OMB Staff Working Comments on EPA's draft Benzo[a]pyrene (BaP) Toxicological Review (page numbers refer to the draft dated June 2011) **and Draft Charge to External Reviewers**

July 7, 2011

General Science Comments:

- While we appreciate that this assessment is streamlined (only 250 pages excluding appendices), it appears to be missing some critical pieces.
    - For instance, instead of including uncertainty tables, as have been in past assessments, the discussion of uncertainties/limitations are very brief (only 1-2 pages) and sections 5.3 and 5.4.5 are minimal. More robust discussion, as well as tables would be helpful to bolster the discussion. It is not clear why EPA has dropped the more robust discussion and tabular presentation that exists in other previous assessments. This is of particular concern because the uncertainties are so large in this assessment. Similarly, EPA does not provide any discussion of the confidence in the cancer values. Such a discussion, including discussion of uncertainties, would be helpful and informative, particularly since EPA likely wants to use BaP as an index chemical for other PAHs.

    - We didn't see a clear integrative weight of evidence evaluation for each of the derived values. Such a discussion would be helpful. Associated with this it would be helpful to see some standardized evidence tables, as well as concise statements of the criteria used to exclude, include and advance studies for consideration. Chapter 7 of the recent NAS formaldehyde report provides some helpful suggestions that are consistent with our recommendation.

    - Section 5.1.1 does not even mention the chosen principal study and critical effect. Additionally, there is no rationale, justification or weight of evidence evaluation. The section appears to simply describe some of the attributes of the available studies. It would be helpful to provide more details on EPAs systematic approach to reviewing the available studies. A clear listing of strengths and limitations of all the studies, perhaps in a table, would be useful.
        - Page 209, provides the first mention of the Xu study for the RfD. However no rational for the choice of this study/endpoint over other studies is provided. In addition, as EPA has chosen a subchronic study, over available chronic studies, it would be helpful to provide a rationale for this choice as it then requires EPA to apply an additional uncertainty factor. EPA should also present an alternative analysis showing what the final RfD would have been if EPA used a chronic study instead. A charge question taking comment on this would be helpful.

    - Section 6 does not appear to describe the overall confidence level of the quantitative aspects of hazard and dose response. In particular, there is no discussion of the confidence level of the inputs used to inform the final proposed cancer values nor is there discussion of the confidence in the final proposed values.

- To provide context, it would be helpful to provide discussion of what other Federal and international bodies have determined safe levels to be and what others have said about the cancer classification. This type of background information is useful and informative to readers and peer reviewers. For instance, it seems that a recent IARC review determined that BaP was a group 1 Carcinogen, however this appeared to be based on sufficient animal evidence. Similarly, NTP lists BaP as a "reasonably anticipated carcinogen", based predominantly on animal data. EPA however, has determined that there is 'credible human evidence'. It would be helpful for EPA to explain their determination in light of acknowledged values and classifications from other representative organizations.

- Has EPA conducted a 'reality check' and considered if the oral slope factor is consistent with background exposures? Looking at the ingestion values provided in Chapter 3, particularly for toddlers, using the proposed oral slope factor, it seems that at the typical background ingestion rates, a toddlers risk may be in the 10^{-5} risk range solely from diet (and without the application of any adjustment factors for childhood susceptibility.) Is this consistent with available information on cancer incidence in toddlers? It is important for EPA to discuss this information, and provide answers to questions that will arise regarding whether or not EPA is implicitly stating that there currently exists a cancer risk at background exposures, particularly for children.

- EPA determines that BaP is 'carcinogenic to humans' based on evidence of cancer in humans exposed to 'mixtures of PAHs' and extensive and consistent animal evidence. EPA also states that the epidemiological studies provide "credible but limited support for a causative role of BaP in human cancer". We wonder if, as per EPA cancer guidelines, a descriptor of "likely to be carcinogenic to humans" is more appropriate, based on the evidence EPA has presented.
    - We agree with the EPA findings regarding the animal evidence, however it is unclear how human exposure to mixtures of PAHs, containing BaP, provide strong evidence of the carcinogenicity of BaP. How does EPA know that the effects are due to BaP alone and not the other PAHs?
    - It is unclear what is meant by "credible but limited support". It would be helpful for EPA to clarify this. We expect that all studies relied on in the tox review are credible, however it is unclear how exposures in certain industries, where there are exposures to multiple contaminants, including BaP, provides support for a "causative role" for BaP alone. Due to the co-exposures from other contaminants, in all the epidemiological studies, wouldn't it be more appropriate to say that human data on BaP causation is limited and confounded?
    - EPA should provide the details of a causal framework detailing how they believe the human epidemiology provides causation for BaP and human cancer.
    - We note that NTP lists BaP as a "reasonably anticipated carcinogen", based predominantly on animal data.
    - We note that IARC, in a recent review, determined that BaP was a group 1 carcinogen, based on sufficient animal evidence.

