry Council

Attachment
Appendices 1, 2

INTRODUCTION

The Center for Advancing Risk Assessment Science and Policy (ARASP), which is managed by the American Chemistry Council (ACC), is a coalition of independent groups and associations that promotes the development and application of up-to-date, scientifically sound methods for conducting chemical assessments. ARASP also fosters activities to promote adoption of policies and practices, both within and outside the government, that assure the best available science underlies chemical determinations. As part of our mission, we review chemical assessments from the available scientific literature and those developed by local, state or federal agencies, including the evaluation of U.S. Environmental Protection Agency (EPA) Integrated Risk Information System (IRIS) toxicological reviews.

In June 2012 EPA announced, in a press release,¹ the availability of its draft Toxicological Review of Ammonia (draft assessment) and EPA characterized it as representing major progress for EPA in implementing the April 2011 National Academy of Sciences (NAS) recommendations² for improving IRIS assessments. Specifically, EPA stated that the draft assessment “*uses a new streamlined document structure that is more transparent and clear; includes a template for describing the literature search approach; identifies the strengths and weaknesses of analyzed studies; and describes how EPA applied their guidance, methods, and criteria in developing the assessment.*” ARASP disagrees with the characterization presented in the Agency’s press release and describes our concerns more specifically below. ARASP supports science based chemical assessments that utilize transparent data evaluation criteria and methodologies. However, the draft assessment does not effectively implement the NAS’s recommendations³ nor is it sufficiently transparent in providing adequate detail on the data evaluation and synthesis practices used by the Agency. The draft assessment also does not include criteria that are sufficiently transparent to understand the literature search strategy used to reach the conclusions. In addition, it is still unclear how the EPA applied the guidance, methods and criteria that are discussed in the preamble.

ARASP raised specific concerns regarding the draft assessment’s preamble, data tables and the literature search strategy in the July 12, 2012 Listening Session. The comments (see appendices 1 and 2) that were presented during the Listening Session were focused on the overarching improvements needed to these new elements of the toxicological reviews. As noted in our oral

¹ EPA Draft Ammonia Assessment Available for Public Comment / Draft assessment continues agency’s responsiveness to NAS recommendations. Release Date: 06/01/2012.

<http://yosemite.epa.gov/opa/admpress.nsf/d0cf6618525a9efb85257359003fb69d/a8178896daa4af9985257a10005dbf37!OpenDocument>

² National Academy of Sciences (NAS). NRC (National Research Council). 2011. Review of the Environmental Protection Agency’s Draft IRIS Assessment of Formaldehyde. Committee to Review EPA’s Draft IRIS Assessment of Formaldehyde. Board of Environmental Studies and Toxicology. Division of Earth and Life Sciences. Available at http://www.nap.edu/catalog.php?record_id=13142.

³ The NAS, in its 2011 review, made five general suggestions for improving the EPA’s IRIS assessments: (1) replace long narratives about particular studies with informative and standardized evidence tables for all health outcomes; (2) include a description of the search strategies used to identify relevant studies with exclusion and inclusion criteria clearly articulated and a visual display of the search results near the beginning of the document; (3) utilize standardized approaches for evaluation of critical studies with findings presented in tables to ensure transparency; (4) present clear and expanded descriptions of the rationales for selecting the studies upon which toxicity criteria are based along with graphic displays; and, (5) discuss the weight of evidence supporting the toxicity criteria in a rigorous, systematic and transparent fashion.

comments and set forth below, ARASP believes that the EPA has failed to provide sufficient detail on the Agency’s data evaluation and synthesis practices or its literature search strategy. The following detailed comments will focus on the improvements still needed to the preface, preamble, data tables and weight of evidence practices used in the draft assessment. While these comments are being submitted in relation to the ammonia draft assessment, they can apply to existing draft assessments (e.g. trimethylbenzene) and should be applied not only to this assessment but also future toxicological reviews that use this approach.

