

STEP 3 IRIS INTERAGENCY COMMENTS

EPA have a specific charge question on the cancer descriptor EPA chooses, and on applying the supplemental guidance to all the compounds for which a non-zero RPF is created.

- In Section 8, page xii notes that “a description of uncertainties and limitations is crucial to interpretation of the RPF approach...” The discussion of uncertainties in this section (beginning on page 197) could be enhanced by a specific discussion of limitations as well. EPA may want to consider changing the title of Section 8 to be “uncertainties and limitations”.
- The provided relative confidence rating for each derived RPF is very helpful and we appreciate EPA including this information, and we believe it would be useful if EPA could provide some suggested guidance on how this rating, as well as other ratings, should be used by the risk assessor. For instance, is there a suggested mathematical approach that would give less weight to a RPF that is characterized as “very low” confidence? For example, the RPF for dibenz[a,c]anthracene is “very low”, yet the RPF value is the 5th highest of all the RPF’s presented. Similarly the PAH with the highest RPF (60) has a ‘low’ confidence rating. Thus it is possible that at a specific site, a PAH, which has very low or low confidence, could be the quantitative driver in a risk analysis. We also suggest that EPA add a specific charge question on qualitative and quantitative approaches that could be used to incorporate these confidence ratings into the risk analysis.

Specific Science Comments:

- Page v, and elsewhere, EPA states that biomarker studies measuring DNA adducts in humans were excluded. Please confirm whether this statement means that there are no dose-response information relating biomarkers of adducts to cancer endpoints.
- EPA states, page vi and elsewhere, that only studies with positive findings were used for the calculation of RPFs, while negative findings were used only in the weight of evidence (WOE) evaluation of potential carcinogenicity. It would be helpful for EPA to better elaborate on the rationale for not including negative studies in the RPF calculation even if they otherwise meet the appropriate quality criteria (or conversely, please clarify if none of these studies met the qualifying criteria). If it is a methodological limitation (e.g. how to consider or weight a zero value study), it would be helpful for EPA to have an explicit charge question that asks about possible approaches to using this information. We suggest that EPA also have a specific charge question that asks the expert panel if they agree with the approach of not using negative findings (from studies which meet the quality criteria) in the calculation of the RPF.
- Page vii, and elsewhere, EPA states that for data sets that included only a single dose or for those data sets for which no model fit was achieved, a point estimate RPF was calculated. We suggest EPA add a specific charge question that takes comment on this approach in these limited data cases. A discussion of this issue, and the impact it has on the RPFs (eg a statement of how many of the 23 values, and which ones are based on a point estimate) would be likely be very helpful to the peer reviewers. In addition, it would be helpful to ask the expert reviewers about how this should, if it all, impact the relative confidence rating given to the chemical.

STEP 3 IRIS INTERAGENCY COMMENTS

- Page xii, the 5th and 6th columns in table 2 appear to be not consistent with the presentation in Table 1.
- Page 27 notes that the data are sufficient to classify a number of PAHs as “possibly carcinogenic to humans,” but for other PAHs the data are negative or equivocal. We believe it would be useful if EPA were to determine the carcinogenicity potential (using the weight of evidence descriptors in the 2005 Cancer Guidelines) for each of the PAHs for which an RPF is determined. If this is not possible, a good starting point would be clarification of which are “possibly carcinogenic” and which have negative or equivocal data. How this relates to the RPF value would also be helpful information.
- Page 28: should *in vivo* carcinogenicity assays be added to this bulleted list of study endpoints evaluated?
- Page 28, states that BaP is one of the most potent carcinogenic PAHs identified. The RPF approach identifies 5 PAHs with values >1, implying that they are more potent than BaP. After calculating the RPFs, has EPA gone back through the existing available scientific literature to affirm that in fact these 5 compounds are expected to be more potent than BaP (and in some cases more than an order of magnitude more potent)? We believe this would be helpful to validate the RPF approach.
- Page 31, line 9-18, please provide citations for each of the bulleted key events presented (similar to the citations provided on page 30).
- Page 33, line 4, are these data for PAHs or BaP? Please clarify.
- Page 39, the unlabelled figure provided above line 2 seems quite general. Is there a more detailed figure that shows the complexity of proteins binding the AHR (rather than just ARNT), as well as a more detailed list of mRNA’s up-regulated that could be provided and discussed? We also note that it would be helpful to provide citations for the sentences ending on lines 8 and 13.
- Page 43, lines 14-17, notes that recent studies have used QSAR to evaluate PAH structural features and mechanistic events related to carcinogenesis. Has EPA considered comparing the QSAR data and predictions to the results of the RPF approach for the 24 PAHs for which a RPF is provided? This type of cross-walk may provide useful validation of the RPF approach.
- Page 46, line 9, cites Schneider 2002 and the suggestion that RPF estimates be derived separately for different exposure pathways. Has EPA considered such an approach? If so, does EPA have a general idea of what impact would this have on the derived RPF?
- Page 51, table 3-1 provides a very helpful comparison of RPF values derived by other agencies and also existing in the published literature. We believe it would be extremely

STEP 3 IRIS INTERAGENCY COMMENTS

helpful and would facilitate understanding of the document if EPA added a column showing the values derived in the current approach.

