OMB Staff Working Comments on EPA's Response to "Health Risks from Dioxin and Related Compounds Evaluation of the EPA Reassessment" Published by the National Research Council (NRC) of the National Academies (NAS), [dated January 10, 2010] and **Draft Charge to External Reviewers** [dated March, 2010].

April 22, 2010

General Science Comments:

- So as to not confuse readers, EPA may want to consider revising the name of this document. NAS provided many recommendations to EPA and this document is not a full response, but is only a response to comments regarding three key areas (we note that on page xxxiii, line 16, EPA states that this document only addresses issues related to dose-response. Clarification of this in the title would thus be helpful). For instance, EPA does not address recommendations related to assessing human exposures. In addition, EPA does much more than simply respond to the NAS comments, as EPA provides detailed discussion and analysis of new studies, published since 2003 (see Sections 3 and 6). As these new analyses will likely be of great interest to EPA's stakeholder community, a change in title to reflect these new evaluations would be helpful.
    - The current title does not accurately reflect all the work EPA has put into updating the dioxin reassessment. A suggested title might be along the lines of: "2010 Updated Dose Response Analysis of Dioxin and Related Compounds" A subtitle could say: "Including a Response to key NRC Recommendations Relating to These Analyses."

- We would like to thank EPA to deciding to conduct this interagency review under the IRIS process, as that serves to clarify for the public how EPA intends to proceed with the review. As this document is somewhat unique, and includes much new information and analysis that will be undergoing peer review, it may be helpful for EPA to clarify for the public that EPA equates this document as being in the new IRIS process at Step 3. As the new IRIS process is becoming well-known to the public, this would likely facilitate their participation in the document review.

- One of the important NAS comments, related to the cancer risk, a topic discussed in the response to comments, has to do with the classification of Dioxin Like Compounds (DLC). NAS stated (page 4): "The committee agrees with EPA in classifying other dioxins and DLCs as "likely to be carcinogenic to humans." However, because mixtures of DLCs may also contain dioxins, including TCDD, EPA should reconsider its classification of such mixtures as "likely to be carcinogenic to humans" if it continues to classify TCDD as "carcinogenic to humans."" We did not see where this comment was mentioned and discussed in the EPA response. Discussion of this may be helpful (for instance on page xlv where EPA discusses the weight of evidence statement for carcinogenicity, and elsewhere throughout the document).

- In the 3rd paragraph of the charge and elsewhere, EPA states that NRC encouraged EPA to calculate an RfD. While this is true, it is important to note that NAS encouraged this in the context of providing appropriate margin of exposure (MOE) information. NAS stated on page 187 & 197, in the conclusions and recommendations section of the review of the risk

characterization: "The committee encourages EPA to calculate RfDs as part of its effort to develop appropriate margins of exposure for different end points and risk scenarios, including the proportions of the general population and of any identified groups that might be at increased risk." In addition, on page 180, NAS discusses concerns with the EPA use of the ED_{01} and appears more interested in having EPA use traditional point of departure measures (such as the NOAEL, LOAEL or BMD). NAS then continues to stress the importance of the MOE (page 181: "Because the exposures of a proportion of the U.S. population would be above any RfD, it would have been useful for EPA to define the nature and magnitude of the risks at different levels of intake, the groups of the population most at risk, and the major sources of exposure for any at risk groups. Alternatively, if MOEs were calculated for noncancer effects, then the risk characterization should describe the nature of the adverse effect and the uncertainties and variability inherent in both the BMD (ED) estimate and the relevant exposure estimate. It would have been useful if MOE values had been calculated and discussed for different exposure scenarios.").

  - Considering this recommendation, we were disappointed to see that EPA has not calculated MOEs as NAS suggested. While following the NAS recommendation is a preferred approach, at a minimum EPA should add language discussing this context for the NAS recommendation and provide a response. It may also be helpful to have a charge question addressing this particular NAS recommendation.

- As a point of departure (POD) for non-cancer effects, EPA uses a LOAEL which corresponds to thyroid stimulating hormone (TSH) values above 5 uU/ml (page xliv and elsewhere). It would be helpful for EPA to provide further discussion as to whether or not this is an adverse effect or a precursor effect. In Section 4, EPA mentions that the 5uU/ml standard "was established by the World Health Organization (WHO, 1994) as an indicator of potential iodine deficiency (and potential thyroid problems) in neonates." What exactly did WHO intend when saying it was an indicator of potential thyroid problems? What is the rationale of using this as a LOAEL? As WHO mentions iodine deficiency, an evaluation of iodine sufficiency in the control and exposed populations would be extremely helpful.

- In flow charts and text throughout the document describing EPAs inclusion criteria for the epidemiologic studies and the animal bioassays, criteria include for epidemiology studies "dose-response is apparent between TCDD and adverse health effects" "with some suggestion of an exposure-response relationship" and for animal studies "magnitude of animal response is outside range of normal variability." Both these criteria lead to the exclusion of well-designed studies which did not show adverse effects at the dose levels evaluated (which could include the low dose range EPA is interested in). On page 6-15, when discussing quantitative uncertainty analysis, EPA states: "If the analysis is restricted to experiments showing a positive response, the results will be biased."
  - While we understand the problems of quantitatively evaluating studies which have no dose-response, how does EPA's weight of evidence approach consider these studies which may provide informative information in the low-dose range? Are the studies simply excluded? It may be helpful to categorize these studies and perhaps present them in a table such that when discussing the weight of evidence for cancer and non-cancer effects, EPA could include discussion of the well-designed studies (such as the Eskanzi study, for example) which did not show adverse effects when certain key

endpoints were evaluated. A charge question on EPA's approach to handling these studies may be helpful.