COMMENTS

A. Preface

The preface of the draft ammonia assessment notes that it updates a previous 1991 assessment of ammonia which only included an inhalation reference concentration for effects other than cancer. The draft assessment also states that new information has become available, and that the assessment reviews information on all health effects by all exposure routes. The preface of the 2012 draft assessment can be improved by:

- Identifying all the factors that can prompt a chemical review (e.g. EPA statutory, regulatory, or program-specific implementation needs; availability of new scientific information or methodology that might significantly change the current IRIS information) and listing the factors that led to the initiation of the ammonia review. Specifically, if new information has become available then include a sentence or two that denotes the compelling reason for the updated ammonia review (e.g. several new studies on the effects of inhalation exposure have been made available since the 1991 review that prompted a reevaluation of the IRIS values).
- Clearly describing the scope and limitations of an IRIS assessment and how any derived toxicity values should be used. This information should include how derived toxicity values should be used in conjunction with relevant information (e.g. exposure information) to make informed risk management determinations.
- Including information relating to any cooperative agreements, contracts, or memorandums of understanding that the Agency has in place which may have informed the development of the assessment.
- Discussing the findings of other regulatory agencies and why the conclusions and/or derived toxicity values in the IRIS assessment are similar or different. The current draft assessment does mention that other agencies have evaluated ammonia and notes that the other agencies assessments were prepared using different methods but it does not provide any detail on how these methods are different. Specifically, the U.S Agency for Toxic Substances and Disease Registry (ATSDR) and the EPA’s approach are similar and it would be useful for EPA to further explain how the two processes are different.

B. Preamble to IRIS Toxicological Reviews

In the draft assessment, EPA has included a section titled “Preamble to the IRIS Toxicological Reviews” that includes a summary discussion of the scope of the IRIS program, process for developing IRIS assessments, study selection, data evaluation and derivation of toxicity values. In 2011 recommendations by the NAS during its review of the EPA’s draft Formaldehyde assessment, NAS stated:

“Chapter 1 needs to be expanded to describe more fully the methods of the assessment, including a description of search strategies used to identify studies with the exclusion and inclusion criteria articulated and a better description of the outcomes of the searches and clear descriptions of the weight-of-evidence approaches used for the various non-cancer outcomes. The committee emphasizes that it is not recommending the addition of long descriptions of EPA guidelines to the introduction, but rather clear concise statements of criteria used to exclude, include, and advance studies for derivation of the RfCs and unit risk estimates.”

Subsequently, in response, in the draft assessment for ammonia, EPA states:

“Chapter 1 has been replaced with a Preamble that describes the application of existing EPA guidance and the methods and criteria used in developing the assessment. The term “Preamble” was chosen to emphasize that these methods and criteria are being applied consistently across IRIS assessments. The new Preamble includes information on identifying and selecting pertinent studies, evaluating the quality of individual studies, weighing the overall evidence of each effect, selecting studies for derivation of toxicity values, and deriving toxicity values.”

As currently written, the preamble offers an abbreviated view of EPA policies, guidance documents and standard practices but fails to include the detail necessary to provide useful information on how the Agency reviews or weighs the scientific information for inclusion in its toxicological review. Unfortunately, in providing this abbreviated view, critical information has been omitted and the preamble may unduly lead readers to incorrectly interpret EPA guidance. In addition, we do not believe that it is appropriate to use the preamble as a means to communicate new criteria, guidance and approaches, that have not been properly peer reviewed, to the public. The adoption of new approaches should be done through an open and robust process that involves peer review and stakeholder participation before being implemented in an assessment.

Below we provide some specific examples of cases where the language in the preamble is not clear and/or may leave readers, including peer reviewers, with an incomplete understanding of EPA guidance and approaches to risk assessment. We believe that it is important that EPA provide correct information in the preamble, and fixing these specific examples should be the starting point but it will not necessarily resolve our over-arching concerns regarding the use of this preamble.

