

1 **Section 1 – General charge questions**

2
3 *Drs. Buckley, Mocarrelli, Schecter*

4
5 **Overview assessment, details to be found in subsequent charge Qs**

6
7 **1.1a. Is the draft Response to Comments clear and logical?**

- 8
9
 - 10 • Yes, in general, EPA has been effective in developing a clear, transparent, and logical response;
 - 11 • The panel was particularly impressed with the process that EPA used for identifying, reviewing, and evaluating the relevant literature including a public workshop;
 - 12 • Executive Summary is important and provides concise summary;
 - 13 • Issues: Provide better integration across chapters (details in Charge Q2) clear description for inclusion and exclusion of studies/data progressing through the document;
 - 14 • Needs to be more clearly written;
 - 15 • Glossary may be helpful to improve clarity given diversity of users;
 - 16 • The large size of the document diminishes the clarity;
 - 17 • We feel EPA’s response is incomplete in considering nonlinear dose response, mode of action, and uncertainty analysis;

18
19
20
21
22 **1.1.b Has EPA objectively and clearly presented the three key NRC recommendations?**

- 23
 - 24 • In general, EPA has objectively and clearly presented the three key NRC recommendations;

25
26
27 **1.2. Are there other critical studies that would make a significant impact on the conclusions of the hazard characterization and the dose-response assessment of the chronic noncancer and cancer health effects of TCDD?**

- 28
29
30
 - 31 • With respect to hazard characterization, consider a more balanced assessment of negative studies;
 - 32 • With respect to dose-response, there are none that we know of. We need to see what comes out of relevant specific charge questions;

1 **Section 2 – Transparency and clarity in selection of key data sets for dose-**
2 **response analysis**

3
4 *Drs. Lawrence, Hauser*

5
6 **2.1 Is this section responsive to the NAS concerns about transparency and clarity in dataset**
7 **selection for dose-response analysis?**

8
9 Response: SAB generally noted that this section was responsive to NAS concerns about
10 transparency and clarity. EPA improved their approach in addressing these concerns from the
11 original document. The EPA's collaboration with Argonne National Laboratory and invitation to
12 the public to engage in updating the literature search to identify all appropriate studies for
13 evaluation, as well as the conduct of the Dioxin Workshop in February of 2009, were
14 instrumental in enhancing the transparency and clarity regarding the process of selection of
15 studies for the dose-response analysis. The development of clear criteria for study evaluation
16 and inclusion were crucial in resolving the concerns raised by the NAS. Five considerations were
17 used to evaluate the epidemiologic studies. Three inclusion criteria were then used to select
18 studies to use for TCDD quantitative dose-response assessment.

19
20 The SAB found that EPA defined a clear and transparent process and then conducted their
21 review in the document in a manner consistent with what they said they would do. The document
22 presented a clear identification of the process and studies used. For example, the process and
23 criteria used by EPA to select key data for dose response analyses is clearly described in section
24 2.3 of this document and in the Executive Summary. Flow diagrams (e.g., ES-1 and ES-2) very
25 clearly demonstrate how studies were chosen for inclusion. Likewise, Appendix B, which
26 includes a point-by-point evaluation of which epidemiological studies were included and
27 excluded was useful and provides a detailed rationale for why the EPA used the particular
28 studies selected in this document. In addition, the results of the literature search performed by
29 EPA are available online, although clarity could be improved by providing search words used for
30 the MedLine searches. A clear case for including high-quality human studies over animal
31 studies is also made.

32
33 While consensus was not reached, many SAB members emphasized that overall clarity and
34 transparency regarding dataset selection would be greatly enhanced if EPA were to make this
35 section (and the document as a whole) more concise. In its present form, this section was viewed
36 as overly verbose; to the detriment of clarity. Some SAB members found that the tone of the
37 document modestly reduced clarity and transparency, but this was not a consensus opinion.

38
39 Recommendations:

- 40
41 • Careful and extensive editing to revise and consolidate this section (and the document as
42 a whole) are strongly recommended. Specifically, editing should include aspects of
43 grammar, syntax, reduction in redundancies, and efforts to give more succinct responses
44 to NAS concerns and suggestions.
45

- 1 • This Section could be structured such that it is easier to follow a study from one section
2 of the document to another; in other words, improve overall document integration, using
3 Section 2 as the foundation for this integration.
4
- 5 • ML: It would be useful to provide some additional information to help justify the 2
6 critical studies (Baccarelli et al and Mocarelli et al) that were used to establish the PoD.
7 Also include in Summary section a clarification of why the Emond model is not very
8 reproducible for mice.
9

10 Dr. Karl Rozman's Comments

11
12 I must take issue with the EPA claims about transparency in its documents and decisions. This
13 document repeatedly focused on dose-response as justification for most of the major conclusions.
14 I have studied and attained a good understanding of all important dose-responses of dioxins
15 including cancer and hormonal effects. I was not able to validate any of these "so-called" dose
16 responses in this document and was herefore concerned about the specific expertise of the
17 authors in this area. To resolve this, I requested the names of the authors so that I could evaluate
18 their experience and competence with dose-response data. My request was denied which leaves
19 me with reservations about this document and EPA's claims of transparency.
20
21

22 **2.2. Are the epidemiologic and animal bioassay study criteria and considerations** 23 **scientifically and clearly justified?**

24
25 Response: This section of the document was deemed generally responsive to NAS concerns and
26 suggestions. The criteria for epidemiological and animal studies are clearly presented, as is the
27 rationale for the parameters used to include studies (Figures ES-1 and ES-2). The EPA is
28 complimented for efforts to present the nuanced differences and complicating issues surrounding
29 this subject in a comprehensive and logical manner. The majority opinion expressed by the SAB
30 is that the general study criteria and considerations are scientifically justified and clearly
31 described. The five criteria are excellent guidelines, however, two require detailed information in
32 order to evaluate study's feasibility. In several instances, the SAB requested further refinement
33 and clarification. There is a clear and logical description of why a cutoff of 30 ng/kg-day was
34 used (p 2.8 and 2.9) in the selection criteria, and the summary tables (e.g., 2-3 and 2-7) were very
35 useful, providing a detailed but readable format for the study data. However, several concerns
36 were discussed, and are summarized here.
37

38 The rationale for distinct criteria for epidemiological and animal studies should be made
39 stronger, and data set selection for non-cancer and cancer endpoints has room for further
40 clarification and justification.
41

42 While the scientific justification for the inclusion was overall justified and well explained, the
43 rationale for exclusion criteria requires refinement. Excellent studies were excluded for reasons
44 that are not well justified. Several SAB members expressed awareness of other studies, with a
45 mixed sense of whether including them would or would not have a significant impact on the

1 dose-response assessment. What follows are specific points of concern raised by SAB members
2 regarding inclusion/exclusion criteria:

- 3
- 4 • It is not clear why a specific statement of TCDD purity must be made explicitly. This is
5 an important issue because TCDD is available from relatively few commercial sources
6 and those sources certify purity of the chemical (typically \geq to 98% purity). Therefore,
7 inclusion as one of the three major selection criteria seems somewhat arbitrary and the
8 rationale could be clarified.
- 9
- 10 • The explanation that the “study design is consistent with standard toxicological practices”
11 is unclear. It would be helpful to explain what aspects are different in toxicological
12 studies than in physiological studies and the rationale for these differences.
- 13
- 14 • The statement “The study criteria shown below and in Figure 2-3 for animal bioassay
15 data reflect EPA’s preferences for TCDD-specific study inclusion, some of which are
16 based on common practices and guidance for POD selection and RfD and OSF
17 derivation” (p 2-5) does not help the reader understand the rationale for criteria. Please
18 define what these common practices are more clearly.
- 19
- 20 • Section 2.4.1.2.1.5.3 (p. 2-110). In the manner in which it is currently written, the
21 rationale for excluding the studies by Baccarelli et al (2002, 2004) on the relationship
22 between TCDD and immunological effects rests on a rather weak foundation. The text in
23 this section states: “Interpreting the inverse association between TCDD exposure and IgG
24 levels in terms of clinical significance is not possible.” This is predicated on the idea that
25 if plasma IgG levels do not sink down to those measured in immunocompromised
26 individuals, then there is no clinical significance. Some SAB members expressed the
27 expert opinion that current human and animal immunology data would not fully support
28 this. It is possible and likely that there are individuals within a population that may not be
29 diagnosed as immune compromised, but whose immune responses fall outside or on the
30 very edges of the range of normal. However, including these studies would not likely
31 change the outcome of EPA’s dose response analyses, therefore there is no specific
32 recommendation made to add these studies.

33
34 Recommendations:

- 35
- 36 • Given the overall sense of the SAB that perhaps EPA may have been too stringent in
37 exclusion of some excellent studies, two recommendations are made: (1) revise pertinent
38 sections of the document, with the above detailed points in mind and (2) consider adding
39 information to the appendices and/or tables to provide readers with clarification regarding
40 the exclusion of particular studies.
- 41
- 42 • EPA’s Consideration #2 was worded awkwardly and misspecified epidemiologic terms.
43 For instances, define susceptible to important biases which is a qualitative term. Does the
44 text ‘control for potential confounding exposures’ refer only to exposures (such as DLCs)
45 or was it meant to more broadly refer to other exposures as well (NIOSH cohort studies)?
46 This should be clarified. Does the text ‘bias arising from study design’ refer to selection

1 bias or is it used more broadly for how exposure and outcome are measured and covariate
2 data collected? Define what is meant by ‘bias arising from statistical analyses’? It is
3 unclear if bias is the correct term, rather this may refer to model misspecification.
4

- 5 • EPA’s Consideration #3: does the way this is worded preclude inclusion of null
6 epidemiologic studies? There needs to be more discussion and clarity on the exclusion of
7 null epidemiologic studies (for instance for the non-cancer thyroid outcome).
8
- 9 • Inclusion criteria 3: Define ‘reported dose’.
10
- 11 • An expanded discussion on suitability for inclusion of related studies on thyroid and
12 diabetes should be considered. Material in the appendix suggests that the lack of an
13 animal model for diabetes and paucity of published dose response data precludes its
14 inclusion. Are there other reasons that this was not considered a primary endpoint?
15 Thyroid homeostasis is complicated with effects in adults with occupational exposure not
16 clearly understood. I think perhaps a more thorough discussion of these issues may be
17 warranted if thyroid effects in newborns are to be highlighted. (Calvert 1999, Steenland
18 2001)
19
- 20 • ML: An explanation for what is meant by ‘consistent toxicological practices’ and ‘outside
21 normal range of variability’ is needed. The former, however, may have been identified in
22 the earlier exclusion criteria. Number 4 criteria is particular difficult to understand since
23 there are no accepted normal ranges for most biomarkers in animal studies (differences
24 are statistically based) plus a small effect in a key clinical endpoint can be potentially
25 more adverse than a large effect in another clinical marker.
26
- 27 • Regarding non-cancer candidate PODs (Figure ES-4): Pg xxxviii (lines 14-16). This
28 should explain the statement ‘that BMDL modeling was largely unsuccessful due to data
29 limitations’ given the number of animal studies that are available. Maybe provide some
30 general examples of major data limitations. Was this due the fact that the BMDL was at a
31 much lower dose than the LOAEL or were there other reasons?
32
- 33 • Pg xxxvii, Lines 16-19 sentence needs clarification. It sounds like those studies that were
34 eliminated for further analysis would have NOAELs available
35
- 36 • Add an extra column in Table 2-7 listing by number reference the criteria that were or
37 were not met for each study.
38

39 **2.3 Has EPA applied the epidemiology and animal bioassay study criteria and**
40 **considerations in a scientifically sound manner? If not, please identify and provide a**
41 **rational for alternative approaches.**
42

43 Response. In general, the SAB considered that the inclusion criteria for data set selection for
44 dose-response analyses are scientifically sound and well justified. Not only are overarching
45 criteria presented, but there is specific discussion of key data sets describing their shortcomings
46 and justification for exclusion or even inclusion. There seems to be a varying acceptance of

1 DLCs in determining relevant data. ON balance, SAB members expressed concern that the EPA
2 report overstates its ability to use “TCDD-only” as a data set selection criteria, and that DLCs
3 should be used to the extent possible to bolster the weight of evidence regarding assessment of
4 dioxin toxicity.

5
6 The discussion of whether certain end points that are measured represent adverse health
7 effects and the elimination of studies because the authors reject this is in some cases arguable.
8 For example, exclusion of the Sugita-Konishi, 2003 study is justified because the linkage
9 between TCDD and immune function in this case is not demonstrated or clear. Similarly, the
10 2009 ANL-EPA meeting recommended using two other immune studies (one by De Vito, the
11 other from Seveso), which seem to have been ignored because a reduction in IgG is not
12 considered an adverse health effect (the same for thymic atrophy, since the thymus normally
13 atrophies in development).

14
15 (AS) Thyroid homeostasis is a difficult endpoint to consider, for choosing a regulatory level.
16 The most serious consequences of abnormal thyroid circuit activity/production probably happens
17 in utero, and while indication of abnormal levels can be done in Italy due to a public health
18 program that takes heel sticks for thyroid at the start of life. It is unclear that measurements of
19 THS/FT4 etc. are clearly understood as adverse effects.