- Page 4-2 states: "In this document, EPA has developed a strategy for identifying the noncancer data sets and PODs that represent the most sensitive and biologically relevant endpoints for derivation of an RfD for TCDD." The layout of the response to comments document discusses the selections of animal and epidemiology data showing the lowest TCDD doses associated with noncancer effects, then discusses kinetic modeling and estimation of human equivalent doses and then in Section 4 EPA presents the dose response modeling and then derives the RfD (as described on page 4-1). It may be helpful for EPA to have a section discussing the mode of action and scientific plausibility, as well as biological relevance of the potential non-cancer effects. This aspect does not appear to be part of EPAs inclusion criteria and thus an explicit discussion connecting the potential PODs to the potential RfD derivations (including scientific plausibility) may be useful. Page 4-3 cites the NAS stating: "The committee recommends that the Reassessment use levels of change that represent clinical adverse effects to define the BMR level for noncancer continuous end points as the basis for an appropriate POD in the assessment of noncancer effects."
    - Based on the NAS recommendation, it is not clear that the EPA response is rigorous in discussing the clinical adversity of the individual adverse effects that have been modeled. If EPA is unable to determine levels of change that represent adversity, this limitation should be brought forward for those endpoints. It may be helpful to have a charge question addressing this issue.

- Page 5-57, line 3-6, EPA states: "The linear approach is used if the mode of action is not understood (U.S. EPA, 2005)". EPA relies on this, and other information, to support linear low dose extrapolation.This seems like an oversimplification as the Cancer Guidelines state (see page 1-15 and elsewhere): "Where alternative approaches have significant biological support, and no scientific consensus favors a single approach, an assessment may present results using alternative approaches."
    - Page 3-23 of the EPA Cancer Guidelines further states "Nonlinear extrapolation having a significant biological support may be presented in addition to a linear approach when the available data and a weight of evidence evaluation support a nonlinear approach, but the data are not strong enough to ascertain the mode of action applying the Agency's mode of action framework."
    - In the introduction of the Cancer Guidelines EPA states, at page 1-8: "When there are alternative procedures having significant biological support, the Agency encourages assessments to be performed using these alternative procedures, if feasible, in order to shed light on the uncertainties in the assessment, recognizing that the Agency may decide to give greater weight to one set of procedures than another in a specific assessment or management decision.
    - Considering this, it appears that the Cancer Guidelines do not require a full understanding of mode of action to support a nonlinear approach. In light of the NAS evaluation and their recommendations for a nonlinear approach, it would seem that in this case, the nonlinear approach has significant biological support and thus it may make sense to present results using both approaches. As the health assessments are risk analyses (or really hazard assessments) it would seem that EPA should be presenting both approaches and then the risk managers can make the decision

regarding which approach may be most appropriate for their intended use as the EPA Cancer Guidelines suggest. EPA however discounts the nonlinear approach and thus only presents some illustrative examples.

- It may be useful for EPA to present both approaches and to ask the reviewers to comment on whether or not there is significant biological support in their minds for a nonlinear approach. If the reviewers then disagree, EPA can move forward with that recommendation. However, if the reviewers, consistent with NAS, think that a nonlinear approach is supported, then it may be helpful to also provide the reviewers with a rigorous set of questions seeking their comments on all aspects of the nonlinear approach as EPA has presented it in the illustrative examples.

Specific Science Comments:

- Page xlvii, and elsewhere, EPA states that for the Oral Slope Factor (OSF) calculations using Cheng et al., the upper 5% of the exposure range was excluded in estimating the slope. EPA further states: "Because this exclusion reduces the upper portion of the response where the slope is shallow, this likely better represents the slope in the region of the curve where the fatal cancer risk is increasing with dose, which is the equivalent of dropping the highest dose in an animal bioassay." As expert reviewers have commented on previous assessments (see Dale Hattis and Bruce Allen comments on 1,1,2,2 Tetrachloroethane available at: http://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=56732), we are unclear why EPA is dropping dose groups that may be very informative to the shape of the dose-response curve in the low dose region. As Dr. Hattis stated, in the context of BMD modeling, "In the light of Dr. Allen's premeeting comments I would recommend that EPA at least show earlier results with the highest doses and show the lack of fit that led the analyst to make the dose exclusions that were made. I agree that it is probably not correct to exclude the higher doses for all endpoints in summary fashion without analyzing the specific fits for different endpoints, if that is what was done in this case."

- Page 2-13, line 28-31, in discussing the Fingerhut study, EPA states that there was only a modest correlation between duration of employment and cumulative TCDD exposures. EPA also states that duration of employment is a surrogate measure of exposure to other chemicals and that the lack of correlation between employment and TCDD exposure suggests that confounding due to coexposures was unlikely to have biased results. It would be helpful if EPA could clarify why duration of employment is an appropriate surrogate for exposure to other chemicals and how a lack of relationship between TCDD exposure and employment suggests that coexposures were unlikely to have biased the results.

- Page 2-14, line 27-30, EPA states that when the authors adjusted for smoking, there were only small changes in the SMR for lung cancer in the overall cohort and in the higher exposure cohort. EPA should also note, looking at the confidence intervals (CI) for the SMRs both before and after smoking adjustment, that none of the SMRs were statistically significant. As the CI's spanned one it appears that all these relationships could be due to chance.

- Page 2-16, line 9 states that the study is not subject to important sources of bias, however, line 25 states that the study did not allow for the accounting of other possible confounders. Wouldn't the presence of confounders create a bias that could materially alter the findings of the analysis?

- Page 2-18, line 19, EPA refers to the results as "borderline statistically significant." It is unclear what this term means. As the CI's include one for all cancer sites combined, shouldn't EPA be clearly stating that the results were not statistically significant?

- Page 2-31, line 11, EPA states that statistically significant excesses were found for leukemia. As the CI's included one, shouldn't EPA be stating that the leukemia's were not statistically significant?

- Page 2-32, line 27-29, EPA states that no dose-response effects were detected and thus the study was inadequate for further analysis. We note that page 2-31 states that there was a statistically significant exposure-response trend for soft tissue sarcomas. Thus shouldn't this study have been included since it met EPA's criteria?

- Page 2-41, line 29, EPA states that for breast cancer mortality, the SMR was of "borderline statistical significance." As the CI spans one, shouldn't EPA be stating that the SMR was not statistically significant?