  - As per the cancer guidelines (see page 2-54) to use this descriptor, EPA should provide the information suggested by the guidelines: "Exceptionally, this descriptor may be equally appropriate with a lesser weight of epidemiologic evidence that is strengthened by other lines of evidence. It can be used when all of the following conditions are met: (a) there is strong evidence of an association between human exposure and either cancer or the key precursor events of the agent's mode of action but not enough for a causal association, and (b) there is extensive evidence of carcinogenicity in animals, and (c) the mode(s) of carcinogenic action and associated key precursor events have been identified in animals, and (d) there is strong evidence that the key precursor events that precede the cancer response in animals are anticipated to occur in humans and progress to tumors, based on available biological information. In this case, the narrative includes a summary of both the experimental and epidemiologic information on mode of action and also an indication of the relative weight that each source of information carries, e.g., based on human information, based on limited human and extensive animal experiments."

- In section 4.1.3, EPA states that molecular epidemiology studies provide support for the involvement of BaP in human cancers. It was not clear how section 4.1.3.1 supports this. In the studies described, authors used human cancer cells (from cancers likely due to other causes) and then exposed them *in vivo* to BaP. It is not obvious how an increased BaP sensitivity would prove that BaP is involved in the development of cancer. More clarity is needed here.

- In section 4.1.3.2 and subsections, EPA discusses the different cohort studies. Each of these industries had exposures to BaP as well as other PAHs and other contaminants. It would be helpful for each of these sections if EPA more clearly discussed the confounders and co-contaminants that may also be impacting the health risks identified.

- For section 4.3, and other sections that discuss possible endpoints, it would be helpful to have a table which shows the different studies, their specific exposure route (eg gavage, diet, air), the NOAEL/LOAEL values, and the main strengths/weaknesses of the studies. EPA has provided tables like this in the past for each section and it was helpful to readers and reviewers.
  - Throughout this section, and other similar sections, it is difficult to determine if the NOAEL/LOAEL values were determined by the study authors or by EPA. More clarity on this, in each of the cases where such a value is provided would be helpful. This would be particularly helpful in section 5.1.2 as well.

- Section 4.5, is entitled "Mechanistic Data and other studies in support of the MOA". This section seems to discuss only genotoxicity, metabolic pathways, mutagenesis and tumor formation and progression. There is no discussion of MOA to support any of the non-cancer effects such a development reproductive effects or other endpoints. In particular a discussion of MOA as it relates to the endpoints chosen for the RfD and RfC should be added or the section should be renamed.

- For the oral exposure pathway, EPA seems to treat gavage and dietary exposure studies similarly and it is not clear that the science supports such treatment. As many of the effects

seen are at the site of exposure, the gavage exposure may be having a quite different effect than a similar dietary exposure may have. EPA should clearly discuss the impacts of gavage vs. dietary exposure, discuss the implications for relevance to human exposure, and also provide information on the strengths/weaknesses of each. In addition, changes to figures and tables to clarify the specific route of oral exposure will be helpful. More clarity in table 5-2, and discussion in Chapter 5 of exposure pathway effects is also needed.

- In discussing non-cancer and cancer effects, there is much reliance on forestomach lesions and forestomach tumors. As humans do not have a forestomach, it would be helpful to include discussion about the relevance of these effects and tumors to humans. For cancer risks in particular, it seems that these tumors are driving the risk numbers. It would be helpful for EPA to add a discussion and also have a specific charge question, in the cancer section asking reviewers to comment on the applicability and relevance of these tumors, as caused by BaP, to humans.

- Section 4.6.4 and subsections, for each subsection, it would be helpful to have some conclusory sentences about the MOA for these non cancer effects. It seems that the MOA is unknown for each of the endpoints discussed. More clarity on the plausibility of these endpoints actually occurring, at environmental BaP exposure levels would be helpful.