20
21 (AS) Some further review of this rejection of diabetes and assorted immunological endpoints vs.
22 utilizing thyroid endpoints might be in order. I do not think that inclusion will affect the
23 BMD/RfD determinations by more than an order of magnitude (and probably much less), but the
24 stronger the weight of evidence for the numbers derived, the better for acceptance.

25
26 Recommendation:

- 27
28 • As noted above, data set selection could be further justified by editing the text. For
29 example, edits could be made to make it clearer to readers why certain studies were
30 excluded. To be clear, this suggestion does not mean a different approach is needed, but
31 that the approach used should be explained more effectively and clearly.
32
33 • Two criteria require further clarification. These are (1) confounding and other potential
34 sources of bias, and (2) statistical precision, power, and study power. For some studies,
35 these criteria may not have been consistently applied. Specific examples:

- 36
37 (1) Confounding and other potential sources of bias are addressed: The differences
38 between males and females with regard to TCDD half-life are discussed, but
39 the description of the number of males and females in each study population
40 are often missing or very difficult to track down. Also, in the occupational
41 cohort studies, the possibility of men and women performing different job
42 tasks also increases the possibility that the men and women were exposed at
43 different levels. However, when the job categories with assigned TCDD
44 exposure levels are presented, there is often no discussion of the numbers by
45 gender in the categories. For example, the Manz et al. study (1991) of the
46 Hamburg cohort (1,583 men and 399 women) does not describe the TCDD

1 categories by gender. In addition, the validity of the TCDD exposure levels
2 assigned to the categories was examined “in a group of 48 workers who
3 provided adipose tissue samples.” (Page 2-41, lines 18-19). How were these
4 workers selected? How many were approached but refused to provide a
5 sample? Assessment of selection bias in this and other similar circumstances is
6 lacking in some of the studies. This is particularly notable in the lack of
7 overall response rates reported for several of these studies. Inclusion of these
8 factors in the study review would be very helpful.
9

- 10 (2) Statistical precision, power, and study follow-up are sufficient: This can be
11 difficult to determine with the smaller sample size populations, but there are
12 studies that can be very useful even given the small samples. For example,
13 the relative risks calculated for increasing TCDD exposure and risk of breast
14 cancer in the Seveso study were greatly increased in the 3rd and 4th highest
15 exposure categories, but the RRs were not statistically significant (page 2-56,
16 lines 1-8). However, as the EPA document states: “Although statistical
17 significance was not achieved for either category, likely because of the small
18 number of cases, the greater than three-fold risk evident in both categories is
19 worth noting.” This needs to be kept in mind for additional evaluations in
20 other studies as well.
21
22
23
24
25
26
27

1 **Section 3 – The Use of Toxicokinetics in the Dose-Response Modeling for**
2 **Cancer and Noncancer Endpoints.**

3
4 *Drs. Fisher, Rozman*

5
6 *3.1 The 2003 Reassessment utilized first-order body burden as the dose metric. In the draft*
7 *Response to Comments document, EPA used a physiologically-based pharmacokinetic (PBPK)*
8 *model (Emond et al., 2004, 2005, 2006) with whole blood concentration as the dose metric*
9 *rather than first-order body burden. This PBPK model was chosen, in part, because it includes a*
10 *biological description of the dose-dependent elimination rate of TCDD. EPA made specific*
11 *modifications to the published model based on more recent data. Although lipid-adjusted serum*
12 *concentrations (LASC) for TCDD are commonly used as a dose metric in the literature, EPA*
13 *chose whole blood TCDD concentrations as the relevant dose metric because serum and serum*
14 *lipid are not true compartments in the Emond PBPK models (LASC is a side calculation*
15 *proportional to blood concentration).*

16
17 *Please comment on:*

18
19 *3.1.a. The justification of applying a PBPK model with whole blood TCDD concentration as a*
20 *surrogate for tissue TCDD exposure in lieu of using first-order body burden for the dose-*
21 *response assessment of TCDD.*

22
23 The use of body burden in the 2003 Reassessment represented an improvement over the usual
24 default metric of administered dose (mg/kg/d), because the default metric would not properly
25 reflect the accumulation of dioxin in the tissues over time. However, because the accumulation
26 of dioxin in liver is dose-dependent, body burden would not serve as a direct surrogate for tissue
27 exposure. The use of whole blood concentration is a better choice than body burden, because it
28 is more closely related to the biologically relevant dose metric: the free concentration of dioxin
29 in the target tissues (liver, fetus, etc.). Blood concentrations are routinely used to estimate
30 biologically effective exposures for pharmaceuticals.

31
32 The rationale for the use of whole blood concentration rather than lipid adjusted serum
33 concentration (LASC) should not be based on the Emond model structure. It would be trivial to
34 change the model so that LASC could be predicted. Indeed, the model is apparently used to
35 estimate LASCs in the RfD calculations (e.g., p. xli, line 21). The question that should be
36 addressed is only whether whole blood concentrations or LASCs provide better surrogates for
37 cross-species and cross-study comparisons of free dioxin concentration in the target tissues.
38 LASC is the preferred measure for reporting dioxin biomonitoring data, and is the measurement
39 reported in most of the human epidemiological studies. A metric that considers blood lipid
40 content is also more likely to reflect free dioxin concentration in the plasma, and hence free
41 concentration in the target tissue. The EPA points out (p. xxxiv) that the LASC is related to the
42 whole blood concentration by a scalar; however, they incorrectly conclude that the metrics are
43 equivalent. They later (p. 3-511, line 6) discuss the fact that the relationship between them is
44 subject to inter-individual and inter-species variation. If the LASC is used to drive the
45 distribution of TCDD to tissues, the pharmacokinetic outcome would be different than whole

1 blood as the driver because the tissue:blood ratio would differ. If the blood fat:blood and tissue:
2 blood values are accounted for in the model the use of whole blood and LASC would be similar.

3
4 It's not clear at this point how this issue is addressed in the dose metric calculations.
5 Consideration of this issue appears unlikely to drastically affect the outcome of the risk
6 calculations, but it would be important for a quantitative uncertainty analysis.

7
8 **Recommendation: Whole blood metric is the PBPK model is ok.**

9
10 *3.1.b. The scientific justification for using the Emond et al. model as opposed to other available*
11 *TCDD kinetic models.*

12
13 The Emond model provides the best available basis for the dose metric calculations in the
14 assessment. It is the product of a high-caliber, multi-year research effort at EPA/NHEERL led
15 by Linda Birnbaum and Mike Devito, and represents a significant effort in terms of data
16 collection. This model builds on prior PBPK modeling efforts conducted by Drs. Andersen and
17 Clewell. However, additional discussion of other published models and quantitative evaluation
18 of the impact of model selection on dose metric predictions should also be provided.

19
20 **Recommendation: This discussion should address how the model is intended to be used in**
21 **the assessment, which would then dictate why a particular model was selected. That is, for**
22 **the intended purposes, was the Emond model more robust and/or simpler than other**
23 **models, and did it contain sufficient details for biological determinants deemed important**
24 **by the Agency.**

25
26 *3.1.c. The modifications implemented by EPA to the published Emond et al. model.*

27
28 The EPA modifications (p. 3-44, account for volume of plasma and describe urinary clearance
29 using blood concentration and not a lumped compartment) are minor and appropriate.

30 **Recommendation: Model changes are fine.**

31
32 *3.1.d. Whether EPA adequately characterized the uncertainty in the kinetic models.*

33
34 The EPA document presents a reasonably thorough qualitative characterization of the uncertainty
35 in the kinetic models, sufficient to support their use in the assessment. A more quantitative
36 uncertainty analysis is needed, using Monte Carlo techniques (as in the vinyl chloride IRIS
37 Technical Support Document). It is critical to demonstrate the dependence of human HED and
38 risk predictions on uncertainty and variability in the model parameters, particularly those with
39 high sensitivity (Evans and Andersen, 2000). Moreover, dose metric uncertainty needs to be
40 determined under the same exposure conditions that dose metrics are calculated: both for the
41 various studies that serve as the basis for the dose-response assessments and for human
42 exposures at the corresponding HEDs and risk specific doses.

43 The Hill coefficients for cyp1a1 and cyp1a2 induction used in the Emond model are 1.0 and 0.6,
44 respectively, based on fitting of kinetic data from single doses of dioxin (Wang et al, 1997 and
45 Santostefano et al 1998). However, Walker et al (1999) subsequently estimated a Hill coefficient
46 of 0.94 for both cyp1a1 and cyp1a2 induction using chronic exposures which are more relevant

1 to the use of the Emond model in the dioxin risk assessment. The value of 0.6 used in the
2 Emond model is well outside the confidence interval of 0.78 to 1.14 reported by Walker et al
3 (1999). The use of a Hill coefficient value well below unity leads to a nonlinear model behavior
4 that is biologically implausible (hypersensitivity to induction at doses near zero). As a result,
5 when the human model is used for extrapolation to lower doses (as in the calculation of risk-
6 specific doses) the model will tend to estimate a lower exposure level for a given blood
7 concentration. This effect can be seen in Table ES-1 of the EPA response document, where a 5
8 order-of-magnitude change in risk is associated with a 6 order-of-magnitude change in risk
9 specific dose. That is, the model-estimated risk specific doses in the vicinity of 10⁻⁶ risk are
10 about a factor of 10 lower (more conservative) than linear extrapolation. The evidence for this
11 parameter needs to be carefully reviewed and the reasonable range of values determined. At the
12 least, the human Emond model calculations will need to be repeated with multiple values to
13 characterize the resulting uncertainty in the estimates. When this is done, the agency should also
14 consider increasing the fat:blood partition in the human model from 100 to 200 to be more
15 consistent with the human data (Patterson et al 1988, Iida et al 1999, Maruyama et al 2002). The
16 Hill coefficient is not likely to have as significant an effect on calculations with the animal
17 models, since low-dose extrapolation is not performed in the animals, but this should also be
18 verified by sensitivity/uncertainty analysis of the animal models.

19

20 **Recommendation: We recommend additional efforts to fully characterize the uncertainty**
21 **in the models.**

22

23 3.2. *Several of the critical studies for both noncancer and cancer dose-response assessment*
24 *were conducted in mice. A mouse PBPK model was developed from an existing rat model*
25 *in order to estimate TCDD concentrations in mouse tissues, including whole blood.*

26

27 *Please comment on:*

28

29 3.2.a. *The scientific rationale for the development of EPA's mouse model based on the*
30 *published rat model (Emond et al., 2004, 2005, 2006).*

31

32 An appropriate approach was used to develop the mouse model on the basis of the published rat
33 model and the available mouse kinetic data.

34

35 **Recommendation: An external peer review of the mouse model should be performed,**
36 **since this model has not been published in the peer-reviewed literature, which is typically a**
37 **requirement for models to be used by the Agency.**

38

39 3.2.b. *The performance of the mouse model in reference to the available data.*

40

41 The mouse model performs reasonably well, apart from under-prediction of urinary excretion
42 data. The urinary excretion data can be improved by taking into account that urine contains
43 metabolites only, which partition differently from the parent compound. The model appears to be
44 adequate for use in estimating dose metrics for the assessment, but with greater uncertainty than
45 the rat and human models. This is considered a reasonable approach to solve a deficiency in
46 published PBPK models to meet the needs of this assessment.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

Recommendation: The EPA’s suggestion in the RfD chapter that the clustering of mouse PODs at the lowest doses is due to mouse model failure is inappropriate and should be rewritten.

3.2.c. Whether EPA adequately characterized the uncertainty in the mouse and rat kinetic models. Please comment specifically on the scientific justification of the kinetic extrapolation factor from rodents to humans.

The EPA provides an adequate characterization of the qualitative uncertainty in the mouse and rat kinetic models, sufficient to justify their use, together with the human model, to estimate rodent-to-human extrapolation factors. On the other hand, formal recalibration of the PBPK model parameters using a Hierarchical Bayesian approach such and Markov chain Monte Carlo analysis is not considered necessary or particularly useful.

Recommendation: A more quantitative uncertainty analysis is needed, using Monte Carlo techniques (as in the vinyl chloride IRIS Technical Support Document) to estimate the propagation of uncertainty from the PBPK model parameters to the dose metric predictions.

3.3 Please comment on the use of Emond et al. PBPK model to estimate human intakes based on internal exposure measures.

The modified Emond model is the best available approach for estimating exposures on the basis of internal exposure measurements. Nevertheless, there is considerable uncertainty associated with attempting to reconstruct prior exposures in a human population (e.g., Seveso).

Recommendation: The modeling of the Cheng, Moccarelli, and Bacarelli studies needs to be described in more detail and the impact of model parameter uncertainty and exposure uncertainty in these studies should be evaluated quantitatively.

3.4 Please comment on the sensitivity analysis of the kinetic modeling (see Section 3.3.5).

The EPA document only presents the sensitivity analysis published by Emond et al. 2006, which is not entirely adequate for the purposes of this assessment. It leaves out the Hill coefficient, which is one of the most important parameters in the model for low-dose extrapolation (Evans and Andersen, 2000). Moreover, model sensitivities are species, dose, and dose-scenario dependent, so they need to be determined under the same exposure conditions that dose metrics are calculated: both for the various studies that serve as the basis for the dose-response assessments and for human exposures at the corresponding HEDs and risk specific doses.