1. The Scope of the IRIS Program (Section 1, page xi) – In this section, EPA has noted that “*IRIS assessments critically review the publicly available studies to identify adverse health effects from long-term exposure to chemicals and to characterize exposure-response relationships.*” However, a paragraph should be added that discusses the four elements of risk assessment (i.e. hazard identification, dose-response assessment, exposure assessment and risk characterization) and what elements the IRIS program addresses. An IRIS assessment includes only hazard identification and dose-response assessment and this should be clearly stated in the preamble. As well, there should be additional discussion included on what role IRIS assessments play in providing scientific information for use in federal regulatory activities.

2. Process for developing and peer-reviewing IRIS assessments (Section 2, pages xi –xii) – In this section the Agency has provided an overview of the May 2009 revised process for developing IRIS assessments.⁴ In step 4 of the development process (external peer review and comment) it notes a timeline of 3 ½ months or more but does not specify specific time frames for public input prior to the draft assessment being released or denote a time frame for delivery of public comments to the peer review panel prior to the peer review meeting. EPA should include an opportunity for public input as the draft assessment is being developed (e.g. after the data call in and the initial scientific information has been gathered but prior to the completion of the weight of evidence data integration). Currently, when the draft toxicological reviews are released by the Agency they are near final – decisions about the main conclusions are presented as a *fait accompli*, which tends to stifle input that the Agency may find valuable. Involving the public and other stakeholders earlier in the process will enable a more meaningful dialogue that can contribute to the development of the draft toxicological review. This engagement with stakeholders should include the identification of useful mode of action information, applicable data evaluation frameworks to synthesize the scientific information being reviewed as well as other relevant topics.

3. Selecting pertinent studies and evaluating the quality (Sections 3 - 4, pages xii –xiii) – These sections provide a summary of the basic search strategy the Agency utilizes to gather scientific information for inclusion in the toxicological review and offers the key considerations used to select pertinent epidemiological and experimental studies. However, there are several areas where this section could be greatly improved.
 - Section 3.2 provides some key considerations for selecting epidemiological studies and specifically states that “*Cohort studies and case-control studies provide the strongest epidemiological evidence, as they collect information about individual exposure.*” However, not all cohort studies collect information based on individual exposure level; one example of this is cohort air pollution studies that are based on group level exposure (e.g., ambient monitoring). This section

⁴ U.S. EPA (U.S. Environmental Protection Agency). (2009). EPAs Integrated Risk Information System: Assessment development process. Washington, DC <http://epa.gov/iris/process.htm>.

should provide clear guidance as to what type of information would generally be given more or less weight in the data evaluation framework.

- Section 3.2 and 3.3 of the draft assessment purport to provide the key design considerations for selecting pertinent epidemiological and/or experimental studies from the results of the literature search and notes exposure route and duration as key considerations. However, this section does not provide the criteria used by the Agency for selecting studies. These sections should include all the considerations EPA utilizes in selecting a study for inclusion in the toxicological review and which of the criteria are deemed most necessary. Furthermore, EPA does not provide information that would allow the public to replicate EPA’s literature selection process for the chemical being assessed. NAS specifically requested this clarity.
4. Evaluating the quality of individual studies (Section 4, pages xiii – xiv) – This section provides basic information on how the assessment evaluates various design and methodological aspects of the data that could increase or decrease the weight given to a study in the overall evaluations. Some examples listed in this section include: documentation of study design, exposure classification, disease classification and sample size. However, it is not clear from reading this section which elements EPA deems most valuable for a study to possess for use in its data evaluation. EPA can improve this section by:
- Providing the specific elements or characteristics that would increase or decrease a study’s weight (e.g. does a low sample size decrease the weight of a study in the overall evaluation of the available scientific information). This section should include at least one paragraph that identifies a list of the design or methodological aspects that increase weight and one paragraph that list the aspects that decrease weight.
 - Expanding the discussion on the use of historical controls. The draft assessment should clearly note that in the EPA’s 2005 cancer guidelines⁵ discussion on the use of historical controls it clearly states: “*However, caution should be used in interpreting results.*”
5. Weighing the evidence and derivation of toxicity values (Sections 5 – 7, pages xiv – xx) – These sections discuss how the Agency evaluates the scientific evidence as a whole to determine the extent to which any observed association may be causally linked to the chemical of interest. EPA’s narrative notes that positive, negative and null results are given weight according to the study quality and provides some aspects to consider in making that association to causality (i.e., strength of association, temporal relations,