- Page 2-43, line 1-5, in discussing the Manz study, EPA states that since there were no changes in the production process after 1954, the exposures to other carcinogens (benzene etc) would be comparable for both time frames and thus confounding due to coexposures is unlikely to explain the dose-response relationship between all cancer mortality and TCDD exposure. The rationale for this statement could be clarified as its unclear why a consistent level of coexposures between exposure periods would negate the impact of confounding.

- Page 2-46, line 3-7, please clarify EPA's rationale for determining that other exposures were unlikely to have biased the results. As there were known coexposures to benzene and other compounds it is unclear why EPA is excluding them from having an impact.

- Page 2-59, line 26, EPA states that since the average age was 45.2, most women have not reached the age when breast cancer is typically diagnosed. It would be helpful for EPA to provide a citation here to the source of information regarding typical diagnosis. While NCI web pages state that the majority of cases occur in women over 50, it is not clear that the risk of developing breast cancer at age 45 is trivial. The follow-up period may have been long enough to detect some of these cancers and thus it is not clear that EPA should so quickly discount these findings.

- Page 2-86, lines 1-8, states: "Although the risk estimates produced for all cancer sites have important limitations and uncertainties, the data are far more consistent in terms of the magnitude of an association and latency intervals. The IARC evaluation has put forth the possibility of a pleuripotential mode of action between TCDD and the occurrence of cancer. Despite the criticism of this assertion by some (Cole et al., 2003), the general consistency of

an increased risk for all-cancer mortality across the occupational cohorts when latency intervals have been incorporated, provides adequate justification for dose-response quantification of all cancer sites combined." As this is an area where EPA acknowledges a divergence of opinions as well as limitations and uncertainties, it may be helpful for EPA to have a specific charge question asking the experts to comment on EPAs chosen approach.

- Page 2-89, lines 11-13, EPA seems to discount the Collins study because no increased risks were seen for non-cancer outcomes. Consistent with previous text in this section, wouldn't a better rationale for discounting this study be because it only looked at mortality effects which EPA does not consider appropriate for RfD development?

- Page 2-94, line 13, as the CI spans one, shouldn't the effect be characterized as not statistically significant (rather than as having marginal statistical significance)? A similar edit is suggested on page 2-95, line 2. As such, in discussions on page 2-94, line 24, it seems as though the Eskanazi study does not suggest that exposures impact menstrual cycle characteristics in women exposures before menarche.

- Page 2-96, line 16-18, as the study discusses coexposures to other DLC's, please clarify why EPA (or perhaps the authors) thought that most TEQ exposures were attributable to TCDD.

- Page 2-103, line 30-31, it appears that EPA is not conducting further evaluation of the uterine leiomyoma endpoint because an inverse dose response relationship was seen. We note that for many of the Eskenazi endpoints an inverse or U shaped relationship is seen. While this is a dose response relationship, and thus should meet EPA's criteria, it is unclear why EPA is not evaluating this study further. Leaving out all endpoints or evaluations which do not show a dose-response relationship in the direction EPA expects, may be discounting valuable information, in particular, information that could inform mode of action as well as dose-response. It may be helpful if EPA were to provide discussion and perhaps a table of studies which meet the criteria, but are not amenable to further analysis due to odd shaped dose-response curves (for instance EPA could also include the findings of Baccarelli 2004 on immunologic effects which showed an inverse dose-response relationship). This information, when evaluated as a whole, may be useful. It may also be helpful for EPA to have a specific charge question asking the expert reviewers to comment on how EPA is handling these studies.

- Page 2-104, lines 23-27, EPA discusses statistically significant changes in sperm count, progressive sperm motility and total number of motile sperm relative to the comparison group.
    - As EPA relies on these endpoints for the RfD, it would be helpful to have a discussion regarding how the measured values in the exposed group compare to national averages and how they relate to clinical effects (for instance, were the number of sperm in a range that today would lead a man to be deemed infertile?).
    - We also note that EPA finds that men aged 10-17 at the time of exposure (rather than 1-9 at time of exposure) showed opposite effects. It may be helpful for EPA to discuss whether or not these "opposite effects" were statistically significant. If they

are statistically significant, it may be helpful for EPA to have an expanded discussion of the clinical significance of all the findings.

- It is not clear which exposure groups showed the changes in serum E_2 levels and FSH concentrations. In addition, EPA may want to clarify if these changes were statistically significant and of clinical relevance.
- It would be helpful to ensure that the expert panel includes some reviewers who have clinical expertise in diagnosing and understanding male infertility. In addition, these and other experts may be able to inform a discussion of the mode of action on these effects to understand whether or not causation is plausible and would be expected.
- As these effects have been measured 40 years post exposure, and since the lifespan of a sperm is quite short (65-90 days) and they are created throughout a males life, it would be helpful to examine the population for any confounders (eg smoking, occupational exposures, medications such as those that treat blood pressure, obesity, alcohol use, etc) that are known to have impacts on male fertility. EPA should present a discussion of confounders and it would likely be very helpful for EPA to have a specific charge question asking reviewers to comment on EPA's choice of relying on this endpoint.
- Page 2-105, line 4, EPA states that the findings "could contribute to reported decrease in sperm quality in young men in the industrialized world." Its not clear that enough information has been presented for EPA to make this statement. In particular, the men examined here were over 40 yrs of age and its not clear that any possible confounders have been examined.
- As the exposure scenario in this study is quite different than the exposures of US women today, it may be helpful to specifically ask the peer reviewers to comment on this aspect of the exposure differences. Charge queston 2 in section 4 mentions this as background. It may also be helpful to ask the peer reviewers to comment on the implications of this for the development of a POD an RfD for todays exposures. Interestingly, when discussing the Alaluusua study, which also used the Seveso cohort, on page 2-112 EPA states: "For example, it is difficult to discern whether these health effects are a consequence of the initial high exposure during childhood or a function of the cumulative exposure for this entire exposure window beginning at the early age. If the latter is true, averaging exposure over the critical window would add considerable uncertainty to effective dose estimates given the large difference between initial TCDD body burden and body burden at the end of the critical exposure window."