Recommendation: Provide a sensitivity analysis of the model to authentication the model for its intended purpose.

1 *3.5 Both EPA's noncancer and cancer dose-response assessments are based on a lifetime*
2 *average daily dose. Did EPA appropriately estimate lifetime average daily dose? If not, please*
3 *suggest alternative approaches that could be readily developed based on existing data.*
4

5 We agree with the average daily dose calculation approaches described in the EPA document.
6

7 **Recommendation: The predictions of the model in the perinatal period need to be re-**
8 **evaluated for the possibility that the change in exposure associated with birth might lead to**
9 **transient changes in peak blood concentration.**

10

11

12 NOTE:

13 It should be noted that the NAS recommendation to use human data for dose metric, which could
14 be done because dose-dependent elimination of TCDD has been described in humans, albeit in
15 just a few cases. Dose –dependent elimination has been reported repeatedly in animals and the
16 model reflects that. Using CYP1A2 data from humans (caffeine metabolism) and mice offers an
17 opportunity to validate and/or adjust the mouse model.
18

1 **Section 4 – Reference dose**

2
3 *Drs. Luster, Petersen, Silverstone, Sweeney*

4
5 **Question 4.1** *Is the rationale for the choice of Mocarelli and Baccarelli scientifically justified*
6 *and clearly described? Please identify and provide the rationale for any other studies that should*
7 *be selected, including the rationale for why the study would be considered a superior candidate*
8 *for the derivation of the Rfd. Also comment on whether the selection of male reproductive effects*
9 *and changes in neonatal thyroid hormone levels was scientifically justified and clearly*
10 *described.*

11
12 In general there was consensus for the use of the Mocarelli et al 2008 and Baccarelli et al 2008
13 studies as identifying “co-critical” effects for the RfD calculation. These are human
14 epidemiological studies that were well thought out and designed. The endpoints of changes in
15 sperm count and TSH levels are of public health relevance and therefore of interest for
16 determining an RfD.

17
18 Collectively, there was support for these endpoints within the context of the broader dioxin
19 literature. There was discussion on whether the magnitude of these changes would represent an
20 adverse health effect. The committee discussed that the shifts observed in TSH levels and sperm
21 counts may or may not pose a significant health effect in a single individual, but a shift on a
22 population basis would constitute potential adverse health outcomes, such as hypothyroidism or
23 reduced reproductive function.

24
25 Some of the strengths of the human studies included the use of a well characterized human
26 cohort, assessment by dioxin epidemiology experts and the fact that similar PODs were found
27 across a broad spectrum of other reported dioxin toxicities in multiple species. However in
28 isolation from each other or from a lack of a consistent signal from the supportive animal and
29 epidemiological studies they are less useful for setting RfD. **The committee emphasized to**
30 **EPA the need to think of these within context of the weight of the database on TCDD.**

31
32 A strong voice from the committee was given for looking at the comprehensive data base of both
33 animal and human epi studies together due to a consistent and integrative signal of toxicity
34 across species and endpoints for TCDD. This “collective” impact of the studies is stated in the
35 document but needs to be made stronger as it represents the contextual framing for understanding
36 dioxin health impacts. For example the type and dose-response relationships for dioxin would
37 strengthen if EPA would include more studies – including studies that used DLCs in their test
38 mixtures. The strength of the RfD should not be based solely on these two human epidemiology
39 studies but rather should be supported by integration with other similar supporting Dx and DLC
40 studies.

41
42 **The committee recommends the EPA authors strengthen the rationale for their selection of**
43 **these two studies by providing a better description of both the strength and weaknesses.**
44 **They should also discuss these studies in the context of other animal and human Dx and**
45 **DLC studies with comparable endpoints.**

1 Although there was concern expressed on the sample size for sperm number and
2 known variability in the biological endpoint, the sample collection was conducted
3 consistently across subjects and the difference in groups apparent. Regarding neonatal
4 TSH levels, the document could better describe the consequences of transient
5 hypothyroidism (e.g., see Anbalagan, J., A. M. Sashi, et al. (2010). "Mechanism
6 underlying transient gestational-onset hypothyroidism-induced impairment of
7 posttesticular sperm maturation in adult rats." *Fertility and Sterility* 93(8): 2491-2497.)
8

9 Numerous times the committee referred to Figures 4.3 and 4.4 that showed quantitative
10 comparisons across the RfDs and BMDLs calculated from the animal and epidemiological
11 studies as being useful in understanding the quantitative similarities in these calculations. The
12 committee also noted that since this figure did not have an indication of endpoints the
13 consistency in signal was not as readily apparent as it could be. **The committee encourages
14 EPA to make this more explicit in the figure and supportive text.**
15

16 **Question 4.2.a.i** Please comment on EPA's approach for identifying the exposure window and
17 calculating average exposure for this study.

18 The committee discussed extensively both as part of the deliberations on Section 4 but also as
19 part of the discussion on section 3 that the pattern of exposure from Seveso poses some
20 extrapolation issues for the EPA. Issues raised include the question whether the same endpoints
21 and or dose response would be expected from such exposure scenarios with high acute exposures
22 when extrapolating to low-dose chronic exposures
23

24 **Question 4.2.a.ii** Please comment on EPA's designation of a 20% decrease in sperm count (and
25 an 11% decrease in sperm motility) as a LOAEL for Mocarelli et al. (2008).
26

27 There was also general support for EPA's approach to use the WHO reference value for
28 determining TSH levels **and there was strong suggestions that further discussion on WHO
29 reference values for male reproductive parameters should be included if available.**
30

31 **Question 4.2.b.i** Please comment on EPA's decision to use the reported maternal levels and the
32 appropriateness of this exposure estimate for the Baccarelli et al. study.
33

34 .
35 The group discussed and generally supported the EPA's decision to use the Baccarelli et al
36 estimates of the relevant effective doses however additional discussion from the reviewers for
37 the kinetics section is needed in order to respond to this part of the question.
38

39 **Question 4.2.b.ii**
40

41 **Question 4.3** Please comment on the rationale for the selection of the uncertainty factors (UFs)
42 for the RfD. If changes to the selected UFs are proposed, please identify and provide a
43 rationale.
44

1 The committee agreed with the EPA that the appropriate UFs were included but suggested that
2 **EPA provide justification for not including an UF for data quality for the two Seveso**
3 **studies.**

4
5 **Question 4.4** *Please comment on whether the decision not to consider biochemical endpoints is*
6 *scientifically justified and clearly described.*

7 In general terms P450 activation, increased oxidative stress and changes in certain other
8 biochemical endpoints may be used to establish PoDs, particular when the quantitative
9 relationship with an adverse outcome is available. However, with respect to TCDD we agree
10 with EPA that other endpoints are more appropriate as these associations with health outcomes
11 are more clear.

12
13 **Question 4.5** *Please comment on EPA's approach for averaging exposures including*
14 *intermittent and one-day gestation exposure protocols.*

15
16 For animal studies acute exposure could give different results than from chronic exposure. For
17 TCDD, however, it's persistence will negate some of these potential differences. In Baccarelli et
18 al (2008) there was extensive discussion regarding the use of the exposure average time for the
19 TCDD concentrations. This is of biological significance as several papers have indicated that the
20 unique aspects of high peak exposure of TCDD as occurred in Seveso and in several of the
21 animal studies. The endpoints affected as a result from these peaks does not always translate to
22 impacts from lower chronic exposures.

23 **Two considerations to address this issue are offered – first, conduct a series of sensitivity**
24 **analyses to evaluate the impact of averaging time on the RfDs and second, return to the**
25 **broader animal literature with DLCs to see if biological support for the two critical**
26 **endpoints could be added. Time and dose-response studies from the broader DLC**
27 **literature could be informative**

28
29 **Question 4.6** *Please comment on the benchmark dose (BMD) modeling conducted by EPA to*
30 *analyze the animal bioassay data and EPA's choice of points of departure (PODs) from these*
31 *studies.*

32 In general the committee's limited discussion would suggest agreement with the BMD modeling
33 approaches used in this section for these two endpoints. **As indicated previously the EPA**
34 **authors need to more specifically cite the endpoint guidance that is present within EPA**
35 **documents for defending these approaches and application of BMD models for the critical**
36 **effects.** Expanded discussion on known human variability in conducting sperm counts and
37 neonatal TSH levels would be helpful.

38
39 **Question 4.7** *Please comment on whether the kinetic extrapolation at the level of the POD prior*
40 *to applying the uncertainty factors was scientifically justified and clearly described.*

41 The approach of EPA to apply the kinetics on the actual data present at the POD is preferred in
42 this assessment.

43
44 **Question 4.8** *Please comment as to whether EPA's qualitative discussion of uncertainty in the*
45 *RfD is justified and clearly described.*

46 Discussed in Section 6.

1
2 **Section 5 – Cancer assessment**

3
4 *Drs. Clewell, Hakansson, Persky*

5
6 In general, panel members were impressed by the extensive work presented by the Agency in
7 their response to the NAS comments on cancer assessment. Comments below are supporting the
8 Agency in further developing section 5 and to transfer some of its contents to other sections of
9 the draft.

10
11 **5.1. Weight of Evidence Cancer Descriptor:** The 2003 Reassessment concluded that TCDD is
12 a “known human carcinogen.” In the current draft Response to Comments document, EPA
13 concluded that under the 2005 Guidelines for Carcinogen Risk Assessment (U.S. EPA, 2005)
14 TCDD is “carcinogenic to humans.” Is the weight-of-evidence characterization scientifically
15 justified and clearly described?

16
17 Comment:

- 18
19
 - Panel members agreed on the classification that “TCDD is carcinogenic to humans”.

20
21 Recommendations:

- 22
23
 - The agency should provide more discussion of the power of studies used and the
24 difficulties involved when assessing rare tumors. Thoroughly addressing these aspects
25 will make the weight of evidence characterization in this section more clear and
26 transparent.
 - In the weight-of-evidence characterization, the agency should build on all the available
27 data to support the decision. It needs to be made clear how different types of data (in
28 vitro, in vivo, human) support each other; or not.
 - The agency should consider including studies with substantial DLC exposure where
29 TEFs can be calculated.
 - The agency should attempt to characterize the uncertainty regarding the carcinogenicity
30 of TCDD at low human exposures, since the minimum dose at which carcinogenic effects
31 would be expected to occur cannot be clearly delineated from the current epidemiological
32 human data. The agency has concluded that AhR activation is a necessary but not
33 sufficient precursor event in the carcinogenic activity of TCDD. Therefore, it would be
34 beneficial if the agency could evaluate available data on AhR activation and related
35 effects in human cells and animal models to help inform the doses at which these
36 precursor events are observed for comparison with the epidemiological data.

37
38 **5.2 Mode of Action:** The mode of action of a carcinogen can inform identification of hazards
39 and approaches used for a dose-response assessment. The mode of carcinogenic action for
40 TCDD has not been elucidated for any tumor type. EPA concluded that, while interaction with
41
42
43

1 the Ah receptor is likely to be a necessary early event in TCDD carcinogenicity in experimental
2 animals, the downstream events involved are unknown.

3
4 5.2.a Are the available data related to mode(s) of action for the carcinogenicity of TCDD
5 appropriately characterized and clearly presented?
6

7 Comments:
8

- 9
- 10 • Panel members appreciated the attempts by the Agency to further develop cancer mode-
11 of-action concepts based on available dioxin liver, lung, and thyroid toxicity data. Such
12 innovative and explorative work is clearly fundamental to the continued need of further
13 developing risk assessment sciences and to make more detailed and integrated use of
14 already existing and published data.
 - 15 • Panel members applaud the agency for providing an up-to-date dioxin cancer mode-of-
16 action section in its response to NAS comments. It could, however, be improved by
17 incorporating additional data on linear and nonlinear modes of action in different target
18 tissues and life stages.
19

20 Recommendation:
21

- 22 • The agency should further expand the discussion of mode of action data available to
23 delineate linear versus nonlinear modes of action and effects in different target tissues at
24 different life stages.
25

26 5.2.b. Do the available data support EPA's conclusion that the overall mode(s) of action for
27 TCDD-induced carcinogenesis is largely unknown? Please comment on whether this evaluation
28 is clearly described.
29

30 Comments:
31

- 32 • Panel members pointed out that much is known about TCDD toxicity and mode-of-
33 action. Some panel members felt that the characterization should be "reasonably well
34 known" rather than "largely unknown." Nevertheless, the panel agrees that the exact
35 mechanism-of-action has not been fully delineated for any distinct TCDD-toxicity end-
36 point.
37
- 38 • A large amount of data related to the mode of action for the carcinogenicity of TCDD is
39 described, but the focus appears to be on presenting evidence that supports the use of a
40 default linear approach rather than providing a balanced evaluation of alternative mode-
41 of-action hypotheses.
- 42 • The discussion of the likely dose-response for receptor mediated processes focuses only
43 on the first step, binding of the agonist to the receptor, which is ultimately linear at low
44 concentrations. However, no discussion is given to the nature of the dose-response for
45 the down-stream sequelae of receptor activation, for which there is evidence of
46 nonlinearity. It is, in fact, the fundamentally nonlinear nature of the dose-response for