- As BaP was one of the chemicals that had specific data for EPA to rely upon when deriving default values and preparing the 'Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens', more clarity is needed as to why EPA is suggesting that users apply the default adjustment factors, rather than using the data that exist specifically on BaP. EPA cites the specific studies, but it is not clear why EPA is not recommending their use. A specific charge question on this would be useful.

- In section 5.1.2, where the BMD analysis is presented for the possible endpoints, it would be helpful to have some discussion of the biological plausibility for each of the fitted models. In particular a rational for why one may be preferred over another would be useful to inform the determination of what is the best model/endpoint to choose.

- Non-cancer critical effects
    - Section 5.1.3, mentions that decreased ovarian weight is the critical effect. Please also provide some discussion regarding whether or not this is an adverse effect or precursor effect. Is there a typical range of normal ovarian weights in female rats and did the Xu study find weights outside the range of normal? We also suggest having a charge question on EPA's determination (eg, if EPA clarifies that this is an adverse effect, please ask reviewers to comment on this).
    - Section 5.2.1, notes that Archibong was chosen as the critical study as it showed biologically significant effects at the lowest dose tested. EPA notes that they RFC is based on fetal survival and body weight decreases. Are the body weight decreases considered adverse? Are the changes at 75 ug/m^3 statistically significant from the carbon black control, or just the unexposed control? It would be helpful to clarify and discuss this. As no changes were seen at the 25 ug/m^3 dose, it is not clear why EPA is stating that body weight changes are the critical effect. Shouldn't 25 ug/m^3 be a

NOAEL, rather than a LOAEL if this endpoint were used? Please have a very specific charge question for reviewers on EPAs choice of endpoint (at this point it is not clear to us why both are used when effects are seen at different levels-perhaps the fact sheet we were provided is leading to this confusion). We would also suggest clarifying that in reality, EPA is only relying on decreased pup survival.

- For the decreased survival in section 5.2.2, EPA notes that there were heterogenous variances in the data. It is not clear where this was discussed previously and what implications it may have for the choice of study/endpoint. We suggest having more discussion of this as it is the chosen EPA endpoint.
- In section 5.2.3, it may be helpful to provide information on what the UF and final RfC would have been had EPA used 25 ug/m^3 as a NOAEL for decreased weight rather than a LOAEL for survival.

- Uncertainty factors and confidence in the values.
  - For the RfD, EPA applies UF= 3000. According to the description provided, it is not clear why a full 10x factor for database uncertainty is not used instead of a 3x factor. EPA appears to be missing a robust evaluation of all the available data to support the use of a 3x factor. If EPA used 10x, then the cumulative UF's would be above the threshold for which EPA typically derives RfDs. Thus we question whether or not an RfD should be derived. If derived, we suggest that the confidence be called "low" not low-medium as EPA suggests in Section 6. It is not clear what would make this anything but low since there are uncertainties in 4 areas and the maximal UF value is applied.
  - For the RfC, EPA applies UF= 1000. According to the description provided, it is not clear why a full 10x factor for database uncertainty is not used instead of a 3x factor. EPA appears to be missing a robust evaluation of all the available data to support the use of a 3x factor. Additionally we suggest that the confidence be called "low" not low-medium as EPA suggests in Section 6. It is not clear what would make this anything but low since there are uncertainties in 4 areas and three orders of magnitude of UF's are applied.
  - As EPA does with the non-cancer values, it would be useful, in Section 6 to have some discussion of the confidence of the cancer values. Considering the limitations in the available studies, and the assumptions needed to extrapolate to humans, it appears that the confidence in these values should be low.
  - It would be helpful to have a clear charge question asking reviewers to comment on each of the confidence ratings for each of the values proposed. This is particularly important since the BaP value may be used as an index for other PAHs. Reviewers should also be made aware of this potential future use of the proposed values.
  - According to IRIStrack, EPA began this assessment in FY04. Considering the length of time EPA has been reviewing the data, it is unfortunate that the uncertainty factors applied are still so high and the confidence in the values are so low. We would hope that EPA could identify data gaps more quickly and thus help inform further research that is needed to strengthen the scientific database.