- Page 2-108, lines 14-through page 2-109, throughout this section (including in its header) EPA refers to an examination of "neonatal thyroid function." It is important for EPA to be clear about what is a change in a hormone level verses a change in function. As the thyroid responds to positive and negative feedback of hormones, fluctuations in hormone levels are not always equated with a change in function that is physiologically or clinically significant. It may be more accurate to characterize this study as evaluating hormone levels rather than evaluating thyroid function. If EPA agrees, editorial changes are suggested throughout the document.
    - As the NAS ensured when they evaluated perchlorate, it would be very helpful for EPA's review to include clinical endocrinologists and developmental biologists who

can comment and provide insights as to the clinical significance of the changes EPA is relying upon for the RfD derivation.

- EPA states that "Apart from iodine deficiency, no other environmental exposure has been associated consistently with reduced neonatal thyroid functioning." What about compounds like thiocyanates, perchlorate or cigarette smoke? Aren't these compounds which can impact thyroid function? We note that table 2-5, page 2-213, states that there was limited evidence of confounding. It would be helpful for EPA to clarify which relevant confounders were adjusted for.
- Page 2-109, line 11, it would be helpful for EPA to explain why a cutpoint of 5 uU/ml was used to assess associations. Is this level associated with clinical adversity?
- EPA should provide discussion of potential coexposures that could act as confounders (eg smoking) and how they were addressed.
- EPA states that the mean TSH levels were positively correlated with average soil TCDD levels in the three zones. It would be helpful to clarify if these correlations were statistically significant.
- It looks like the odds ratio's were only statistically elevated in the highest exposure zone (zone A) compared to the reference population.
- As the exposure scenario in this study is quite different than the exposures of US women today, it may be helpful to specifically ask the peer reviewers to comment on this aspect of the exposure differences. Charge question 2 in section 4 mentions this as background. It may also be helpful to ask the peer reviewers to comment on the implications of this for the development of a POD an RfD for todays exposures.

- Section 2.4.2, it seems that many of the studies discussed are more informative as mode of action studies rather than studies that help determine the non-cancer endpoint of concern. For instance, Devito 1994 evaluates TCDD impacts on EROD induction and phosphotyrosyl protein, Hassoun 1998 evaluated impacts on lipid peroxidation and oxidative stress, Slezak 2000 looked at GSH levels as well as oxidative stress, Crouch 2005 evaluated IGF signaling, and Rier 2001 evaluated TNFalpha. In many of these studies, it is not clear that the magnitude of response is outside the range of normal variability, thus it is not clear how the studies met EPA's criteria for inclusion in the evaluation. While the EPA document does not have a section addressing mechanism of action, has EPA considered a separate section that would include studies that inform mode of action?

- Page 3-52, line 8-12 for the human equivalent dose (HED) in the gestational exposure scenario, EPA assumed that pregnancy begins at 45 yrs of age. EPA acknowledges that this is health protective as the daily exposure achieving the target blood concentration is smaller than for earlier pregnancies. As studies show that the birth rate for women ages 25-29 is 116 births per 1000 women and 0.6 births per 1000 women for women aged 45-49, with the average age of pregnancy being between 25-27, it would be helpful for EPA to also provide the calculations that use the average aged woman. EPAs cancer modeling approach is typically health protective and in the RfD calculations EPA adds an uncertainty factor to protect sensitive individuals, thus it would be very helpful for EPA to provide the impacts of having the gestational exposure scenario use the average aged woman. While understanding the impacts of the most sensitive populations is critically important, it is also important that risk managers understand the impacts to the average (or more average) individual. EPA

should provide a clear discussion of the impacts in overall values when EPA uses the average verses older woman model. It may also be very helpful for EPA to have a specific charge question asking the experts to comment on this parameter. It may also be very helpful for EPA to have a specific charge question asking EPA to comment on the approach of determining the HED based upon the continuous daily intake that would result in the target concentration over peak 5- year period (as per lines 12-31). EPA states that this is health protective. It would be helpful for EPA to provide information to the reviewers so that they can see the impacts of this choice compared to other choices.

- Page 4-8 , line 30-32, states in discussing the epidemiological studies: "EPA did not conduct Benchmark Dose modeling because the covariates identified by the study authors could not be incorporated by modeling the grouped response data." As a BMD approach is always preferred over a NOAEL/LOAEL approach, we are wondering if EPA did any outreach to the study authors (as has been done for other assessments) to try to gather the data points that would allow for BMD modeling. Would such data be informative and helpful?

- Page 4-9, line 24, in discussing Mocarelli EPA states that men who were 10-17 were not affected. However, according to the previous discussion and the Mocarelli publication it looks like this age group showed statistically significant increases in sperm count, motile sperm counts, FSH and also reduced estradiol. Shouldn't EPA be discussing these impacts and their potential clinical significance as well? For both age groups, an enhanced discussion of potential mode of action and clinical and physiological significance would be helpful. We suggest that EPA ensures that there are fertility experts on the review panel. A specific charge question relating to the opposing effects (depending on age of exposure) would be helpful. EPA assumes this is due to the need for a 10 yr critical exposure window, but the discussion for the scientific justification for this, considering the lifespan of a sperm, is not provided. Such a discussion would be useful and would give the expert reviewers something to react to during their review. Similar comments apply to the discussion on page 4-22, line 1-3.

- Page 4-15, line 3-18, EPA describes the BMD modeling approach. While this information is all in Appendix E, it may be helpful for EPA to create a table which arrays the potential BMD/ BMDL values as well as their AIC values so that readers can easily see the implication of the choices EPA made. In previous assessments we have seen that the choice of the lowest AIC, which may differ by only 0.5 from a non-chosen AIC value can have big implications for the chosen BMD/BMDL value. EPA states that if the BMDL values were within a factor of 3, EPA chose the model with the lowest AIC value, otherwise EPA chose the model with the lowest BMDL. Does this mean that the lowest BMDL value, rather than the AIC value drove the model choice? More clarification on what this looks like and its implications for the chosen BMDL values would be helpful. A specific charge question on this aspect of EPAs approach may be very helpful as this could end up being a key driver in choosing the BMDL value.