1 receptor mediated processes that underlies the conviction of a large segment of the
2 scientific community, that a nonlinear approach should be preferred for the risk
3 assessment for dioxin.
4

5 Recommendation:

- 6
- 7 • The agency should provide a balanced discussion of the evidence for possible modes of
8 action, including both linear and nonlinear alternatives
9
- 10 • The description of the nature of a receptor mediated dose-response needs to be expanded
11 by including more evidence regarding the nonlinearity of the receptor mediated dose-
12 response for dioxin (e.g., Andersen et al 1997).
13

14 **5.3 Data selection.** Is EPA's approach for selecting data sets from the key epidemiologic studies
15 and animal bioassays identified for cancer dose response modeling scientifically justified and
16 clearly described?
17

18 Comments:

- 19
- 20 • The panel agrees with the inclusion of the Cheng study, which incorporates information
21 on gradation of exposure.
22
- 23 • Expanded discussion of several other studies would support the weight of evidence for
24 carcinogenicities in less common cancers such as lymphomas and soft tissue sarcoma.
25
- 26 • Panel members discussed the possible value of including studies with DLCs in the
27 evaluation of the weight of evidence, in light of the small number of studies involving
28 primary exposure to TCDD.
29

30 Recommendation:

- 31
- 32 • The agency should present in a clear and visible format, for example in a table, which
33 studies were carried forward or not, and the reasons for the decisions made. The weight
34 of evidence discussion should be expanded to include evidence from studies of individual
35 cancers for which precise gradation of exposure data is lacking.
36

37 **5.4 Animal bioassay data.** For the animal bioassay data, potential cancer oral slope factors
38 (OSFs) were calculated by linear extrapolation (using a linear, non threshold cancer approach)
39 from the point of departure (POD). EPA also estimated the composite risk of the occurrence of
40 several tumor types from the animal cancer bioassay data.
41

42 5.4.a. Please comment on whether the approach for estimating cancer risk, including the use of
43 tumor modeling of the TCDD animal cancer bioassay data, is scientifically justified and clearly
44 described.
45

46 Comment:

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

- The panel agreed that the approach for estimating cancer risk from animal studies was scientifically justified and clearly described.

5.4.b. Please comment on the choice of using a BMDL01 as the POD for the development of candidate oral slope factors derived from the TCDD animal cancer bioassays.

Comment:

- Panel members noted the consistency of the selection of the BMDL01 as the POD with agency guidelines and had no further comments.

5.5 Cheng et al. EPA selected Cheng et al. (2006) – an analysis of the NIOSH occupational cohort – as the critical study for oral slope factor (OSF) development. This study was chosen because it considers dose-dependent elimination of TCDD rather than first-order kinetics.

5.5.a. Please comment on whether the rationale for this selection is scientifically justified and clearly described. Please identify and provide the rationale for any other studies that should be considered and provide a critical evaluation of the study and of its suitability for meeting the goals of a quantitative cancer assessment.

Comment:

- Panel members agreed that Cheng et al is the appropriate study, and the selection of this study is well described.

5.5.b. Cheng et al. (2006) analyzed all-cancer mortality. Please comment on the use of all-cancer mortality as the basis of the OSF.

Comment:

- Panel members agreed that it is appropriate to use all-cancer mortality in this case, because of the extensive dose-response information.

5.5.c. Please comment on whether the use of the Emond PBPK model in the estimation of risk-specific doses from the Cheng et al. dose-response modeling results is scientifically justified and clearly described.

Comment:

- Panel members agreed that the use of the Emond model is scientifically justified and clearly described.

1. 5.5.d. EPA elected to use the log linear relationship of fat concentration and rate ratio to estimate risk-specific doses at all risk levels. EPA could have estimated a POD for cancer

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

Comment:

- Panel members agreed that the Agency has chosen the appropriate extrapolation model and that using the oral slope factor to arrive at the POD was correctly done.

5.5.e. The slope factor derived from Cheng et al. (2006) was extrapolated below the background TCDD exposure levels experienced by the NIOSH cohort. Please comment on this extrapolation.

Comment:

- The ability of the Cheng study to be informative regarding risks below current background exposure levels is not completely clear.

Recommendation:

- The agency should expand the discussion to consider the possibility that mode of action considerations could help to inform whether linear extrapolation of the Cheng data to obtain risk estimates in this range of exposures is appropriate.

5.6 OSF derivation. Please comment on whether EPA has clearly described the major qualitative uncertainties in the derivation of the OSF.

Comment:

- The panel found the description of qualitative uncertainties in the derivation of the OSF to be clear and adequate.

5.7 DLCs. EPA did not consider dioxin-like compounds (DLCs) in the cancer dose-response modeling because the occupational exposures in the available cohorts were primarily to TCDD. Background DLC exposures were not incorporated in the dose-response modeling because EPA judged that it was not possible to disaggregate the responses from background exposure to DLCs and occupational exposure to TCDD. Please comment on whether this approach is scientifically justified and clearly described.

Comments:

- While the panel members felt it was important to include DLC studies in the weight of evidence analysis, they were conflicted on their use as a source of dose-response estimates for TCDD.
- Several panel members pointed out the scientific importance and regulatory relevance of including a coordinated TEQ/DLC-discussion in the response. Including TEQ/DLC-

1 aspects in the evaluation would allow for the use of additional studies with dose-response
2 information that more closely mirror environmental exposures.

- 3
4 • On the other hand, the panel recognizes the complications associated with developing a
5 TCDD risk estimate that is dependent on current TEF values.
6

7 Recommendation:

- 8
9 • DLC studies should be considered in the weight of evidence discussion.
10

11 **5.8 Non-linear approach.** The NRC suggested that EPA consider nonlinear approaches for the
12 assessment of TCDD carcinogenicity. In the Response to Comments, EPA presents two
13 illustrative nonlinear approaches for cancer, but considers both inappropriate to use because lack
14 of MOA information.
15

16 5.8.a. Please comment on these two illustrative nonlinear approaches including EPA's
17 conclusions regarding the limitations of these approaches.
18

19 Comments:

- 20
21 • The EPA document does not respond adequately to the NAS recommendation to
22 adopt "both linear and nonlinear methods of risk characterization to account for
23 the uncertainty of dose-response relationship shape below the ED01." Instead of
24 adopting both linear and nonlinear methods, the EPA argues that only a linear
25 approach can be justified, and derives two examples of RfD development using a
26 nonlinear approach that they characterize as an illustrative exercise only.
27

- 28 • The choice not to include both linear and nonlinear risk assessment approaches
29 for TCDD is inconsistent with the EPA (2005) cancer guidelines (p.3-23/24):
30

31 "Nonlinear extrapolation having a significant biological support may be presented
32 in addition to a linear approach when the available data and a weight of evidence
33 evaluation support a nonlinear approach, but the data are not strong enough to
34 ascertain the mode of action applying the Agency's mode of action framework."
35

36 "In the absence of data supporting a biologically based model for extrapolation
37 outside of the observed range, the choice of approach is based on the view of
38 mode of action of the agent arrived at in the hazard assessment. If more than one
39 approach (e.g., both a nonlinear and linear approach) are supported by the data,
40 they should be used and presented to the decisionmaker."
41

42 Recommendation:

- 43
44 • The EPA should present both linear and nonlinear risk assessment approaches. They can
45 still conclude that EPA policy dictates that, in the absence of a definitive nonlinear mode

1 of action, the linear option should be preferred in order to assure protection of the public.
2 The examples in the current document should be formalized and extended.

3
4 5.8.b. Are there other nonlinear approaches that could be readily developed based on existing
5 data for the assessment of TCDD carcinogenicity? If so, please suggest alternative approaches
6 and describe their utility and suitability for meeting the goals of a quantitative cancer assessment.

7
8 Recommendation:

- 9
- 10 • Since the EPA nonlinear analysis only used studies in S-D rats that were
11 identified in Section 2 for potential noncancer dose-response modeling, additional
12 alternative PODs should be added. For example, Simon et al (2010), which is
13 cited in the EPA document, provides a number of alternative PODs for a
14 nonlinear approach that should be included in the EPA risk assessment.

1 Section 6 – Uncertainty analysis

2
3 *Drs. Ferson, Cox, Small*

4 5 **Question 6.1**

- 6 • Chapter 6 is clearly presented, but not scientifically justified.

7 8 **Question 6.2**

- 9 • Quantitative uncertainty analysis (QUA) is possible.
- 10 • There are several ways one could do QUA without expert elicitation. These include
 - 11 ○ Probability tree (model choice tree, similar to Sielken’s “comprehensive
 - 12 realism”),
 - 13 ○ Sensitivity studies, even if not completely comprehensive,
 - 14 ○ Bounding approaches, such as
 - 15 ▪ **Interval analysis** (Moore 1966; Neumaier 1990) which has been widely
 - 16 used for decades and can be applied to complex models and even blackbox
 - 17 models (Kreinovich and Trejo <<>>),
 - 18 ▪ **Nesting of intervals** (an approach which philosophers sometimes call
 - 19 “supervaluation” in the sense of van Fraassen),
 - 20 ▪ **Probability bounds analysis** (Ferson and Long 1995; Ferson 2002;
 - 21 Ferson et al. 2003) including Bayesian p-boxes (Montgomery 2009),
 - 22 which has been used in a variety of applications (Aughenbaugh and
 - 23 Paredis 2007; Dixon 2007; Karanki et al. 2009; Minnery et al. 2009;
 - 24 Regan et al. 2002a; 2002b), including assessments at two Superfund sites
 - 25 (EPA 2007; 2002-2005),
 - 26 ▪ **Info-gap decision theory** (Ben-Haim 2006) which has been used in
 - 27 several applications (Davidovitch et al. 2009; Hall and Harvey 2009;
 - 28 Regan et al. 2005; Rout et al. 2009; Yokomizo 2009), and
 - 29 ▪ **Robust optimization** (Bertsimas et al. 2009, 2010).
- 30 • Bounding analysis *is* an uncertainty analysis technique. At a minimum, EPA could
- 31 propagate simple bounds. Selecting precise probability distributions may be hard, but
- 32 ranges are easier.
- 33 • An appropriate QUA is possible, though EPA may decline to do one on other grounds.
- 34 • Value of information (Raiffa 1968) approaches should be used to clarify whether
- 35 modeling uncertainties and disagreements significantly affect risk estimates.
- 36 • Model uncertainty, including uncertainty about dependencies, can also be addressed with
- 37 the methods mentioned above.
- 38 • Validation, e.g., via a ‘reality check’ against the total number of cancers predicted versus
- 39 observed in a population, should be discussed in the chapter.
- 40 • Epistemic uncertainty (page 6-5) is not what the document says it is. Epistemic
- 41 uncertainty reflects imperfect knowledge, such as from limited data or imperfect causal
- 42 understanding about a system. It does not imply that a quantity about which there is
- 43 epistemic uncertainty is necessarily fixed.
- 44 • The word ‘exotic’ should be excised from the document. More generally, the tone of
- 45 chapter 6 seems condescending and should be strongly edited to be more neutral.
- 46

1 **Question 6.2a**

- 2 • Purge the document of the notion of ‘volitional uncertainty’. Display the different
3 modeling choices and the consequences of making them.
4

5 **Question 6.3**

- 6 • The sensitivities are useful and they provide part of the foundation for a QUA.
7
8
9

10 **Recommendation 6.1: Consider omitting or strongly revising chapter 6, particularly its**
11 **argument that quantitative uncertainty analysis is unfeasible for the dioxin assessment.**
12
13

14 **Recommendation 6.2: Reconsider the argument for not doing a quantitative uncertainty**
15 **analysis, or undertake one. EPA could follow NAS’ recommendation on this point by using**
16 **the techniques suggested above.**
17
18

19 **Recommendation 6.2a: Purge the document of the notion of ‘volitional uncertainty’.**
20 **Display the different modeling choices and the consequences of making them.**
21
22

23 **Recommendation 6.3: Keep and expand the sensitivity analyses.**
24
25

1
2 *6.1: Please comment on the discussion in this section. Is the response clearly presented and*
3 *scientifically justified?*

4
5 Chapter 6 is generally clearly presented, but it is not scientifically justified, although EPA's
6 decision to not do a QUA may be justified on grounds of practicality.

7
8 The EPA response is clearly presented, although one panel member felt that the whole section
9 should be rewritten to be accessible by non-statisticians. Some phrasing and words choices in
10 the text should be reconsidered, including 'exotic methods', 'volitional uncertainty', and
11 'epistemic uncertainty'. One panel member thought the definition of 'quantitative uncertainty
12 analysis' was overly narrow and should be expanded to embrace methods other common and
13 useful methods.

14
15 The arguments in section 6 are not scientifically justified. Although EPA's decision to not do a
16 quantitative analysis might have been justified on grounds of practicality, the panel feels that
17 quantitative uncertainty analysis is an integral part of any good assessment, and many issues in
18 this case beg for explicit consideration in the context of an uncertainty analysis. The panel
19 thought that EPA should be methodical and balanced about what variables and components of
20 the assessment would be included in the analysis. The uncertainty narratives and sensitivity
21 analyses already in the document are an excellent beginning and may constitute the lion's share
22 of the work necessary to implement quantitative uncertain analysis based on simple bounding.