- Page 58, line 4-10, provides a comparison of TEFs and RPFs. However, at this point in the document still may be unclear to readers how the RPFs should be applied (eg to what exposure routes and to what durations). More clarity here would be helpful.
- Page 60, line 5, states that the RPF method produces risk estimates that are significantly lower than those derived from epidemiology studies. Could a comparison be done with the epidemiology studies to determine if the relative rankings in RPF values are similar to the relative rankings of risk determined from the epidemiology data? If so, we believe such a comparison would be useful and informative.
- Page 61, line 9-11, provides a list of studies that were excluded from the database. We suggest EPA take specific comment on whether or not expert reviewers agree with the determination to exclude these studies. EPA could add this question to the charge section which addresses chapter 4.
- Page 89, line 5, EPA refers to non-positive bioassays. Please clarify if these means the BaP results were negative or the PAH test chemical results were negative. Are both options plausible?
- Page 108, we believe it would be helpful if EPA provided some discussion of the types of tumors that were included in the quantitation. It appears that EPA treated benign and malignant tumors the same. We suggest adding a charge question for reviewers on this aspect of EPA's approach. Perhaps a different approach would have EPA treating the tumors differently depending upon the specific type of tumor, its likelihood of progression to malignancy, and/or its relevance to humans (for some specific tumors expert reviewers have commented that the animal data are not relevant). We suggest having a charge question which asks expert reviewers to comment on such alternative approaches. These questions could be added to the section of the charge which discusses chapter 5, and this treatment of the data should be discussed in the uncertainty section in chapter 8.
- Page 108, line 34-38, discusses how EPA sequentially eliminated high-dose data groups in an effort to achieve adequate model fit. EPA states that these high-dose groups are not as informative and can be eliminated, with the justification that the lower dose is the region of interest. Did EPA develop an understanding of any data anomalies and a sound theoretical basis for either including or excluding them before looking at the level of the dose of the groups? We believe this step should have preceded—not followed—model fitting. If the model is fit first and then adjustments are made, the practice may be subject to the criticism, and we don't believe from the write-up that EPA has provided a robust rationale for dropping data at the highest doses, especially for the explicit purpose of improving the model fit. In practice, we believe it may be difficult to predict which levels may be useful from a policy perspective; therefore, the full dose-response relationship may be helpful for the risk manager who may need information across a wider range of dosages than just near the BMR. We suggest that EPA add an explicit charge question addressing this methodological issue.

STEP 3 IRIS INTERAGENCY COMMENTS

Also, to inform the peer review, EPA could, perhaps in an appendix, provide information on the number of PAHs for which an RPF was determined based on dropping high dose data. As a comparison, EPA could also provide information using a point-estimate approach, with a charge question on the usefulness of this approach. Peer reviewers could also be asked to comment on the confidence rating of such values, as compared to values derived from a BMD approach.

- Page 110, line 1-6, EPA should clarify the rationale for why they find the BMD to be more stable than the BMDL, in order to facilitate peer review of the rationale.
- Page 113, line 1-4 mentions that in several studies the test conditions were not necessarily optimal for the selected PAHs. It may be useful to also discuss this in Section 8. It may also be helpful if EPA clarified how the test conditions impacted, if at all, the confidence ratings that were derived.
- Page 113, line 28, EPA states that a point estimate approach using the peak response for the selected PAH was used. We suggest that a fuller discussion of the impacts of this approach may be helpful, and a charge question on this aspect of the approach may also be helpful.
- Page 114, line 18 states that PAHs with adequate evidence to suggest that they have little or no carcinogenic potential were assigned an RPF of 0. This suggests that the higher the RPF the greater the potential carcinogenicity. More discussion on how the RPF rating equates with the cancer weight of evidence descriptors in the cancer guidelines would be helpful. For example, does an RPF of zero mean that the compound has been determined to be “not likely to be carcinogenic”? For those which EPA eliminated from consideration, would the rating be something like “Inadequate Information to Assess Carcinogenic Potential”?
- Page 115, line 22, EPA states that when the data were considered adequate for a given PAH, it was selected for inclusion in the RPF approach. Further information on how EPA determined adequacy, and the criteria used, would be helpful.
- Page 116, EPA’s weight of evidence approach looks for positive results in tumor bioassays but does not appear to consider the type of tumor, the location of the tumor or the potential relevance of the tumor to humans. If this is correct, discussion of this aspect of EPA’s approach and its potential impacts should be included in section 8 on uncertainties. EPA should also consider a charge question on this aspect of the approach.
- Page 118, it seems that EPA’s weight of evidence evaluation (figure 6-1), is more of an examination of data sufficiency, rather than an evaluation of potential carcinogenicity. The evaluation seems to determine whether or not the PAH has a data set which allows the PAH to be included for evaluation, yet the evaluation does not determine what the relative ranking should be. More clarity here, and perhaps a possible name change for the evaluation may be helpful. As figures 6-2 through 6-35 do provide helpful graphic arrays, which include presenting the RPF values as part of the weight of evidence evaluation, it might be helpful to have figure 6-1 mention the determination of the RPF as part of the evaluation process. More