- For the dermal slope factor, EPA relies on mouse data as the mouse was the most sensitive animal in the available studies. It would be helpful for EPA to provide some discussion of

whether or not the mouse is the best biological model for dermal absorption compared to rabbits, rats and guinea pigs. It was our understanding that for dermal absorption EPA should compare each of the animal models to what is known regarding human absorption (perhaps some in vitro skin models could help with this). Discussion of the biological support for each animal model, rather than simply choosing the most sensitive model, would help to strengthen the scientific basis for the derivation.

  - Page 229, briefly mentions that EPA evaluated the strengths and limitations of the data sets, however it is not clear where the findings of the evaluation are presented.

- We recognize this is a very important assessment and the implications of a new RfC, RfD, IUR, and Oral Slope Factor and new Dermal Slope factor will be broad and far-reaching with potentially large regulatory implications. In addition BaP is often used as the index chemical for a large family of PAHs. Therefore, in light of concerns and implications raised, we suggest that EPA consider an SAB or NAS review. The review panel should include broad and diverse expertise to encourage vigorous scientific discussion and debate. In addition, the charge should include questions about the confidence in each of the derived values and whether they are sufficiently robust to be used as an index chemical for the family of PAHs.

Editorial Comments (with Scientific Impacts):

- Page 18, line 4-8, it is unclear how the results imply that BaP was completely absorbed from the GI. For the sentence on line 6-8, please provide the results of the studies mentioned.

- Page 20, line 19-20, it was not clear where EPA actually discussed and presented the findings showing the presence of DNA adducts in humans.

- Page 22 , line 36, please note that the type of contaminated environmental material also matters; not all contaminants are equal.

- Page 43, line 7-8. Please clarify if these respiratory tumors are in humans or animals.

- Page 64, line 16, Please clarify which components exceed BaP toxicity.

- Page 77, line 13-18, it is not clear how an association of higher intake of BaP and cancer leads to a conclusion that BaP plays a role in the etiology of these tumors. As there were so many other co-exposures, how were confounders considered? For the final sentence here, it would be useful to clarify if any statistical significant differences were found.

- Page 78, line 36-38, it is interesting that the exposure levels here was so high (3400 ug/m3) yet no dose response was seen and exposures were not statistically significant from controls. How does EPA incorporate studies like these into a weight of evidence evaluation?

- Page 81, line 2, please clarify if the risks were statistically significant.

- Table 4.8, it would improve readability if EPA had clear separate sections in the table for studies that were done by gavage vs dietary ingestion. While both are oral, the specific route can have large impacts on effects.

- Page 106, please include the Clement 1990 analysis in an appendix. It is not clear from the references that it is publicly available. Also, has the calculation and approach used been independently peer reviewed? Suggest adding a specific charge question on this as it is important input for the derivation of the IUR.

- Page 116, EPA relies on Sivak 1997, for the dermal slope factor. However the discussion of this study is very short (11 lines). We suggest adding a fuller discussion of this critical study, including strengths and limitations, to section 4.

- Page 139, line 35-36, please provide a citation for this statement.

- Table 4-27, for clarity and readability, we suggest breaking this into 4 tables. One table for Subchronic and gavage, one table for subchronic and dietary exposure, one table for chronic and gavage, one table for chronic and dietary exposure. It is not clear why EPA mixes both the routes and duration in one table. It is harder for the reader to visually look at the data and do comparisons.

- Figure 4-2, please separate the gavage and dietary routes of exposure into two different tables. It is not clear why EPA is treating them similarly.

- Page 179, table 4-28, notes that the Archibong 2002 study and Wu 2003 study both used carbon black as a carrier. We did not see where the potential impacts of this were discussed. Was carbon black also used in the zero dose group? How were the potential impacts of carbon black exposure controlled for? Since Archibong 2002 is used to support the RfC, suggest having a specific charge question for reviewers to ensure that any potential impacts of the carrier exposure are addressed directly.

- Section 4.8, for each of the subsections, it would be useful to have some concluding sentences about whether or not the data support a concern for the specific susceptibility and if yes, the extent of the concern. In addition, the relevance of the susceptibility to the RfD, RfC and cancer values should also be clarified. It is unclear if the approach EPA has taken is one that leads to protection of sensitive populations, or if EPA thinks that there are populations that may be at risk at the levels derived. We also suggest adding a charge question regarding the applicability and utility of this section.

- Page 205, line 25, and elsewhere, please clarify (perhaps in a footnote) that this document is just a draft and does not represent official agency guidance.