23
24 **Recommendation: Consider omitting or strongly revising chapter 6, particularly its**
25 **argument that quantitative uncertainty analysis is unfeasible for the dioxin assessment.**

26
27 *6.2: Please comment on EPA's overall conclusion that a comprehensive quantitative uncertainty*
28 *analysis is not feasible.*

29
30 The panel rejects EPA's argument that a quantitative uncertainty analysis is unfeasible. Many on
31 the panel felt that the present circumstances warrant a compromise approach that would be
32 simple and achievable with modest effort by the agency. Various bounding approaches,
33 sensitivity studies, and event trees (probability trees without the probabilities) were suggested as
34 possible approaches that could be used. With such methods, legitimate and comprehensive
35 uncertainty analyses (including even fully probabilistic analyses) are possible. They would be
36 useful and sufficient to respond to NAS' criticism.

37
38 The panel generally agreed with EPA that expert elicitation would be problematic and should be
39 off the table. However, many on the panel further suggested that value-of-information methods
40 would also be very useful, although feedback from EPA included reservations about this idea. A
41 discussion of value of information methods appears as an appendix to these comments.

42
43 During its discussions of the other charge questions, the panel assembled several important
44 issues into a *parking lot* that should be addressed in the eventual uncertainty analysis.

45

1 As requested by the SAB, Roger Cooke sent us a document on bounding analysis. It focused on
2 the features of interval analysis, although this is hardly the only method that might be useful in
3 the context of the dioxin assessment. He mentions one issue that could be construed as a
4 disadvantage of this simplest bounding approach. It is idea that the ranges are supposed to be
5 absolute bounds on the possible values of each input variable. So the only thing you can say
6 about a percentage is that it is between zero and 100%, or the only thing you can say about a
7 dispersal distance is that it is between zero and the circumference of the Earth. (These are his
8 examples.) But I think this represents a misunderstanding by Roger of the word "absolute".
9 Vacuous (e.g., physically limiting) bounds are not the only bounds that can be used in interval
10 analysis. In fact, they are meant to be informed by observed study results.

11
12 Furthermore, we are not necessarily limited to interval ranges and interval analysis. As was
13 mentioned last time and elaborated upon in our written comments, there are a variety of methods
14 that, with proper application, could be useful and informative, including nested ranges, info gap
15 methods, p-boxes, probability trees, robust optimization, etc. These are non-trivial, potentially
16 valuable, alternatives to traditional probabilistic uncertainty analysis, able to provide insights on
17 critical uncertainties in the assessment endpoints and the ongoing and future research needed to
18 achieve their resolution.

19
20 What they could have said:
21 Guidance doesn't require it
22 Did do one
23 Doing one shouldn't delay formalization
24 Parking lot, but two-decade delay

25
26 **Reconsider the argument for not doing a quantitative uncertainty analysis, or undertake**
27 **one, perhaps using suggested list of techniques.**

28
29 *6.2a: Please comment on the discussion in Section 6 regarding volitional uncertainty an how this*
30 *type of uncertainty limits the ability to conduct a quantitative uncertainty analysis.*

31
32 The panel felt the term 'volitional uncertainty', which might also have been called 'decisional
33 uncertainty', should be dropped. The EPA should focus instead on uncertainties about the state
34 of world and display the different modeling choices and the consequences of making them. The
35 decisions mentioned in the chapter's discussion of volitional uncertainty are modeling choices,
36 and they should be dealt with using techniques for model uncertainty. Standard tools and
37 techniques for analysis of model uncertainty can be applied.

38
39 **Recommendation: Purge the document of the notion of 'volitional uncertainty'. Display**
40 **the different modeling choices and the consequences of making them.**

41
42 *6.3: Throughout the document (including the Appendices), EPA presents a number of limited*
43 *sensitivity analyses (e.g., toxicokinetic modeling, RfD ranges, cancer OSF ranges, cancer RfD*
44 *development). Please comment on the approaches used, and the utility of these sensitivity*
45 *analyses in clarifying potential significant uncertainties.*

46

1 6.3: *Utility of the sensitivity studies?* The utility of the sensitivity studies is very good, but they
2 are not integrated and they need to be.

3
4 So what should they do? Well, we've dumped a lot of issues from our consideration over the last
5 two days of the other sections onto the uncertainty analysis. It might be odd to discharge them
6 all now by suggesting that EPA doesn't need to conduct one. Will a QUA change the outcome
7 of this assessment? Josh Cohen, one of our public commenters who was on the NAS committee,
8 seems to think it would or at least could. I don't know. I'm not sure we can tell without doing
9 one. But maybe EPA's analysts know.

10
11 Should the absence of QUA further delay the finalization of this superannuated assessment? I'm
12 not sure that it should. Maybe we should consider this question and weigh our desire for an
13 uncertainty analysis in light of this. We want them to do a better job, but even more we want
14 them to do the job. Are we "past the time for reasonable debate and robust science", as a public
15 commenter said yesterday?

16
17 The panel congratulates EPA on the sensitivity studies that it has already done and considers
18 them to be very useful. The panel felt these studies should be integrated and unified in an overall
19 uncertainty analysis. The panel emphasized that EPA has already done the lion's share of the
20 effort needed already in their considerations described in the uncertainty narratives. The panel
21 feels the agency should take credit for this hard work and extend them to respond fully to the
22 NAS criticism.

23
24
25 6.2 (Infeasibility of a comprehensive quantitative uncertainty analysis)

26
27 Although a completely comprehensive quantitative uncertainty analysis might indeed be too
28 much to expect, I think that it is both possible and practical to provide readers with much more
29 useful information about uncertainty. A policy maker might reasonably expect the report to
30 provide insight into major uncertainties and questions such as the following:

- 31 • How likely is it that TCDD is not a human carcinogen at current exposure levels? Full
32 discussion of this uncertainty may help to overcome probability neglect and action bias (Patt and
33 Zeckhauser, 2000, <http://www.springerlink.com/content/k47064873365w720/>).
- 34 • What is the probability that reducing TCDD exposures would not reduce cancer risk at all,
35 based on recent epidemiological studies and updates such as Pesatori et al., 2009,
36 www.ncbi.nlm.nih.gov/pmc/articles/PMC2754980/?tool=pmcentrez&report=abstr
37 act?
- 38 • What is the probability that reducing TCDD exposures would reduce cancer risk by less
39 than 1 excess cancer case per decade (or per year or per century) in the whole US population,
40 under current conditions?
- 41 • What is the probability that reducing TCDD exposures would increase cancer risk (e.g., if
42 the dose-response relation is J-shaped or U-shaped)?
- 43 • What is the decision-analytic value of information (VoI) from collecting more information
44 on Ahr kinetics and dose-response before making risk management decisions? Although many
45 members of the public believe that it is imprudent and/or morally wrong to delay tighter
46 regulation of TCDD exposures (perhaps reflecting beliefs that TCDD is a potent carcinogen,

1 developmental toxin, etc.) EPA should provide a thorough quantitative decision analysis that
2 makes explicit the current uncertainties and trade-offs and that shows the conditions under which
3 acting now or postponing action are the optimal actions. Without such quantitative analysis, risk
4 management decisions for TCDD will not be adequately informed, and principles other than
5 those of rational decision-making (e.g., the biases discussed in Sunstein and Zeckhauser, 2010,
6 <http://www.hks.harvard.edu/fs/rzeckhau/Sunstein4-6-09.pdf>) may dominate risk management
7 decisions for TCDD. EPA's uncertainty analysis should provide the (decision and management
8 science) scientific basis for improved decision-making. The current decision to, in effect, punt
9 on quantitative uncertainty analysis is not adequate for informing responsible risk management
10 decision and policy-making, and is not justified.

11
12 While I agree with EPA that a quantitative uncertainty analysis is challenging, I do not think that
13 it is impractical to undertake one. It may well be true that we lack an adequate empirical basis
14 for Monte-Carlo propagation of input distributions, but there are many other options available
15 (e.g., Info-Gap analysis, uncertainty set analysis, consideration of alternative assumption sets and
16 their implied constraints on possible risks, etc.) that could at least provide useful bounds on the
17 plausible risks and on the VOI of reducing uncertainties further (especially, perhaps, on whether
18 the dose-response relation has a threshold – a topic still not settled, despite the pages of
19 discussion.)

20
21

22 **Person's responses to charge questions concerning section 6**

23

24 The arguments in section 6 are clearly written, mostly coherent, and perhaps fairly reasonable. I
25 had a lot of preliminary comments, including comments on the document's wording, some of
26 which is strongly at variance with the literature on uncertainty analysis. So I incorporate those
27 here by reference [see "Elaborated responses..." and "Minor comments" in the following
28 sections].

29

30 I was befuddled by the argument EPA used to justify not doing a unified QUA. If the blunt
31 answer to the question of why they didn't is that they couldn't specify precise marginal
32 distributions and dependence functions from existing data, then I reject this reasoning and
33 conclude EPA has not been responsive to the NAS criticism. If you're saying EPA guidance
34 doesn't require a QUA, then I would agree and say that the NAS criticism is perhaps itself
35 unreasonable. Or, if you say that you did do an uncertainty analysis in the form of UFs and the
36 limited sensitivity studies that you've done, then I might agree that's a reasonable position, even
37 if it's old-fashioned or dubious. Or even possibly, if you say that mounting a QUA is a
38 significant and controversial undertaking itself and that doing one shouldn't delay the
39 finalization of the report, that I could get behind just on grounds of practicality in the face of a
40 two-decade-long delay.

41

42 Here is the reasoning that I would have to reject: EPA asserts that "Data are the ultimate arbiter
43 of whether quantitative uncertainty analysis ... has sufficient evidentiary support". This flies in
44 the face of how uncertainty analyses are normally conceived. Of course, the absence of data is
45 never a substantive reason *not* to conduct an uncertainty analysis; it is the reason *to* do one.

46

1 EPA says it needs an “underlying distribution from which to sample” in order to conduct a
2 quantitative uncertainty analysis. I think this is a misunderstanding. And it is facile to shrug off
3 a call to characterize and account for important uncertainties in the assessment process on these
4 grounds alone. If you can *estimate* the value of a quantity, you should be able to express the
5 uncertainty about the value, otherwise you don’t really have a scientific measurement in the first
6 place. And, keep in mind, we are not forced to identify precise probability distributions and
7 dependence functions for everything that is to be characterized as uncertain. Even when the
8 uncertainty is volitional, there can be relevant ranges that are interesting to decision makers and
9 stakeholders. In some cases, the analysis may be formally closer to a sensitivity analysis, but
10 some appropriate response is usually possible, if not always practicable. To their credit, EPA has
11 acknowledged the legitimacy of the call by NAS and undertaken some efforts in this direction,
12

13 EPA calls uncertainty analysis an “emerging area in science” and this is inarguably true, but I
14 don’t believe it is true that methodological research is necessary for EPA to do anything more
15 comprehensive to respond to NAS’s criticism, even if we disallow the use of expert elicitation.
16

17 I’m entirely sympathetic to the idea of having analyses be *data-driven*, but it is still possible to
18 do something that’s useful, even it’s not precisely distributional. There are a variety of ways to
19 conduct a quantitative uncertainty analysis, even an entirely probabilistic one that obeys the
20 Kolmogorov axioms that require neither a bunch of data nor expert elicitation. I’ll provide a list
21 of various ways, with appropriate references [see the bulleted list in the summary of July
22 discussions about section 6]. The list includes simple interval analysis that just propagates the
23 plausible ranges, and the supervaluation approach that uses nested inner and outer intervals, with
24 the inner range representing the values that most everyone considers to be plausible values and
25 the outer range representing conservatively broad ranges. There’s also a continuous and
26 unbounded version of nesting intervals in an approach known as info-gap analysis that would be
27 useful if we cannot come up with finite bounds on some of the inputs. You can also propagate
28 bounds on distribution functions, so if you know some but not perfect information about each
29 input variable’s distribution or some information about some dependence function between the
30 variables, you can fashion bounds on distribution functions and conveniently propagate them
31 through calculations.
32

33 Does using these approaches require EPA to make judgments? Yes, it would, in the same way
34 that developing any analysis requires judgments. This does not mean that analysts would be
35 required to make up stuff or elicit any expert opinion. Does it necessitate a lot of extra work?
36 Not necessarily. These methods can be simple to develop, and they are mostly computationally
37 trivial. Of course, the more comprehensive it is, the harder it is. But the analysis does not have
38 to be fully comprehensive to be useful.
39

40 Nevertheless, I agree that an uncertainty analysis is not an absolute good. If the answer is
41 already clear, it can be a waste of time and resources. I don’t support wasting time and
42 resources. If the analysis is done poorly, or without appeal to available evidence from the real
43 world, it can be misleading. If the analysis is used *strategically* to avoid rendering or finalizing a
44 decision that is proper, it can be counterproductive.
45

46 The following are synoptic answers to the four charge questions of section 6:

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

Findings

The report addresses a broad range of philosophical and methodological issues in conducting an uncertainty analysis for TCDD toxicity, specifically for estimates of cancer oral slope factors and noncancer reference doses. The Section is successful in identifying the challenges involved in assessing uncertainty in toxicity estimates based on:

- **A small set of available models for toxicokinetics, dose-response relationships, and low dose extrapolation, with limited application, testing, and verification; and**
- **A small set of animal bioassay, epidemiological or clinical/case studies, many with differing endpoints, dose metrics, and (in the case of the human studies) uncertain exposure and subject data.**

As such, the Section provides many useful insights for EPA’s Reassessment. However, in its discussion of available methods, the report is somewhat biased in its treatment of certain statistical methods which could address some of these issues (though it does note their potential contribution at the end of the Section, as part of ongoing or future studies) and overly pessimistic regarding our ability provide improved quantitative estimate for certain portions of the toxicity assessment. This is unfortunate since, in other Sections of the Reassessment, the report provides a very credible discussion of the range of scientific uncertainty in current knowledge regarding TCDD toxicokinetics and toxicity.