STEP 3 IRIS INTERAGENCY COMMENTS

clarity on where the weight of evidence approach comes into play in determining potential carcinogenicity would be helpful.

- Page 189, we believe it would be very helpful to have a summary table which presents the inputs to the RPF derivation for each compound. For instance, for each PAH, there could be columns which provide data on: # of positive bioassays, # of negative bioassays, # of positive studies on other endpoints, # of negative studies on other endpoints. Such a table would allow a summary view of the strength of the dataset for each compound. This would also allow for comparison among the 26 compounds.
- Page 190, because for most of the PAHs (17/23) there were 3 or less RPFs derived, instead of presenting the range of RPFs, did EPA consider simply presenting all the RPF values in the table? This would be very helpful as it would make all the data available and would let users and reviewers understand the full extent of how variable the RPF values may be for certain compounds.
- Page 190, line 17-19, mentions that calculating a weighted average might increase uncertainty. It would be helpful if EPA could explain why this might be the case, as we could imagine a method of weighting the studies with more relevant exposure routes or perhaps higher quality data sets could be a promising approach.
- Page 191, line 3-5 discusses how when separate RPFs were calculated for different endpoints in the same group of animals, the higher value of the two RPFs was included in the calculation and the lower RPF was dropped. It would be useful to include discussion of why this approach was taken and its impacts on the final derived RPFs. For example, it seems that EPA could have included this lower RPF value in the range of RPF values derived for the PAH.
- Page 191, in the discussion of confidence ratings, we suggest EPA more clearly present which criteria and inputs led to which confidence ratings. A table would be useful here. For instance, it is not clear how the number of RPFs based on BMD modeling played into the confidence rating, nor is it clear how the number of datasets helped determine the ranking. For instance, Fluoranthene has 6 datasets, 5 of which had RPFs based on BMD modeling. This is the highest number of BMD values for any compound, yet the confidence rating is low. For example, this perhaps could have been driven by the exposure route, or a hierarchy that EPA applied (eg exposure route then number of studies, etc) when deriving confidence ratings. More clarity on this would be very useful.
- Page 197, line 34-37, EPA focuses discussion of improvements on the comprehensive review that was conducted. While this is very important, are there also methodological improvements in the derivations of the RPFs that EPA could also discuss?
- Page 198, in discussing uncertainties related to the dose-response, as noted above, this may be a good section in which to include discussion of EPA's approach of treating all tumors and all exposure routes equally. Similarly, EPA treated RPFs derived from point estimates and

STEP 3 IRIS INTERAGENCY COMMENTS

BMD modeling equal as well. It would also be helpful, in this section, to have discussion of which decision points may lead to overestimates of potency and which may lead to underestimates of potency. Page 202 mentions some of these aspects in bullets, but a discussion of impacts would be useful.

- Page 201, line 30-31, talks about situations where bioassays that did not include BaP data were considered. However, from previous text in the document, it appeared that only studies that also included simultaneous BaP data were included for evaluation. More clarification here would be helpful.
- Page 204, line 9, discussion of why EPA finds that providing order of magnitude estimates was not considered superior to providing simple means would be helpful.
- Page 205 provides a very helpful table of the PAHs and their confidence ratings. See comments above on this issue.
- Page 212, line 18-21, EPA makes mention of the California RPF scheme which uses a hierarchy of bioassay data based on relevance of exposure route. Other than the preference for exposure routes that target the respiratory track, EPA states that the basis for prioritizing is not evident. It would be helpful if EPA could elaborate on why the agency did not employ a similar approach to the California RPF method.
- Page 213, line 28, EPA states that the available information provides “some support for cross-route extrapolation.” It is not clear what is meant by “some support.” This is an important issue as it provides the basis for drawing data from multiple exposure routes. We suggest adding a charge question on this issue.