- Page 219, line 2-5, please use a direct quote from the cancer guidelines. It is also important to note that support for this approach is only in the appendices of the cancer guidelines and it is clearly noted that it is a default approach. When data exist to inform the appropriateness of combining benign and malignant tumors for a specific tumor type, this information should be used over a default. Page 219 should clearly note that EPA is invoking a health protective default. Page A-5 of the cancer guidelines state: "*The default is to include benign tumors*

observed in animal studies in the assessment of animal tumor incidence, if such tumors have the capacity to progress to the malignancies with which they are associated."

- Page 220, footnote b, hasn't US EPA 1988 been updated? Page 222, refers to US EPA 1992- a document we could not find in the references. Is this the same as US EPA 1988?

- Page 221, line 1-6, please use a direct quote from the cancer guidelines. It is not clear that the interpretation provided here is entirely consistent with the language in the guidelines.

- Page 224, line 17, what is US EPA 2010? This is not in the appendix. Without knowing what EPA is referring to it is difficult to comment. However this seems like a new approach and we suggest adding a charge question to specifically take comment on what EPA has done. This document is also mentioned on page 227.

- Page 241, section 5.4.5.3, shouldn't this also discuss some of the assumptions mentioned in earlier text?

- Page 242, line 22, EPA states that human evidence was previously inadequate. It would be helpful if EPA could clarify what human evidence EPA now finds to be adequate and compelling. We are only aware of studies where exposure to BaP was confounded by the presence of other environmental contaminants.

Comments on the Draft Charge:

[Note: some suggestions for charge questions are provided in comments in the above sections. Many of those comments have not been reiterated here, but should be considered as equally important.]

- In the introductory section of the charge, it would be helpful to make peer reviewers aware of two important things:
  1) the relationship between the proposed values and background exposures in those cases where the proposed value suggests a risk at background exposures. We think this may be particularly important for the oral cancer value.
  2) whether or not any or all of these values will be used as an index value for PAHs in other contexts. If EPA does plan such use, as mentioned before, asking reviewers to specifically comment on the confidence of the values and their appropriateness for use will be of critical importance to the review.

- General Questions 2: It is unclear how reviewers will be able to tell if additional studies "would have a significant impact on the conclusions." Suggest reframing this to simply ask about relevant studies and then EPA can conduct further evaluation to determine if the studies will have a significant impact.

- A2, please clarify if the critical affect is adverse and ask reviewers to comment on this.

- Section A: It seems that EPA calculated time weighted average doses in order to use the Xu study. Suggest adding a charge question for reviewers to comment on this derivation. Please also add a question taking comment on EPAs confidence determination for the RfD.

- B2: it seems EPA only relied on fetal survival for the RfC. Please clarify this in the charge question.

- Section B: Please add a specific question on adjustments made to derive the RfC. This would include taking comment on the approach used to account for discontinuous daily exposure, the use of the RDDR, and the use of the multi-path particle dosimetry model. Please also add a question taking comment on EPAs confidence determination for the RfC.

- C1: please add a specific question regarding EPAs treatment of co-exposures in the discussion and determinations regarding the weight of evidence given to the human epidemiology in EPAs determination that BAP is 'carcinogenic to humans'.

- C2: it may be helpful to clarify that you are looking for a response for all the 3 different exposure pathways. Reviewers may have different answers for different routes.

- Section C: additional suggested questions for peer reviewers:
    - Please add a question about the application of the child specific adjustment factors
    - Please add a question regarding EPAs confidence in each of the cancer values and take comment on whether or not the values should be used as an index value for other PAHs.
    - Each of the specific questions presumes that EPA was correct to calculate an IUR, or oral/dermal slope factor. Suggest taking comment on the suggestion that perhaps the database is not robust enough to support generating each value. EPA should leave the door open to the reviewers to suggest that the available data do not support deriving these values due to uncertainties and limitations. It is not clear that this option is reflected in any of the questions. A similar question would be helpful for the RfD and RfC as well, although we think this is most important in the context of the dermal slope factor as that derivation is most novel and relies heavily on assumptions. In the case of the dermal slope factor, it would also be helpful to take comment on all the assumptions underlying the analysis.

- For the IUR, please also take comment on the Clement analysis which calculated the average lifetime daily dose. It would be helpful to show this analysis in an appendix.

- For the dermal slope factor, please take comment on EPAs proposed approach for how and when (<20ug/day exposure) the value should be used.