Methods that should be given a more extensive and balanced discussion, including more citations to the literature include:

Bayesian Hierarchical Modeling (for combing information from multiple studies):

Axelrad DA, Bellinger DC, Ryan LM, Woodruff TJ. Dose–response relationship of prenatal mercury exposure and IQ: an integrative analysis of epidemiologic data. *Environ Health Perspect* 2007;115:609–615.

Choi, T., M. J. Schervish, K. A. Schmitt and M. J. Small. 2010. Bayesian hierarchical analysis for multiple health endpoints in a toxicity study. *Journal of Agricultural, Biological, and Environmental Statistics*. Available online at:
<http://www.springerlink.com/content/2h416p2581210773/fulltext.pdf>

Coull B., Menzetti M. and Ryan L. (2003) A Bayesian hierarchical model for risk assessment of methylmercury, *Journal of Agricultural, Biological and Environmental Statistics*, 8, 3, 253–270.

Ryan L. Combining data from multiple sources, with applications to environmental risk assessment. *Stat Med* 2008: 27(5): 698–710.

1
2 **Bayesian Model Averaging (for considering more than one dose-response equation,**
3 **allowing the data to weight their *relative* likelihood and contribution to the estimate):**
4

5 Morales, Knashawn H., Joseph G. Ibrahim, Chien-Jen Chen, and Louise M. Ryan. 2006.
6 “Bayesian Model Averaging With Applications to Benchmark Dose Estimation for Arsenic in
7 Drinking Water.” *Journal of the American Statistical*
8 *Association* 101 (473): 9–17.
9

10 Viallefont, V., Raftery, A.E. and Richardson, S. (2001) Variable selection and Bayesian model
11 averaging in case-control studies. *Statistics in Medicine* 20: 3215-3230.
12

13 Wheeler MW, Bailer AJ (2007). Properties of Model-Averaged BMDLs: A Study of Model
14 Averaging in Dichotomous Risk Estimation." *Risk Analysis*, 27, 659-670.
15

16 Wheeler, M. W., Bailer, A. J. (2009). Comparing model averaging with other model selection
17 strategies for benchmark dose estimation. *Environmental and Ecological Statistics* , **16** (1): 37–
18 51.
19
20

21 **Note: These Bayesian methods should not be referred to as “exotic”. For example, in**
22 **agreeing with the Section 6 authors that these methods should be pursued in ongoing and**
23 **future case studies, White et al. (2009) refer to them as “advanced”, rather than exotic.**
24 **Specifically, they recommend that health scientists should:**
25

26 **Explore statistical approaches to model selection**

27 Improvements to statistical approaches for model selection, such as model
28 averaging, should be pursued. Case study applications of these advanced
29 statistical approaches will identify potential strengths and weaknesses of these
30 approaches and their significance for risk characterization.

31 White et al. (2009)
32

33 R.H. White, I. Cote, L. Zeise, M. Fox, F. Dominici, T.A. Burke, P.D. White, D. Hattis, J.M.
34 Samet, State-of-the-science workshop report: issues and approaches in low dose–response
35 extrapolation for environmental health risk assessment, *Environ. Health Perspect.* 117 (2009)
36 283–287.
37
38

39 **Distributional (Probability Tree) Methods for considering alternative assumptions and**
40 **models at various stages of the toxicity assessment.**
41

42 These methods do rely upon expert judgment, but can provide a basis for ongoing integration and
43 value of information assessment as new studies and knowledge accumulate over time (Brusick
44 et. al., 2008). As described in Small (2008):
45

1 The distributional approach for characterizing uncertainty in cancer risk assessment was
2 developed by Evans, Sielken, and co-workers beginning in the 1990s⁽²⁻¹⁰⁾ and has also
3 been referred to as information analysis, weight-of-evidence analysis, the comprehensive
4 methodology, and comprehensive realism.⁽⁸⁻¹⁰⁾ The method has since been acknowledged
5 in a number of reviews of cancer risk assessment practice and research needs,⁽¹¹⁻¹³⁾ and
6 applied in various forms for risk assessment of different chemical compounds.⁽¹⁴⁻¹⁹⁾

7 The motivation for the distributional approach is the recognition that the use of a single
8 set of assumptions for the components of a cancer risk assessment, whether default,
9 conservative, or otherwise, fails to capture the full range of plausible or likely
10 relationships, how these relationships depend upon our current state of knowledge, the
11 implications for computed values of potency or unit risk, and the opportunities for
12 improved estimates. The distributional approach thereby enables consideration of a
13 "portfolio-of-mechanisms" that may contribute to carcinogenesis.⁽²⁰⁾

- 14 • 2. Holland, C. D., Sielken, R. L. Jr. (1993). Quantitative Cancer Modeling and Risk
15 Assessment . (Chapter 7). Englewood Cliffs , NJ : Prentice Hall.
- 16 • 3. Evans, J. S., Graham, J. D., Gray, G. M., Sielken, R. L. Jr. (1994). A distributional
17 approach to characterizing low-dose cancer risk. Risk Analysis , 14 (1), 25–34.
- 18
- 19 • 4. Evans, J. S., Graham, J. D., Gray, G. M., Sielken, R. L. Jr. (1995). A distributional
20 approach to characterizing low-dose cancer risk. In S. Olin, W. Farland, C. Park, L.
21 Rhomberg, R. Scheuplein, T. Starr, J. Wilson (Eds.), Low-Dose Extrapolation of Cancer
22 Risks (pp. 253–274). Washington , DC : ILSI Press.
- 23
- 24 • 5. Sielken, R. L. Jr. (1993). Evaluation of chloroform risk to humans. The Toxicology
25 Forum, 1993 Annual Winter Meeting, February 15–17, 1993, The Capitol Hilton,
26 Washington, DC.
- 27
- 28 • 6. Evans, J. S., Gray, G. M., Sielken, R. L., Jr., Smith, A. E., Valdez-Flores, C., Graham,
29 J. D. (1994). Use of probabilistic expert judgment in distributional analysis of
30 carcinogenic potency. Regulatory Toxicology and Pharmacology , 20 (1), 15–36.
- 31
- 32 • 7. Sielken, R. L., Jr., Valdez-Flores, C. (1999). Probabilistic risk assessment's use of trees
33 and distributions to reflect uncertainty and variability and to overcome the limitations of
34 default assumptions. Special Issue of Environmental International on Modeling and
35 Simulation , 25 , 755–772.
- 36
- 37 • 8. Sielken, R. L. Jr. (1990). A weight-of-evidence approach to quantitative cancer risk
38 assessment: Information analysis. In G. Schettler, D. Schmähl, T. Klenner (Eds.), Risk
39 Assessment in Chemical Carcinogenesis . New York : Springer-Verlag. Proceedings of
40 the Satellite Symposium on Risk Assessment in Chemical Carcinogenesis, Heidelberg,
41 Germany August 24–25, 1990.
- 42

- 1 • 9. Sielken, R. L., Jr., Bretzlaff, R. S., Stevenson, D. E. (1995). Challenges to default
2 assumptions stimulate comprehensive realism as a new tier in quantitative cancer risk
3 assessment. *Regulatory Toxicology and Pharmacology* , 21 , 270–280.
4
- 5 • 10. Sielken, R. L., Jr., Valdez-Flores, C. (1996). Comprehensive realism's weight-of-
6 evidence based distributional dose-response characterization. *Special Issue of the Human
7 and Ecological Risk Assessment on Theoretical, Toxicological and Biostatistical
8 Foundations for Deriving Probability Distribution Functions for Reference Doses and
9 Benchmark Doses with Application to Carcinogens and Noncarcinogens* , 2 (1), 175–193.
10
- 11 • 11. Boyce, C. P. (1998). Comparison of approaches for developing distributions for
12 carcinogenic slope factors. *Human and Ecological Risk Assessment* , 4 (2), 527–577.
13
- 14 • 12. Moschandreas, D. J., Karuchit, S. (2002). Scenario-model-parameter—A new method
15 of cumulative risk uncertainty analysis. *Environment International* , 28 (4), 247–261.
16
- 17 • 13. Zeise, L., Hattis, D., Andersen, M., Bailer, A. J., Bayard, S., Chen, C., Clewell, H.,
18 Conolly, R., Crump, K., Dunson, D., Finkel, A., Haber, L., Jarabek, A. M., Kodell, R.,
19 Krewski, D., Thomas, D., Thorslund, T., Wassell, J. (2002). Improving risk assessment:
20 Research opportunities in dose response modeling to improve risk assessment. *Human
21 and Ecological Risk Assessment* , 8 (6), 1421–1444.
22
- 23 • 14. Humphreys, S. H., Carrington, C., Bolger, M. (2001). A quantitative risk assessment
24 for fumonisins B1 and B2 in US corn. *Food Additives and Contaminants* , 18 (3), 211–
25 220.
26
- 27 • 15. Rai, S. N., Bartlett, S., Krewski, D., Paterson, J. (2002). The use of probabilistic risk
28 assessment in establishing drinking water quality objectives. *Human and Ecological Risk
29 Assessment* , 8 (3), 493–509.
30
- 31 • 16. Kirman, C. R., Sweeney, L. M., Teta, M. J., Sielken, R. L., Valdez-Flores, C.,
32 Albertini, R. J., Gargas, M. L. (2004). Addressing nonlinearity in the exposure-response
33 relationship for a genotoxic carcinogen: Cancer potency estimates for ethylene oxide.
34 *Risk Analysis* , 24 (5), 1165–1183.
35
- 36 • 17. Starr, T. B., Matanoski, G., Anders, M. W., Andersen, M. E. (2006). Workshop
37 overview: Reassessment of the cancer risk of dichloromethane in humans. *Toxicological
38 Sciences* , 91 (1), 20–28.
39
- 40 • 18. David, R. M., Clewell, H. J., Gentry, P. R., Covington, T. R., Morgott, D. A., Marino,
41 D. J. (2006). Revised assessment of cancer risk to dichloromethane II. Application of
42 probabilistic methods to cancer risk determinations. *Regulatory Toxicology
43 Pharmacology* , 45 (1), 55–65.
44

- 1 • 19. Crump, K. S. (1994). Risk of benzene-induced leukemia: A sensitivity analysis of the
2 Pliofilm cohort with additional follow-up and new exposure estimates. *Journal*
3 *Toxicology and Environmental Health* , 42 , 219–242.
- 4
- 5 • 20. Cox, L. A. (2006). Quantifying potential health impacts of cadmium in cigarettes on
6 smoker risk of lung cancer: A portfolio-of-mechanisms approach. *Risk Analysis* , 26 (6),
7 1581–1599.
- 8
- 9

10 **Elaborated responses to charge questions concerning section 6**

11
12 The arguments in section 6 are coherent and fairly reasonable, although they overstate some
13 issues and underserve some others. This section carefully considers the surprisingly detailed
14 criticisms from the National Academy of Science (NAS) review committee of the 2003
15 Reassessment concerning the need for quantitative uncertainty analysis. EPA has declined many
16 if not most of the particular suggestions of NAS about uncertainty, and it argues that undertaking
17 the suggested analyses would necessitate further fundamental research in uncertainty
18 quantification. Although I find some of its arguments to be compelling, I also wonder whether
19 EPA has really been responsive to the central criticism about uncertainty. Despite my own
20 strong disposition in favor of quantitative uncertainty analysis in general, it is possible to
21 conclude the agency’s judgments on this matter have been thoughtful and defensible.

22
23 The following are several comments aimed at improving the text.

24
25 The meaning of the phrase ‘epistemic uncertainty’ given on page 6-5 is plainly incorrect.
26 Epistemic uncertainty is the uncertainty that arises from imperfect knowledge such as from
27 limitations on the amount or quality of data available or deficiencies in our causal understanding
28 about a system. It is not true that a quantity about which there is epistemic uncertainty is
29 necessarily fixed. Although I can see how one might come to this mistaken impression, as far as
30 I know, no researchers use the phrase to imply that the underlying quantity has no variability
31 (although all would admit that this could be the case given our ignorance about it). This mistake
32 echoes in a couple of other places throughout this section.