⁵ U.S EPA (2005a) Guidelines for carcinogen risk assessment (EPA/630/P-03/001F).
<http://www.epa.gov/cancerguidelines/>.

biological plausibility). However, the section does not indicate how EPA assigns weight to studies or whether, for instance, studies of similar quality are given equal weight regardless of whether the study’s results are positive, negative or null. EPA’s weighting scheme should be discussed in more detail and clear criteria should be provided for increasing and decreasing weight. Information should be included in this section on how positive, negative and null studies are evaluated and weighted (i.e. are they given equal weight). The preamble also does not clearly identify which weight of evidence approach(es) EPA supports or utilizes. EPA should provide a listing of data evaluation practices that are used in the toxicological review. Additional examples where the sections could be improved are provided in the bullets below:

- Section 5.1, cites several references for weighing and synthesizing but it is unclear how or if EPA followed the references noted in this section. For instance, CDC 2004 is cited as an example of a way to make clear how epidemiological evidence contributes to the overall weight of evidence using specific descriptors. It was not apparent whether this framework was applied in the draft assessment.
- Section 5.1 begins a discussion regarding the criteria for causality, however later the discussion moves away from causality to focus on determining whether or not an ‘association’ exists. IRIS assessments should retain a focus on whether evidence of causality exists for compounds of interest.
- On page xv, that draft assessment states “*Negative results carry less weight , partly because they cannot exclude the possibility of effects on other tissues*”, however this appears to over simplify the International Agency for Research on Cancer (IARC). IARC 2006 actually states that: “*Negative results in tests for mutagenicity in selected tissues from animals treated in vivo provide less weight, partly because they do not exclude the possibility of an effect in tissues other than those examined.*” By citing IRAC, EPA should clarify whether it plans to adopt IARC guidance as EPA guidance. As stated, it is unclear what should be considered EPA guidance and what approaches EPA utilizes.
- Section 5.3 states that “*Information suggesting quantitative differences in doses where effects would occur in animals or humans is considered in the dose-response analysis but is not used to determine relevance. Similarly, anticipated levels of human exposure are not used to determine relevance.*” These statements are not in the EPA’s 2005 cancer guidelines thus it seems inappropriate for EPA to use the preamble to create new guidance. This statement should be removed from the preamble. It is also unclear why dose- response analysis would not be used as part of the weight of the evidence evaluation to determine relevance.

- In section 5.3, Guyatt et al. 2008a⁶ appears to be misquoted. The paper does not state that *“the credibility of a series of studies is reduced if evidence is limited to studies funded by one interested sector.”* It is not clear why EPA has decided to quote one sentence, but does not appear to implement the overarching GRADE approach that is discussed in multiple Guyatt publications. Nevertheless it is unclear if EPA intends to adopt this and other publications as Agency guidance.
- In section 5.3 EPA states: *“Key data include the ability of the agent or a metabolite to react with or bind to DNA, positive results in multiple test systems, or similar properties and structure-activity relationships to mutagenic carcinogens (U.S. EPA, 2005a).”* This statement does not appear to be included in EPA’s cancer guidelines. Is EPA implying that negative data would not be equally considered if it was of equal quality? EPA should not set new guidance using this preamble. This sentence should be removed from this and future preambles.
- In section 5.4, it is unclear why EPA focuses only on the cancer descriptors, when the cancer guidelines state: *“Users of these cancer guidelines and of the risk assessments that result from the use of these cancer guidelines should consider the entire range of information included in the narrative rather than focusing simply on the descriptor.”* (emphasis added by EPA). The cancer guidelines correctly note that the complete narrative *“preserves the complexity that is an essential part of the hazard characterization.”* EPA should change this language so that it is consistent with the 2005 cancer guidelines.
- To ensure completeness in stating EPA policy regarding when the Agency would derive a toxicity value, Section 6 should also include a statement from the cancer guidelines that states: *“When there is suggestive evidence, the Agency generally would not attempt a dose-response assessment, as the nature of the data generally would not support one; however, when the evidence includes a well-conducted study, quantitative analyses may be useful for some purposes, for example, providing a sense of the magnitude and uncertainty of potential risks, ranking potential hazards, or setting research priorities. In each case, the rationale for the quantitative analysis is explained, considering the uncertainty in the data and the suggestive nature of the weight of evidence. These analyses generally would not be considered Agency consensus estimates. Dose-response assessments are generally not done when there is inadequate evidence, although calculating a bounding estimate from an epidemiologic or experimental study that does not show positive results can indicate the study’s level of sensitivity and capacity to detect risk levels of concern.”*