- Page 4-15, line 23 states: "On occasion, high doses were dropped and the models were refit." As per comments above (see comments regarding page xlvii), it would be helpful for EPA, at a very minimum, to provide information, perhaps in the BMDL summary tables, that makes

it clear when dose groups were dropped to improve model fit. A charge question on the impacts of this approach may also be useful.

- Page 4-18, line 24-30, EPA states: "Baccarelli et al. (2008) reported increased levels of TSH in newborns exposed to TCDD in utero, indicating a possible dysregulation of thyroid hormone metabolism." This appears to be the first sentence EPA has put out describing potential impacts. More detail here, and whether or not the changes put the hormone levels in a range that is clinically or physiologically believed to be associated with dysregulation would be helpful. What is the citation that Baccarelli provides for their statement? More discussion of the science regarding the impacts associated with the changes seen would be extremely helpful. Similarly, EPA notes the findings for men who were 1-9 yrs old in the Mocarelli study, but does not discuss the findings of the older group or the clinical or physiological implications of the changes.

- Page 4-21, line 14-19, as per comments above, please provide more scientific information to support the statement regarding dysregulation of thyroid hormone metabolism. EPA mentions that the 5uU/ml standard "was established by the World Health Organization (WHO, 1994) as an indicator of potential iodine deficiency (and potential thyroid problems) in neonates." What exactly did WHO intend when saying it was an indicator of potential thyroid problems? What is the rationale of using this as a cutpoint for adverse effects? Or does EPA believe this is a precursor to adverse effects. As WHO mentions, iodine deficiency, did the study authors also looks at iodine levels in the children? Including this information would be extremely helpful.

- On page 4-22, line 1-18, EPA further describes the Mocarrelli studies. As per comments above (page 4-9, line 24), edits to the discussion of the men in the older age group are suggested. EPA states: "Although a decrease in sperm production of 20% would not have clinical significance for an individual, EPA considers a 20% shift in the population mean to be of biological significance." What is the scientific rationale behind this statement and supporting the determination that a 20% shift in the population mean is of biological significance. EPA has previously mentioned that RfD's are based on endpoints that are adverse or precursors to an adverse effect. Is this referring to an individual response or a population based response? More clarity would be helpful as it seems as though using a POD that is known to not be of clinical significance to an individual seems like a novel approach. A charge question on this specific aspect of choosing the Mocarrelli study for the RfD derivation would be helpful.

- Page 5-5, line 4-6, please clarify that the relative risk for all cancer mortality could be due to chance since the CI includes 1.0. Similarly, in lines 12-15, please clarify if the increases were statistically significant. It would be helpful to clarify this for risks presented throughout section 5.1.2.1.1

- Section 5.2.3.1.2, as NAS recommended that "EPA should compare cancer risks by using nonlinear models consistent with a receptor mediated mechanism of action and by using epidemiological data and the new National Toxicology Program (NTP) animal bioassay data (NTP, 2006)," it is not clear why EPA has not provided an updated nonlinear evaluation of

the epidemiology studies presented in the 2003 assessment. It is clear that EPA has conducted a robust evaluation of two newer studies, however providing the modeling that NAS recommended for these older datasets may be very informative and it is not clear why EPA has not provided these evaluations as part of their update.

- Page 5-43, line 30-31, EPA states: "The bioassay-based cancer dose-response assessment in this section has used the multistage model which is the standard model choice for such assessments and has been the basis for most of EPA's cancer risk assessments. In that sense, there is no associated uncertainty for model choice." Please explain the scientific rationale for this statement.

- In section 5.2.3.4.1.3, EPA in discussing receptor theory modeling, states that TCDD will show linear dose response binding in the 1-10% receptor occupancy region. It may be helpful if EPA were to discuss what the expectations are for receptor occupancy at current background US exposure levels.

- Page 5-67, line 30, EPA recommends and OSF when the target risk range is 10^{-5} to 10^{-7}. It is unclear why EPA has chosen this target risk range. As most environmental statutes refer to the 10^{-4} to 10^{-6} risk range, it would be extremely helpful for EPA to provide a recommendation for this risk range.

- Page 6-3, lines 1-4, instead of misrepresenting the OMB efforts (for example draft proposals are not retracted) it would be more helpful to readers if EPA were to cite the current state of the OMB and OSTP recommendations regarding risk assessment (similar to how EPA describes the NRC reports). The following language is suggested: In 2007, the US Office of Science and Technology Policy (OSTP) and the Office of Management and Budget (OMB) released a Memorandum on Updated Principles for Risk Analysis. These principles provide Federal Agencies with recent guidance from the scientific community, the Congress, and the Executive Branch concerning generally-accepted principles for risk analysis. Of particular interest, the principles state that: "When something more than a superficial analysis can be conducted, quantitative uncertainty analysis, sensitivity analysis, and a discussion of model uncertainty can greatly inform risk management decisions." (memorandum available at: http://www.whitehouse.gov/omb/assets/omb/memoranda/fy2007/m07-24.pdf)

- Section 6.1.3, it may be helpful for EPA to provide a citation to the source that led EPA to discuss the specific basic requirements that are presented in this section. If one looks at the section headers in this section, it is not clear how these aspects were determined to be the basic requirements.

- Page 6-10, line 25-28, EPA interprets the cancer guidelines to imply that to use a nonlinear model," it must have a preponderance of evidence" to override the default choice. This language is not in the Cancer Guidelines and appears to be an over-interpretation. As per our general comment above, EPA may wish to revise this (perhaps using quotes from the cancer guidelines) to reflect that when alternative approaches have significant biological support, they may be presented.

- In section 6.4 EPA spends a great deal of time discussing the lack of feasibility of conducting a quantitative uncertainty analysis. This is done in the context of responding to the NAS. It

seems that what could be much improved is EPA's discussion and presentation of qualitative uncertainties. It would greatly benefit this updated dose-response assessment if EPA provided further discussion on these qualitative uncertainties.