33
34 There is some strange text on the subject of dependence. Lines 30-32 on page 6-5 and section
35 6.1.3.3 are also incorrect that the “[i]ssues involving...epistemic and aleatory uncertainty
36 translate into issues of dependence”. This is just wrong (even under their unusual definition of
37 ‘epistemic’). Likewise, the last paragraph on page 6-7 extending onto the next page should be
38 rewritten. The example is reasonable and important, but the discussion about it is confused. The
39 first sentence is incorrect. The uncertainty mentioned in the second sentence may be epistemic,
40 but the sentence is erroneous in its claim. In the following sentences, the words ‘variable’ and
41 ‘fixed’ (or ‘constant’) should be used rather than ‘aleatoric’ and ‘epistemic’. I believe it is
42 nonsense to say that a kinetic constant is “completely correlated across individuals”. It’s not
43 correlated; it is invariant. This case is not an example of a dependence issue. There is no
44 correlation between a distribution and a fixed quantity (even if it’s uncertain). Correlation is
45 defined between *varying* quantities. If the number is fixed, whether or not we know what it is,
46 then you cannot say it’s correlated with anything. The authors may have come to this twisted

1 language because they're thinking of the uncertainties in terms of how they might plan to
2 quantitatively characterize them in a Monte Carlo simulation (repeatedly selecting a random
3 deviate for the kinetic constant but assigning it to every individual). Of course, variables such as
4 body fat, age, and smoking, on the other hand, can and do exhibit correlations that definitely
5 should be accounted for in the quantitative assessments. Likewise, the constancy of particular
6 quantities about which we may not know the precise value is also important to keep track of.
7 These two issues should be untangled and discussed in a less confusing way.

8
9 It is not clear to me what the authors take to be the difference between epistemic uncertainty and
10 what they call 'cognitive uncertainty'. It seems that the latter phrase was introduced because the
11 meaning of 'epistemic uncertainty' had been misunderstood. Normally, the phrase 'cognitive
12 uncertainty' would refer to an individual person's uncertainty about the validity of the results of
13 his or her own information processing. The assertion that cognitive uncertainty may be
14 represented by probability (i.e., by precise probability measures) is unnecessary and may be
15 misleading. In fact, researchers in human cognition and neuroscience have shown that humans
16 process this kind of uncertainty (which they often call 'ambiguity') separately and differently
17 from what we think of as probability or frequentist risk (Hsu et al. 2005; Glimcher 2003). I
18 suggest that the section can omit the phrase 'cognitive uncertainty' altogether and use in its place
19 'epistemic uncertainty'. There are slight differences between the two ideas (e.g., epistemic
20 uncertainty could be shared by members in a group, whereas cognitive uncertainty is always
21 personal), but these appear to be unimportant in this context.

22
23 The assertion (on line 10 of page 6-5) that the frequentist and Bayesian interpretations are not
24 mutually exclusive may be misleading. They are mutually exclusive in the sense that it would be
25 improper to mix and match components of each into an analysis. I believe it would be
26 appropriate to omit the clause with the phrase 'mutually exclusive', although it is surely fair to
27 say that subjective probabilities can and do track relative frequencies.

28
29 Section 6.1.3.2 starting on page 6-6 discusses a way to address uncertainty for sample data. This
30 Spartan treatment does not mention that sampling uncertainty is not the only kind of uncertainty
31 that can be associated with data, nor that it may not even be the largest kind of uncertainty.
32 Mensurational uncertainty (including the plus-minus part of a measurement, and censoring) may
33 be more important. In some cases, the family or shape of the marginal distribution may be
34 unknown, which is a kind of model uncertainty. As suggested on page 6-35, such uncertainties
35 can be significant. The section suggests only a resampling approach to expressing the
36 uncertainty, but fails to mention the often severe limitations of such approaches, and says
37 nothing about what one might do if there is no relevant sample data.

38
39 The first paragraph of section 6.1.3.4 seems to be saying that one can sometimes express model
40 uncertainty as parametric uncertainty, which simplifies its handling. This could be said rather
41 more plainly. It would be helpful to mention that this trick cannot always be used (as when the
42 possible models cannot be listed). It might also be especially helpful to mention that this trick is
43 not so much a way to propagate model uncertainty as a way to sweep it under the rug. Model
44 averaging, including Bayesian model averaging, erases model uncertainty in the same way that
45 averaging variable quantities erases their variation. Bayesian model averaging is mentioned
46 several times in the document, including on page 6-36, lines 3ff. I believe that this method has

1 substantial disadvantages that may disqualify it for consideration here, even as an “exotic”
2 method. Having said this, I would hasten to emphasize that addressing model uncertainty is
3 often useful, and could be useful here as well despite the pessimism of 6.4.2.8. Even a restricted
4 sensitivity analysis, although clearly not comprehensive, can still be informative.

5
6 Section 6.1.3.6 starting on page 6-9 might also mention *graphs*, and other traditional
7 communication tools other than correlation indices.

8
9 Overall, I think the arguments in section 6 are fairly reasonable, or at least tenable. Although I
10 cannot completely subscribe to the document’s conclusion that a reasonably comprehensive
11 quantitative uncertainty analysis is not yet *possible* owing to a lack of models on which to hang
12 the analysis and unavailability of key empirical evidence, I agree that a serious effort in this
13 direction requires further development that may not be justified on *practicality* grounds in this
14 case.

15
16 EPA may be overstating the argument a bit, and some text should perhaps be softened. The
17 assertion “Data are the ultimate arbiter of whether quantitative uncertainty analysis with
18 uncertainty factors, as currently envisions, has sufficient evidentiary support” (page 6-21,
19 lines 12-14) flies in the face of how uncertainty analyses are normally conceived. Of course, the
20 absence of data is never a substantive reason *not* to conduct an uncertainty analysis; it is the
21 reason *to* do one.

22
23 Nevertheless, I agree that an uncertainty analysis is not an absolute good. If the answer is
24 already clear, it can be a waste of time and other resources. If it is used strategically to avoid
25 rendering a proper decision, it can be counterproductive. If it is done poorly, or without appeal
26 to available evidence from the real world, it can be misleading. Surely, if it is worth doing, it is
27 worth doing well and doing something well can be resource-intensive. The idea, mentioned in
28 footnote 66 on page 6-20, of arbitrarily converting uncertainty factors to independent lognormal
29 random variables in a scattered attempt to mount a quantitative uncertainty analysis would entail
30 a suite of unjustified and probably untenable assumptions rendering the exercise nearly pointless.

31
32 The pessimistic conclusion on page 6-31, line 24, may be a bit strong. Any *estimate* made from
33 data is amenable to a quantitative uncertainty analysis so, if you’re measuring anything, you can
34 propagate uncertainties such as mensurational uncertainty, sampling uncertainty, and perhaps
35 even surrogacy uncertainty. I don’t think it’s quite as hard to get quantitative models as the text
36 here seems to suggest. Likewise, the similarly dour conclusion on lines 13-14 of page 6-32
37 leaves me confused. You could do a sensitivity analysis in this case, couldn’t you? If so, it
38 seems that some kind of uncertainty analysis is clearly possible. The caveat on line 29 of page 6-
39 37 is also overwrought. I think exploring relevant alternative values in a sensitivity analysis
40 could constitute a quantitative uncertainty analysis, even if the exploration is limited.

41
42 It is important to keep in mind that, in general, we are not necessarily limited to identifying
43 precise probability distributions for everything that is to be characterized as uncertain (as seems
44 to be suggested on line 30 of page 6-37). Simple intervals about uncertain quantities can support
45 a straightforward, albeit crude, interval analysis that propagates uncertainty about parameters and
46 other model choices to statements about the range of possible results. Similarly, an approach

1 based on interval probabilities, probability boxes, or general imprecise probabilities (Walley
2 1991) can combine such intervals with precise distributions if they are known for some other
3 inputs, and with structures that are intermediate between coarse intervals and delicate probability
4 distributions when some but incomplete knowledge is available. If the inputs are profoundly
5 uncertain, the results from such analyses are likely to be wide in reflection of these uncertainties.
6 In pretty much all cases, it is possible to be entirely rigorous without necessarily being precise
7 and without completely specifying each probability distribution.

8
9 There does not need to be a specified “underlying distribution from which to sample” (page 6-37,
10 line 31) in order to conduct a quantitative uncertainty analysis. I think it is a bit too facile to
11 shrug off a call to characterize and account for important uncertainties in the assessment process
12 on these grounds alone. Even when the uncertainty is volitional, there can be relevant ranges
13 that are interesting to decision makers and stakeholders. In such cases, the analysis may be
14 formally closer to a sensitivity analysis, but some appropriate response is usually possible, if not
15 always practicable. To their credit, EPA has acknowledged the legitimacy of the call and
16 undertaken some efforts in this direction, notably Tables 5-18 and 5-19 (although some kind of
17 graphical summary of the results might have been nicer).

18
19 The assertions in section 6.5.2 are rather surprising and questionable. EPA says that uncertainty
20 quantification is an “emerging area in science” and that it is “an area where research could be
21 focused” because “the requisite knowledge does not yet exist” to apply quantitative uncertainty
22 analysis in assessments such as this one for dioxin. The document peremptorily dismisses the
23 utility of “convening a blue-ribbon panel” to identify the proper approach and suggests instead
24 that “multiple approaches should be encouraged”. Are we to infer that the present review panel
25 shouldn’t try to say what the proper approaches to uncertainty quantification are, even if we
26 think the area is more mature than emerging? Do these statements suggest that the agency will
27 support intramural and extramural research efforts in this direction? And, if not, how can we
28 take these pronouncements seriously? Is it not possible that EPA could benefit from some tech
29 transfer efforts as well as basic research on uncertainty quantification? The paragraph beginning
30 on page 6-42 (line 3) mentions a European idea of bench-test exercises to compare different
31 approaches. It may be worth mentioning that this idea has been implemented in the United
32 States as well (Oberkampf et al. 2004; Ferson et al. 2004).

33
34
35
36

1 **Value of Information**

2
3 When human health risk assessments include an explicit representation of uncertainty, the
4 potential value of new information (VOI) can be estimated by its ability to reduce uncertainties
5 that matter most to the assessment target. While methods for determining VOI are most
6 commonly associated with the decision analysis literature in the context of informing
7 management or regulatory decisions (Raiffa, 1968; Keeney, 1982; Winkler and Murphy, 1985;
8 Finkel and Evans, 1987; Taylor et al., 1993; Clemen, 1996; Chao and Hobbs, 1997), there are
9 many steps in a scientific assessment well before (or even without subsequent) decision support
10 and decision making where VOI evaluations can be of benefit in characterizing current scientific
11 knowledge and the potential for its improvement. ORD should integrate these methods into
12 their current and ongoing assessments of dioxin toxicity.
13

14 *When uncertainty in a scientific assessment is measured by the variance of model predictions, a*
15 *first measure of VOI is the extent to which this variance might be reduced by new or additional*
16 *data (e.g., Patwardhan and Small, 1992; Brand and Small, 1995; Abbaspour et al., 1996; Chao*
17 *and Hobbs, 1997; Sohn et al., 2000; Bosgra et al., 2005; Cooke, 2009). The relative*
18 *contribution of different model assumptions and parameter uncertainties to the variance of the*
19 *estimated effect (e.g., the BMD, or the cancer slope factor) provides an indication of which of*
20 *these uncertainties would be most beneficial to address. In addition, a VOI assessment*
21 *considers the potential for the component uncertainties to be reduced, based on the feasibility,*
22 *resource requirements (time and funding), and likelihood of success of the studies that would be*
23 *needed to achieve the necessary improvement in scientific knowledge.*

24
25 A scientific VOI study may also target a key classification inference that results from a risk
26 assessment, for example, whether a compound is genotoxic. Assuming the current assessment
27 leads one to assign an inconclusive probability to this outcome (e.g., between 10% and 90%, so
28 that neither inference can be rejected with a high degree of confidence), then potentially valuable
29 studies are those able to shift subsequent probabilities to high values (e.g., above 90, 95, or 99%)
30 with a positive result (e.g, providing support for genotoxicity) and/or to low values (below 10, 5,
31 or 1%) with a negative result.
32

33 To illustrate, Small (2008) presents a simple probability tree model (a “distributional approach”)
34 for assessing genotoxicity based on studies of DNA damage response caused by naphthalene and
35 its metabolites. In the proposed studies a series of isogenic cell lines deficient in various DNA
36 metabolism pathways are used to characterize the DNA damage responses caused by the targeted
37 compounds. Following results from the cultured cells, mice deficient in the specific DNA
38 damage responses would be exposed to naphthalene. Possible inferences are identified based on
39 the assessed sensitivity and selectivity of study results to the genotoxicity of naphthalene. Study
40 outcomes considered include: i) DNA damage responses in the isogenic cells; ii) increased
41 numbers of stable DNA adducts in the DNA repair deficient mouse lung; and iii) heightened
42 Clara cell toxicity in the DNA repair deficient mouse lung. Illustrative results using Netica are
43 presented as follows:
44
45

1
2

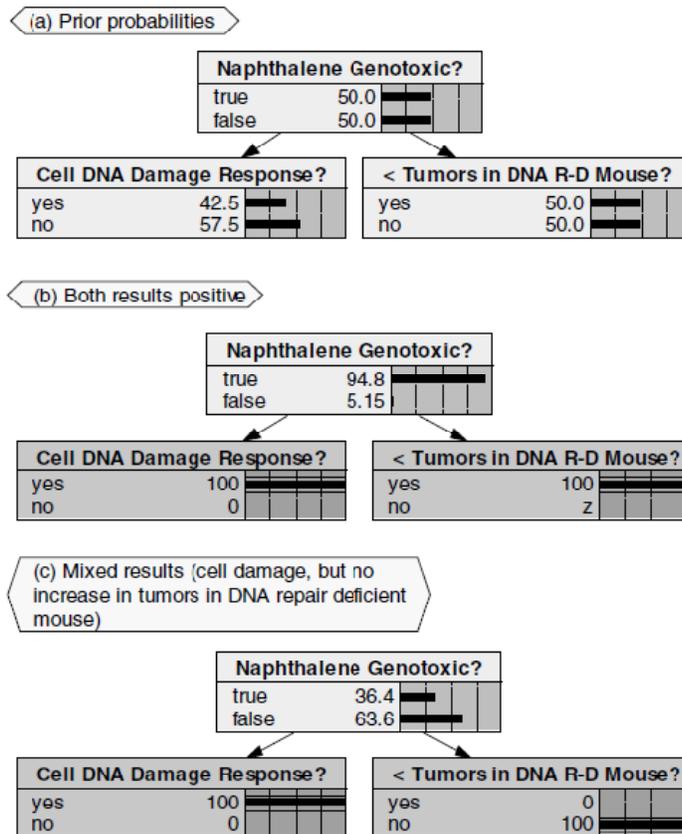


Fig. 5. Illustrative consideration of selected experimental results for naphthalene Study IV (outcomes assumed independent, prior probability of genotoxicity set to 0.5, and sensitivities and selectivities chosen by author solely for illustration of methodology); (A) Prior probability before study; (B) Positive outcomes for both study results; and (C) Positive results for cell DNA damage, but negative results for increased tumors in DNA repair-deficient mouse.