⁶ Guyatt GH, Oxman AD, Vist G, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, for the GRADE Working Group. Rating quality of evidence and strength of recommendations GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-926

- In Section 7.5, it is unclear what EPA means by “suspected carcinogens” and EPA should ensure that this language is consistent with its 2005 cancer guidelines and only include terms from the guidelines in the assessment.
- Section 7.6 does not adequately characterize what an oral reference dose (RfD) or an inhalation reference concentration (RfC) are because the text does not clearly state that RfD and RfC values are estimates, with uncertainty spanning perhaps an order of magnitude. EPA should correct its description in the assessment.
- Additional explanation is needed in the “Conflicting evidence” and “Differing results” sections on page xv. In these sections EPA states that “Negative or null results do not invalidate positive results in a different experimental system. EPA regards all as valid observations and looks to methodological differences or, if available, mechanistic information to reconcile differing results.” However, EPA does not provide information regarding what type of methodological differences it would look for or examples of the types of mechanistic information that would assist the Agency in reconciling the differences in the scientific data. EPA should clarify and provide this information in the assessment specifically because if such data exist the public will be alerted to provide such information to EPA.
- Section 7 of the preamble, which summarizes the Agency’s process for deriving toxicity values, including how it performs dose-response modeling, selects points of departure and applies uncertainty factors, could also be greatly improved by providing added clarity. For example, section 7.3 discusses dose-response modeling and notes that *“For dichotomous responses, the point of departure is often the 95% lower bound on the dose associated with a 10% response, but a lower response that falls within the observed range may be used instead.”* This appears to be new guidance and is not consistent with the recently finalized Benchmark Dose Technical Guidance.⁷
- While the cancer guidelines state that non-linear modeling should be selected when there is *“sufficient data to ascertain the mode of action....”*, section 7.4 appears to misinterpret the EPA cancer guidelines, which states *“Where alternative approaches with significant biological support are available for the same tumor response and no scientific consensus favors a single approach, an assessment may present results based on more than one approach.”* Thus the guidelines imply that having sufficient data to ascertain mode of action, is not the only time when a non-linear model may be considered. EPA should make the conforming changes to the assessment to be consistent with its guidelines.
- Section 7.6 provides some discussion regarding uncertainty factors (UF) however it is unclear what the Agency’s policy is on the application of UFs. In this section,

⁷ U.S. EPA. 2012. Benchmark Dose Technical Guidance.
http://www.epa.gov/raf/publications/pdfs/benchmark_dose_guidance.pdf

EPA has appeared to create new policy in stating that the UF for human variation is reduced only if the point of departure is derived specifically for susceptible individuals. EPA should provide clear criteria for the application of uncertainty factors and discuss how the Agency considers UFs in totality to ensure that any compounding conservatism in the derivation of a toxicity value does not lead to an unrealistic final value.