- Page 6-34, line 19-24, EPA states: "The NAS committee explicitly requested that the uncertainty attending the choice of a BMR be quantified. First of all, simply plugging other values in for the BMR does not constitute a quantitative analysis of uncertainty. The plugged-in values must be sampled from some uncertainty distribution. Since this concerns volitional uncertainty, there is no underlying distribution from which to sample, unless the choice of BMR is related to some claim about the state of the world." While EPA may argue about what constitutes an uncertainty analysis, a more useful discussion might be one that focuses around what EPA has done to address the NAS request. Perhaps a similar approach would be more useful throughout section 6.4 and 6.5. We note that the NAS stated: "When selecting a BMD as a POD, EPA should provide justification for selecting a response level (e.g., at the 10%, 5% or 1% level). In either case, the effects of this choice on the final risk assessment values should be illustrated by comparing point estimates and lower bounds derived from selected PODs."

- Section 6.5, as per comments above, it may be more useful for EPA to present what they have done in response to NAS, rather than to simply argue against the NAS recommendations. In addition, on page 6-37, line 22, EPA makes the argument that rationale does not exist to choose between options. While this may be true, perhaps the point of the NAS comments were to inform the risk managers by providing them with a range of information.

- Page 6-39, lines 7-14, contains a pure policy determination that any pushes forward in the uncertainty arena must be nondisruptive to the traditional process. Is this the official agency position? Considering the number of assessments that EPA is working on that are proposing safe levels of exposure below our natural background exposures, perhaps some out of the box thinking that improves the traditional process, and may be disruptive, could be of use in moving the risk analysis field forward.

Editorial Comments (with Scientific Impacts):

- When quoting the NAS we suggest that EPA provide full sentences without deletions of text from the NAS. For instance (and this is just one example), on page 2-1, lines 30-33, EPA quotes the NAS from page 27 of their report. However this quote does not include the full sentence and this is not clear to the reader. If EPA is going to delete text or end sentences early, at a minimum EPA should note this and include the full sentence in the footnote.

- Page xvii, section 6.5 is entitled "Conclusions" yet it is only responding to some comments addressed in chapter 6. It may be helpful for EPA to provide some overall conclusions and/or recommendations stemming from their updated analysis and response to NAS.

- Page xxxiii, EPA reiterates parts of the dioxin science plan. As EPA has not met their deadlines for some of these items, we suggest deleting the details provided and simply referring readers to the webpage (which hopefully explains and provides updates as to how EPA is doing in meeting the established timelines and provides revised dates).

- Page xxxviii, lines 12 and 28, in discussing the epidemiology and bioassay inclusion criteria and evaluation process, EPA states that "only studies meeting these criteria were included…" Please clarify if studies had to meet all the criteria or just some of the criteria.

- Page xliii, line 20, please provide citation for the dioxin workshop recommendations. Similarly please provide citation for statement on line 26.

- Page li, its not clear that the discussion of Bussard et al, is appropriately presented on line 27. This book chapter presents a summary from a workshop that evaluated the approaches EPA then describes. However, EPA presents these as possible methodologies. It may be helpful to clarify that these approaches do not represent the complete universe of methodologies but only the four evaluated at the workshop.

- Page 2-5, line 10, in citing the draft BMD guidance, EPA has it cited as US EPA 2008. Please note that this document is from 2000. In the references, it is cited both ways and thus some edits are needed here and in the references. In the text it would be helpful for EPA to clarify that they are referring to a draft guidance that does not represent official agency position, nor has the approach been codified as EPA states. We encourage EPA to work towards updating and finalizing this document in the near future.

- Page 2-21, line 10, refers to "the IARC recommendation". As this is the first time this recommendation is mentioned, it would be helpful to provide a citation and clarify exactly what recommendation EPA is referring to.

- Page 2-25, lines 20-25, EPA states that "although the serum-based measures did not fit the data as well as the exposure scores, the authors regarded them as providing reasonable fit." Is there any further information from the authors that may help clarify what was meant by "reasonable fit"?

- Page 2-44, line 2-4, please clarify whether or not this dataset is further analyzed.

- Page 2-49, line 25, as the $p = 0.06$, please replace "marginally statistically significant" with "not statistically significant". Similar comment for line 27 and on page 2-50, line 2 where EPA states "borderline statistically significant".

- Page 2-81, line 3-4, please clarify whether or not this dataset is further analyzed.

- Page 2-112, line 3, EPA refers to "developmental defects", does EPA mean to say "dental defects"?

- Page 2-116, line 1-2, it would be helpful for EPA to provide specific citations back to the specific studies, with positive findings to which they are referring here.

- Section 2.4.2, throughout this section, when discussing the animal bioassays, EPA does not always provide a clear statement as to whether or not the study was considered further for dose-response modeling. As this summary statement was very helpful in the discussion of the epidemiology data, a similar statement for each animal bioassay would be quite useful. Similarly throughout this section, it would be helpful for EPA to clarify who identified the

LOAEL in each study. Many times it is not clear if the LOAEL is determined by EPA or by the study authors.

- Section 2.4.2.6, in the epidemiology section, EPA evaluates cancer studies first and then non-cancer studies. EPA may want to use a similar approach in the animal bioassay section, rather than presenting the animal cancer endpoint studies after presenting the majority of the animal bioassay data.

- Section 3.3.3.3, as this section only addresses smoking, EPA may want to consider renaming the title of this section.

- Page 3-44, line 21, EPA states that the re-estimation of urinary clearance did not result in any significant changes in the fit and performance of the original model. It may be helpful for EPA to numerically present the changes so that readers can better understand what EPA means.

- Page 3-49, line 31, EPA states that absorption and excretion parameters were among the sensitive parameters in the rat. Does this mean that they were the most uncertain? Please clarify.