Comments on the draft Charge:

(To be helpful, we have included some of the specific charge questions that were interspersed in the comments above to assist with consolidation)

- In many cases, EPA asks if the scientific justification or criteria for a choice is appropriately described. Although it may be implied, EPA does not appear to ask the reviewers whether or not they are in agreement with the scientific justification provided. While it is important that the reviewers comment on the transparency and clarity of the guidance, it would likely also be useful to explicitly ask that the reviewers comment on the scientific justifications presented. Below we suggest specific clarifications of a few questions; however, if EPA agrees with this general suggestion this change would apply to many of the questions not mentioned below.
- Since the development of Agency Information Quality (IQ) guidelines required by statute, many agencies have been using charge language that tracks with the standards of their own IQ guidelines. For example, such language often focuses on whether or not the information in question is accurate, clear, complete, transparently and objectively described, and scientifically justified. We believe it may be useful for

STEP 3 IRIS INTERAGENCY COMMENTS

EPA to follow a similar approach and incorporate some of the language from your IQ guidelines into the formulation of the charge questions.

- Chapter 1: Question 1 asks if the report provides adequate context for the use of the RPF approach. It is unclear what is meant by this question. This could be strengthened by asking explicitly how RPFs should be used in assessing human health risks and if there are any limitations in their use.
- Chapter 2:
 - In Question 4 suggest asking if peer reviewers agree with the choice of BaP as the index compound.
- Chapter 4:
 - In Question 7 suggest asking reviewers if they agree with the criteria used to select the studies for inclusion in the assessment.
 - In Question 9 please clarify which tables are being referred to.
- Chapter 5:
 - Suggest taking specific comment on the use of the multistage model as well as the determination to use the linear model for all datasets for which multi-dose continuous data was available.
 - As mentioned above, suggest taking comment on the decision to sequentially eliminate high dose groups to get the models to provide adequate fit.
 - Suggest taking comment on EPA's approach for choosing the BMR based on using the observed response in BaP adjusted for background response.
 - Section 5.4 discusses how a point estimate approach was used for data where only a single dose of BaP and other PAHs was available. We suggest that EPA take comment on whether these single dose studies are robust enough to be used, and if so, how the single dose nature of the studies should be taken into account in the confidence rating and averaging approach used to derive the RPFs.
 - Section 5.4 also mentions that statistical analysis was not always possible for each dose group and in these cases EPA chose the lowest dose that produced a near maximal change in the assay (the highest dose in the linear portion of the dose-response curve as identified visually). We suggest that EPA take comment on this aspect of the methodology. It may also be helpful to take comment on how these studies should be taken into account in the confidence rating and averaging approach used to derive the RPFs.
 - Suggest taking comment on the inclusion of RPF calculations from the newborn mouse assays. In particular, EPA may also want to ask for specific comments on its approach to handling gender and tumor differences in the newborn mouse studies. We also note that section 5.6 is a bit unclear on how the data were treated in the end. Did EPA average the RPF values derived separately from the male and female mice, consistent with EPA's approach of averaging RPFs?
- Chapter 6

STEP 3 IRIS INTERAGENCY COMMENTS

- We suggest EPA should specifically ask reviewers to comment on how EPA implemented the decision tree approach. For instance, do all the compounds included for RPFs belong and were the other compounds appropriately excluded?
- Chapter 7
 - In Question 21, please clarify that the arithmetic mean was used. It may also be helpful to take comment on other approaches in addition to the presentation of a range or confidence interval (for instance our comments above suggest that it may be just as easy to present all the data in some cases).
 - As mentioned above, suggest taking specific comment on EPA's treatment of negative data when determining the RPF.
 - It appears that EPA, when averaging RPFs does not take into account route of exposure in the study. Suggest taking comment on this approach to get suggestions on how such information may or may not be useful.
 - Suggest taking specific comment on the EPA approach, as described on page 191, where EPA uses only the higher value of multiple RPFs from different endpoints in the same group of animals.
- Chapter 8
 - In addition to asking if the most important uncertainties are identified, it may be helpful to ask the reviewers to comment on whether the most important limitations are identified.
 - Suggest taking general comment on the fact that most PAHs evaluated included only 1-3 useable datasets and whether or not this should impact how the data are used.
 - Suggest taking specific comment on options for prioritizing RPFs (rather than averaging). Page 202, in the uncertainty discussion, states that the current state of knowledge does not support a biological basis for prioritizing RPFs. It may be useful to get reviewers comments on this finding.
 - Suggest taking comment on EPA's approach to providing a RPF for 1 PAH based on cancer-related endpoints only. It may also be helpful to ask peer reviewers to comment on how this value, which has a low confidence rating, should be used and considered by risk assessors and risk managers.