C. Literature Search Strategy

EPA’s draft assessment includes an explanation of the literature search strategy and study selection criteria the Agency used to identify studies for inclusion in the draft assessment. Table LS-1 provides the search parameters and terms used to identify studies; Figure LS-1 provides a schematic for how the Agency narrowed the available scientific literature. Table LS-1 provides useful and sufficient detail and should be maintained, in its current form, in future toxicological reviews. We have included below several areas where the transparency of the literature search could be greatly improved:

- Figure LS-1 needs to be further expanded to include more detailed information regarding the criteria EPA used to include or exclude studies from consideration in the ammonia assessment. For example, Figure LS-1 indicates that 220 human studies, 203 animal studies and 599 supporting studies were found for a total of 1022 studies which were considered for inclusion in the draft assessment. 781 of these studies were excluded for various reasons (e.g. inadequate exposure characterization) but no breakdown has been included regarding the number of studies that were excluded for each of the exclusion categories provided. One example where the exclusion criteria are unclear is in the instance where nonstandard animal model (e.g. nonmammalian species, cattle) is noted as an exclusion criteria however, Appendix D of the draft assessment titled “Information in Support of the Hazard Identification and Dose-Response Analysis” includes a table (i.e. Table D-12) which discusses cattle studies. It seems that based on the exclusion criteria this study information would not have been used to support the hazard characterization or dose-response analysis.
- Additional figures for the study selection criteria for human studies, animal studies and supporting studies should be included as individual figures. Each figure should clearly show the specific exclusion criteria for each study type. Including these as separate figures would provide the space needed to include the additional detail on the reasons for exclusion.
- Specifically, on page xii, EPA states that the literature search was conducted following “standard practices.” EPA should clearly explain what is meant by “standard practices.”

D. Data Tables

One of the recommendations of the NAS to improve IRIS assessments was to replace long narratives about particular studies with informative and standardized evidence tables for all health outcomes. EPA has included new evidence tables in the draft assessment that generally address the suggestion of the NAS. Tables 1-1 and 1-2 in the draft assessment provide a summary of the effects observed in animals and humans respectively. ARASP offers the following recommendations for improving the utility of the data tables:

- Information on the statistics is missing from a portion of Table 1-1 on respiratory symptoms; for example, in the summaries of Holness et al. (1989)⁸ and Rahman et al. (2007),⁹ p-values are provided with no indication of which statistical tests were used. As well, the footnotes for Table 1-1 provide information on the confidence in the air sample measurements and more information was provided later in the document; however, it would be advisable to provide a measure of confidence in the exposure quantification, e.g., low, medium, high, somewhere in the tables where this information would be more easily seen than in the footnotes. Narrative about the exposure quantification is provided on page 2-4 but, to follow the NAS’s recommendation for informative evidence tables, this information should also be included in Table 1-1.
- Figure 1-1 should also include the derived EPA reference concentration (RfC) value in order to illustrate where the RfC falls relative to the lowest observable adverse effects or the no observed adverse effects noted in the relevant scientific studies.
- The results from various test species (e.g., rodent, monkey, pig, and dog) were presented in Table 1-2 and Figure 1-1 in a seeming random fashion. Are some species more sensitive to some of the observed effects? The effects observed in animals studies, summarized in Table 1-2 and Figure 1-1, should be ordered in terms of their level of adversity; and their occurrence (e.g. early or late) within the mode of action. Separating the animal studies in terms of test species could show species differences in sensitivity and might also help with understanding the mode of action
- There should be a clear correlation as to how the data tables connect to the literature search strategy. In the draft assessment, 75 human studies were identified in the literature search but how did EPA determine to include only 3 studies in the data tables? Specific information should be provided on how and why studies were selected for further focus.

⁸ Holness, D.L., J.T. Purdham and J.R. Nethercott. 1989. Acute and chronic respiratory effects of occupational exposure to ammonia. *Am. Ind. Hyg. Assoc. J.* 50(12): 646-650.

⁹ Rahman, MH; Bråtveit, M; Moen, BE. (2007). Exposure to ammonia and acute respiratory effects in a urea fertilizer factory. *Int J Occup Environ Health* 13: 153-159.