- Table 4-6 lays out the strengths and limitations/uncertainties associated with the animal bioassays that provided candidate PODs for the RfD. A similar table describing the strengths and limitations/uncertainties associated the epidemiology studies would also be helpful.

- Page 4-11, line 24, mentions the menstrual effects reported in Eskenazi. It may be helpful to clarify what effects in particular EPA is referring to and if the effects were statistically significant and have clinical relevance.

- Page 4-15, line 6, please clarify that the current guidance is only a draft and has not been finalized by the agency.

- Table 4-4, please provide units for the NOAEL/LOAEL and BMD/BMDL values

- Table 4-5, page 4-18 describes the table as presenting a wide area of toxicological endpoints. It may be helpful if for each of the endpoints, EPA defines the effect seen as adverse or as a precursor effect.

- Section 4.4, in the past in addition to having a discussion of the uncertainties and limitations in the RfD derivation, EPA has also provide a table that visually and succinctly shows the impact of the major assumptions and uncertainties. This table has always been very useful to readers and reviewers. EPA may want to consider such a table for this document. For instance EPA could expand table 6-1 to show the impacts that their determinations have on each of the uncertainties presented.

- Page 5-4, line 18-20, as so many studies were reviewed in section 2.4, it may be helpful to specifically cite the recent ones that do address the concerns for potential confounders. Age 5-6, line 3-5 states: “Despite these uncertainties, many of the more recent studies have greatly improved exposure assessments compared to earlier studies of the same cohorts and have addressed the potential for confounding and other types of biases.” Perhaps we missed

this discussion in section 2.4, but it may be useful to clarify exactly what these improvements in exposure were.

- Page 5-6, line 17, is it true that strength of association, is considered "irrespective of statistical significance"? If this is true, its not clear why the statistical community exists if not to inform the determinations regarding strength of association.

- Page 5-6, line 26-30, EPA states that the consistent results are not likely due to chance. As many of the CI include 1.0, isn't this statement inconsistent with the statistical evaluations?

- Page 5-15, section 5.1.2.3.3 includes a section for the hypothesized more of action of TCDD in rodents and then has specific subsections for each type of tumor. It might be helpful to have a similar section for the mode of action of the particular human cancers that have been detected in the epidemiological studies.

- Page 5-22, section 5.2.2 could be improved by clearly listing each of the NAS key recommendations and then giving a short summary of the EPA response with reference to the section of the response document that addresses the comment. For instance it is unclear in this overview section what EPA did to address the NAS key finding that EPA should compare cancer risks by using nonlinear modeling.

- Section 5.2.3.6, in discussing the quantitative uncertainties in the slope factor estimates, EPA may want to include a section which summaries the uncertainties and limitations associated with use of the Emond PBPK model.

- Page 5-44, line 28, to defend EPA's choice of POD, it seems as though EPA should replace the word "fact" with "policy determination".

- Page 6-9, line 23-24 EPA states that the reference values "suggest a biological population threshold beneath which no harm is anticipated." This is an oversimplification as the complete definition of an RfD states that it is: "An estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime." It is important that EPA be clear to readers that reference values are estimates that may contain a wide range of uncertainty and also don't claim to provide no harm but instead use a standard below which there is no "appreciable risk of deleterious effects."

- Page 6-34, line 28-29, it is unclear where EPA presents the suggested research agenda that is referred to.

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 ensuring a rigorous peer review of this highly technical document.)

- It may be helpful to clarify to reviewers, as well as the public, what steps EPA will take after this review is complete. For instance, if EPA plans to finalize the dioxin reassessment after

this review, it may be helpful to state that. Alternatively, if EPA plans to take these comments into account, and subsequently produce a revised draft assessment that will go through a review process before being finalized, EPA may want to state that as well. As the 2009 dioxin plan is a bit outdated, it may be helpful to be very clear about exactly what this review will be informing.

- In the 2nd paragraph EPA states that the draft document is not an assessment per se and that it is designed to supplement the 2003 reassessment. While this is correct, in supplementing the 2003 reassessment, EPA now provides updated recommended non-cancer and cancer values. If EPA is going to continue to provide these recommendations, we suggest that EPA be clearer in characterizing this as an updated assessment of the cancer and non-cancer health effects and their impacts.

General Charge Questions:

- Q1, since this is a response to the NAS comments, EPA may additionally want to ask the experts to comment on whether or not the document adequately captures the predominant NAS concerns related to dose-response assessment.

Section 2:

- It may be helpful to add a charge question to explicitly take comment on the inclusion criteria EPA has developed for both the epidemiology data and the animal bioassays. While it should be helpful to have the experts comment on each element of the criteria, in particular EPA should ask the reviewers to comment on the inclusion of only those studies showing effects. We note that there are differences in the two criteria (epi and animal) in that the epidemiology criteria includes only studies that show "adverse effects" and yet the animal bioassays include studies which show "responses outside the range of normal variability." It may additionally be helpful if EPA defines, and then takes comment on their definition of an adverse effect for the purposes of this document.

Section 3:

- In Section 3 of the document, EPA includes a lengthy discussion of the CADM model, including discussion of its structure, mathematical representation, parameter estimation, model performance and confidence. In charge question 3(1), EPA simply asks the reviewers to comment on the justification of using the Edmond model as opposed to other models. As EPA spends so much time on the CADM model, it may be useful to have a specific charge question in this section which asks the PBPK experts to comment on EPAs review of the CADM model and its utility.

- While question 3(1) asks for comment on EPAs justification for using the Emond model, there are a few areas of EPA's approach where EPA may want to ask for specific comments. These include:
    - Taking comment on EPA approach to addressing concerns that the dose-dependency of metabolic elimination in the model was not calibrated to human data and the

conclusion (as stated on page 3-39) that it is not unreasonable to use the model as it stands.
  - Taking comment on EPAs conclusions regarding the reliability of the model for simulating liver concentrations in the rat and human
  - Taking comment on EPAs conclusions regarding the confidence in the model for different important aspects as described by EPA in 3.3.4.5

- In Q4, in addition to asking the reviewers to comment on EPAs estimation, it may be helpful to specifically ask reviewers to comment on EPA's use of lifetime average daily dose.