E. Weight of Evidence Evaluation for Ammonia

In the preamble, EPA has provided some information regarding how it conducts systematic reviews that evaluate the quality and weight of the scientific information used in an IRIS assessment. However, the draft assessment does not provide sufficient detailed information concerning how EPA used the ammonia literature to derive toxicity values or how the Agency conducted its weight of evidence evaluation. One of the NAS’s recommendations for improving IRIS assessments included presenting clear and expanded descriptions of the rationales for selecting the studies upon which toxicity criteria are based and, discussing the weight of evidence supporting the toxicity criteria in a rigorous, systematic and transparent fashion. In the draft assessment, the rationale used to select the critical studies was presented in section 2.2.1. However; there is little information on how a study’s strengths or weaknesses were used to inform the ammonia weight of evidence. EPA should add a table that specifically denotes the strength and weaknesses of a study and the reasons for excluding a seemingly pertinent study. Several examples where the weight of evidence information was lacking are included below:

- The narrative on page 2-2 of the draft assessment indicates that the evidence for associations of ammonia with toxicity to target organs other than the respiratory system is weak; however, in Figure 2-1, no indication is given as to why the immune system effects or other systemic effects were not chosen.
- The selection of the critical study is not clearly supported. The RfC derived in the draft assessment was based on the Holness et al. 1989 study with the critical endpoint being decrease lung function and increased respiratory symptoms. However; (1) no statistically significant differences were noted between the control and exposure groups for respiratory irritation, (2) no changes in lung function was observed between control and exposures groups and (3) no relationship was demonstrated between chronic ammonia exposure and lung function changes when looking at level of exposure or duration of exposure. In addition, EPA frequently (see for example page xxiv, line 8), mischaracterizes Holness as part of a body of literature which consistently demonstrated an increased prevalence of symptoms. The Holness study did not detect differences in lung function or any other markers of lung function. Thus its selection as the critical study and its inclusion as a study that is shown to support EPA’s determination that the weight of evidence identifies respiratory system effects as a hazard seems unsupported.
- The explanation of how the endogenous production of a chemical is considered in the overall weight of evidence is lacking and needs more detail. The draft assessment includes a section that acknowledges the endogenous production of ammonia and notes that the draft RfC value “*falls within the range of concentrations measured in the mouth or oral cavity.*” EPA then further states that because exhaled breath is diluted in ambient air it would not contribute to ammonia exposure. However, the rationale that EPA has used to justify setting an RfC at a level equivalent to the internal human breath level is unclear. EPA should provide clear

justification for setting an RfC that is within the range of natural human breath levels.

SUMMARY

ARASP supports the development of chemical risk assessments that reflect up to date scientific knowledge, methods and practices. Advancing the technical quality and objectivity of EPA IRIS assessments, particularly by ensuring transparency in what science is being considered and how it is being interpreted, and integrating within an assessment, will go a long way to assuring that potential risks are objectively evaluated. ARASP recognizes that this draft assessment reflects one of EPA’s preliminary steps to implement the recommendations as laid out by the NAS in 2011. However, as described above, this draft assessment still falls short of adequately implementing those recommendations and does not provide transparent criteria for evaluating the scientific data. NAS specifically recommended Stroup et al. (2000)¹⁰ as a way of organizing a meta-analysis of observational studies and EPA can also draw on the Klimisch¹¹ criteria for data reliability, the Bradford Hill¹² considerations and more recent methods for use in a weight of evidence characterization.

In summary, EPA must provide consistent methods for conducting data evaluation and this information should be transparent and specific for the chemical being reviewed. Specifically, all future toxicological reviews should include: (1) detailed information about the specific frameworks EPA utilizes for synthesizing scientific data and how those frameworks are employed for use in toxicological reviews; (2) the specific criteria employed by the Agency for evaluating study quality; (3) a clear listing of all the exclusion and inclusion criteria used in the literature search along with information about which studies were used in the weight of evidence determination; and (4) data tables that array the scientific information using a mode of action framework.

¹⁰ Stroup, DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, and Thacker SB. (2000). Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA. 283(15): 2008-2012.

¹¹ Klimisch HJ, Andreae M, Tillmann U. (1997) A Systematic Approach for Evaluating the Quality of Experimental Toxicological and Ecotoxicological Data. Regul Toxicol Pharmacol 25: 1-5.

¹² Hill AB (1965). The environment and disease: association or causation? Proc R Soc Med 58: 295-300.