Section 4:

- In Q2, EPA mentions that there is uncertainty due to the influence of the high-dose pulse exposure and the fact that the Seveso cohort exposure pattern is different than the exposures experienced by the general population. It would be helpful to ask the expert reviewers to comment on the implications of this and what it may mean for the relevance of the determined POD's and RfDs. If EPA could say anything about the direction of the uncertainty (eg leads to an underestimate or overestimate of risk) this would also be helpful.

- In Q2, it may be helpful to have an expanded charge question that asks reviewers to specifically comment on EPA's approach for estimating the TCDD intake associated with the LOAEL in the Mocarelli study (as described by EPA on page 4-10). Considering the endpoint of concern, a question regarding the use of peak blood concentrations may also be helpful.

- In Q2, EPA states that for the Baccarelli study EPA used reported maternal levels and asks for comment on this. However page 4-9 states: "Therefore, EPA determined the maternal intake at the LOAEL from the maternal serum-TCDD/TSH regression model by finding the maternal TCDD LASC at which neonatal TSH exceeded 5 μU/mL. EPA then used the Emond PBPK model under the human gestational scenario (see Section 4.2.1) to estimate the continuous daily TCDD intake that would result in a TCDD LASC corresponding to a neonatal TSH of 5 μU/mL at the end of gestation, with the resulting maternal intake established as the LOAEL (0.024 ng/kg-day), shown in Table 4-1 as a candidate POD for derivation of candidate RfDs." It may be helpful for EPA to explicitly ask the expert reviewers to comment on this entire approach.

- In Q3, EPA takes comment on the approach of averaging TCDD blood concentrations over the entire dosing period. Page 4-13, line 8, mentions that EPA started by using the initial peak TCDD blood concentration as this was considered to be the most relevant exposure metric. It may be helpful for EPA to also take comment on the use of peak blood concentration. Additionally page 4-14, line 16-19, discusses some critical choices made during EPA's model fitting approach, including the exclusion of supralinear fits and saturated models. EPA may also want to specifically ask the expert reviewers to comment on EPAs approach.

- Q4, it may be helpful to expand this question as EPA typically does to inquire about the reviewers opinions on the chosen response levels.

- Q6, EPA may want to explicitly ask reviewers to comment on EPA's determination of the critical effect, and the use of the NOAEL/LOAEL approach for quantify the POD for the epidemiology studies.

- Q7, as EPA is also taking comment on other studies that reviewers may recommend, EPA may want to expand this question to include asking for recommendations related to the uncertainty factors applied to other potential studies that reviewers may think are more appropriate for an RfD determination.

- Q8, as section 4.4 also discusses the key limitations of the epidemiologic studies, it may also be helpful to ask the expert reviewers to comment on these limitations and any implications they may have for the utility of the RfD EPA has derived using these studies.

Section 5:

- Section 5.2.3.4.1.2, EPA spends some time interpreting the cancer guidelines and applying new terminology to define threshold/nonthreshold responses as well as nonlinear models and defines them in an individual and population sense. It may be helpful if EPA were to ask the reviewers to comment on these definitions. In particular, it may also be helpful for EPA to seek confirmation of interpretations of the "zero slope at zero" model for the population and the impacts that receptor kinetics may have on the ultimate population response. This is a critical point that EPA makes and seems to underlie the EPA response to the NAS as EPA argues against the NAS conclusion which favored a nonlinear model that would include a threshold response.

- In section 5.2.3.4.1.3, EPA concludes that linear low dose extrapolation should be the preferred modeling approach. It would be helpful to have a specific charge question that asks the expert reviewers to comment on this conclusion and its scientific justification. In particular, EPA, on page 5-72, states that this choice was made because EPA determined that the Agency lacked sufficient evidence to support an assumption of nonlinearity. In addition, EPA also states that there was insufficient evidence to support an assumption of a threshold Taking comment on these determinations seems to be of critical importance for establishing consensus for EPA's preferred modeling approach..

- It may be helpful to ask the reviewers to specifically comment on EPAs decision to use cumulative serum concentrations as the primary metric for carcinogenicity (as described on page 5-69).

- In Q1, it may be helpful to clarify what routes of exposure EPA is referring to when providing the weight of evidence cancer descriptor. We presume this is only related to oral exposure but if EPA is going to broaden thisa discussion of the rationale should be added to the document.

- In Q4, it might be very helpful to ask the reviewers to comment on EPAs choice of using a $BMDL_{01}$ (1% excess risk) as the POD for the development of candidate oral slope factors. Similarly, as NAS recommended that EPA consider nonlinear models ("Because the committee concludes that the data support the hypothesis that the dose-response relationship for dioxin and cancer is sublinear, it recommends that EPA include a nonlinear model for cancer risk estimates but that EPA also use the current linear models for comparative purposes."), EPA may want to specifically ask a charge question about their choice of relying on the linear model rather than the nonlinear models.

- In Q10, EPA states that they considered nonlinear approaches and asks the expert reviewers to comment on other approaches that could be developed. It may also be helpful if EPA asks the expert reviewers to provide scientific comments on the two illustrative approaches EPA has provided. It may also be helpful to ask the reviewers to comment on EPA's conclusions regarding the limitations and utility of these approaches.

- EPA may want to consider a charge question that takes comment on EPA's determination to not conduct a meta-analysis.

Section 6:

- Q1 describes the discussion as a response, however much of the discussion reads as a white paper regarding quantitative uncertainty analysis. Thus EPA may want to reframe this question and ask for comments on all aspects of the discussion provided. In particular, EPA may want to ask the reviewers to comment on each of EPA's feasibility determinations as well as the EPA responses to NAS in section 6.5. EPA may also want to ask about whether or not the reviewers find that the aspects of discussion are useful in addressing the concerns articulated by NAS. In particular, EPA could also take comment on the utility of their discussion framed around volitional uncertainty and how this type of uncertainty limits the ability to conduct a quantitative uncertainty analysis.