3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

As noted, the results shown above are intended solely to demonstrate the way in which study results can be combined to support or refute targeted inferences.

Even when the uncertainty tree method is only used to delineate the set of possible outcomes and relationships among steps and assumptions in the risk assessment (i.e., mode of action; dosimetry measures for exposure; the mathematical form of the dose-response relationship; the experimental data set(s) used to fit the relationship; and the procedure used for interspecies extrapolation) *without the assignment of probabilities to the tree branches*, key assumptions and the experiments needed to support or refute them can still be identified. These will typically involve elements of the assessment that, depending on their resolution, effectively restrict the set of possible outcomes to either a positive or a negative inference regarding the endpoint of the risk assessment. Establishing a procedure of this type will allow the Agency to put in place a more formal mechanism for identifying, conducting, and integrating the results of key studies for future assessments.

Minor editorial comments on chapter 6

Unless chapter 6 is omitted entirely from the document, the following minor comments may be useful in revising it.

Page 6-2. Add NRC(1996).

- 1
2 Page 6-3, bottom: The word ‘margins’ should be ‘marginals’.
3
4 Page 6-3, line 26: If you want to use the adverb ‘always’, the phrase ‘as a joint distribution’
5 should be ‘as some characterization of a joint distribution’ to be correct.
6
7 Page 6-4, lines 9-12: This text is strange and off-putting. A reader might ask who wrote this and
8 why. It seems opinionated and unnecessary.
9
10 Page 6-4, line 9: The tone is too pedagogical (“This is not the place . . .”).
11
12 Footnote 54: The discussion of alternatives to strict, single-measure probability theory is ham-
13 handed. Neither interval probabilities nor imprecise probabilities (sensu Walley 1991) depart
14 from probability theory; they follow the Kolmogorov axioms. They are motivationally and
15 essentially equivalent to sensitivity analyses, except they do not make use of sampling strategies
16 and can be more comprehensive.
17
18 Lines 29-30: It is simply untrue that sensitivity analyses have to be systematic. The word
19 ‘systematic’ might better be ‘comprehensive’ and the word ‘essential’ should be weakened,
20 perhaps to ‘advantageous’.
21
22 Page 6-5: I consider epistemic to mean unknown and aleatoric to mean inherently variable. So
23 when (for example) body weight varies across a population, but with a distribution that is
24 unknown, this exhibits both aleatoric and epistemic uncertainty.
25
26 Page 6-5, lines 4-7 and footnote 55: There seem to be only two axioms mentioned in the text,
27 but Kolmogorov needs three to make probability theory.
28
29 Page 6-5: The words ‘aleatoric’ and ‘aleatory’ are both used on this page as (synonymous)
30 adjectives of uncertainty. Actually, in the engineering literature, only ‘aleatory’ is preferred for
31 this use. In any case, please pick one to use.
32
33 Page 6-6, line 20: Maybe the last word of the header should be plural.
34
35 Line 21: Modern practice has replaced ‘error’ with ‘uncertainty’ in this context.
36
37 Footnote 56: You could add ‘or subtracting’ after ‘adding’.
38
39 Page 6-7, line 14: I think you should replace ‘The role of dependence modeling’ with
40 ‘Dependence among variables’.
41
42 Page 6-7. More examples of use of expert judgment for health assessment are available and
43 should be cited.
44
45 Page 6-8. line 13: Omit the unnecessary fancy after the semicolon.
46

- 1 Lines 15-17: This sentence is nonsense, if I understand what a linear low-dose model is. Parsing
2 the sentence, it seems to say “uncertainty over a...slope...may be quantified, but uncertainty...in
3 slope...cannot be captured” which is self-contradictory. I think what you mean to say is that the
4 linearity assumption is not itself subject to uncertainty quantification.
5
- 6 Page 6-9, line 1: The mathematical symbol x should be italicized, as should all Roman letters
7 throughout the document that represent unknown quantities, i.e., are symbols representing
8 something else rather than names like ‘e’ the base of the natural logarithms.
9
- 10 Lines 14 and 16: The prefixes ‘pseudo’ and ‘quasi’ are not words. Hyphens are needed.\
- 11
- 12 Page 6-9, line 18: Provide citations for dependence modeling.
13
- 14 Page 6-10, line 4: Add mention of methods that identify uncertain assumptions/ parameters that
15 are *important* – for determining whether the model is consistent with observed data (Hornberger
16 and Spear) and for affecting a decision that is made as a result of the model (Merz et al.).
17
- 18 Page 6-10, lines 29-30: Do you mean ‘*this* probabilistic language’, referring to the word ‘likely’
19 in the quoted text?
20
- 21 Page 6-11, line 19: Of course there is no guarantee that linear will be protective.
22
- 23 Page 6-13, line 18: Of course it isn’t really apodictic knowledge at all, but rather only an opinion
24 or an assumption. I see your point and agree with it entirely, but perhaps you should use a word
25 other than ‘apodictic’ here since it’s not technically correct.
26
- 27 Page 6-14, lines 33-34: The parenthetical phrase ‘volitional uncertainty’ should be expanded
28 into a sentence that says what you mean to express. The phrase ‘cognitive uncertainty’ does not
29 mean anything to me in this context. Perhaps if you expanded it into a sentence too, maybe
30 making it ‘epistemic uncertainty’ along the way, I would understand what you’re trying to say
31 here.
32
- 33 Footnote 62: ‘Effective’ is misspelled, as is ‘cancer’.
34
- 35 Page 6-16, line 5: And it’s not really a guarantee of course.
36
- 37 Line 8: The word ‘common’ should be ‘predominant’.
38
- 39 Page 6-16, line 20: Perhaps we can say that variability (and uncertainty) in the factors that are
40 used to determine a particular UF can be considered in choosing the particular value of the UF.
41
- 42 Page 6-17, lines 3-14: I disagree with this assertion. This problem can be addressed using a
43 Bayesian analysis with a beta conjugate for the uncertain response probability, p , with
44 informationless (uniform) prior for p . The probability that “an experiment with a null response
45 might have yielded a positive response” can be estimated from the predictive distribution (which
46 will depend on the number of test animals in the original study that yielded zero responses) for

- 1 the next experiment (with any number of exposed animals). I will bring an example to the
2 meeting.
3
- 4 Page 6-17, line 28: The word ‘band’ should be ‘limit’.
5
- 6 Page 6-20, footnote 66: The text starting ‘each have an error factor’ should be followed by ‘of’
7 rather than ‘or’.
8
- 9 Page 6-21, line 6: It would be nice to give a hint about what the concerns are.
10
- 11 Page 6-21, lines 12-14: NAS was not suggesting that EPA use the *uncertainty factors* approach
12 to mount an uncertainty analysis, but rather a more modern approach.
13
- 14 Page 6-22, line 19: And establishes a concomitant reduction in some UFs?
15
- 16 Line 29: The word ‘invokes’ should perhaps be ‘would require’.
17
- 18 Page 6-23, line 33 and passim: The word ‘exotic’ is a poor choice that is unnecessarily and
19 transparently loaded.
20
- 21 Page 6-25, line 29: This sentence is ungrammatical.
22
- 23 Page 6-26, line 24 and Figure 6-1: Would it be helpful to draw the 45-degree line on the graph?
24
- 25 Page 6-27, line 10: The word ‘epistemic’ here is acceptable.
26
- 27 Line 14: The word ‘epistemic’ here should be replaced by ‘fixed across individuals’. And ‘is
28 estimated from’ should be replaced by ‘varies with’. I don’t see how half life’s estimability from
29 data implies that it is variable.
30
- 31 Page 6-28, lines 1-2: You would need the dependence between the variables to proceed.
32
- 33 Line 9: I think that ‘and’ should be ‘although’.
34
- 35 Page 6-29, line 1-2: There are bounding techniques based on the classical Fréchet inequality that
36 do not require any knowledge of or any assumptions about dependencies.
37
- 38 Line 32: Omit ‘to’.
39
- 40 Page 6-33: The example in the text box is great, but the second table seems to say the log-
41 likelihood for LLD is 2.46 and for Hill is 2.16, which would make LLD’s larger than Hill’s,
42 which contradicts what’s said in the text.
43
- 44 Page 6-34, line 4: Shouldn’t ‘*Delivered dose*’ be a new bullet?
45

- 1 Line 8: I don't think this statement is true. Perhaps 'statistically more powerful' should be
2 'typically yield more sensitive'.
3
- 4 Lines 24-25: I don't think it's necessary or helpful to persist with Box's platitude. Model
5 uncertainty is the uncertainty about a model's predictions that arises from doubt about the
6 relevance of that model for making such predictions.
7
- 8 Page 6-37, line 30: This sentence is false. Analytical methods of propagation (convolution)
9 don't "sample" anything, and analyses based on intervals or imprecise probabilities don't depend
10 on uncertainty "distributions" (i.e., precise probability distributions).
11
- 12 Page 6-38, line 30 and passim: The adjective 'data driven' needs a hyphen, as it has elsewhere in
13 the document.
14
- 15 Line 23-24: I think this sentence is true, but, again, sampling from a distribution is not the only
16 way to conduct a quantitative uncertainty analysis.
17
- 18 Line 26: What is '(2.a)'?
19
- 20 Page 6-41, line 23: Omitting the word 'extra' would make the sentence more easily
21 understandable.
22
- 23 Line 31: What does 'How Forward?' mean? Is this idiomatic?
24
- 25 The document's reference list is alphabetically arranged, but seems to go from Z back to A again
26 on page R-33.
27

1 **References**

- 2
- 3 Abbaspour, K. C., R. Schulin, R., E. Schlappi and H. Fluhler 1996. A Bayesian approach for
4 incorporating uncertainty and data worth in environmental projects. *Environmental Modeling and*
5 *Assessment* 1: 151-158.
- 6 Aughenbaugh, J. M., and C.J.J. Paredis 2007. Probability bounds analysis as a general approach to
7 sensitivity analysis in decision making under uncertainty. *SAE 2007 Transactions Journal of*
8 *Passenger Cars: Mechanical Systems, (Section 6)* 116: 1325-1339, SAE International, Warrendale,
9 Pennsylvania. [http://www.srl.gatech.edu/Members/jaughenbaugh/papers_presentations/2007-01-](http://www.srl.gatech.edu/Members/jaughenbaugh/papers_presentations/2007-01-1480.pdf)
10 1480.pdf
- 11 Ben-Haim, Y. 2006. *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*. 2nd edition,
12 Academic Press, London.
- 13 *Bertsimas et al. 2009*
- 14 Bertsimas et al. 2010
- 15 Bosgra, S., Bos, P.M., Vermeire, T.G., Luit, R.J., Slob, W., 2005. Probabilistic risk characterization: an
16 example with di(2-ethylhexyl) phthalate. *Regul. Toxicol. Pharmacol.* 43: 104-113.
- 17 Brand, K.P. and Small, M.J. 1995. Updating uncertainty in an integrated risk assessment: Conceptual
18 framework and methods. *Risk Analysis* 15(6): 719-731.
- 19 Brusick, D., Small, M. J., Cavalieri, E. L., Chakravarti, D., Ding, X., Longfellow, D. G., Nakamura, J.,
20 Rogan, E. C., Swenberg, J. A. 2008. Possible genotoxic modes of action for naphthalene. *Regulatory*
21 *Toxicology and Pharmacology* 51 (2): 43-50.
- 22 Chao, P.T., Hobbs, B.F. 1997. Decision analysis of shoreline protection under climate change uncertainty,
23 *Water Resources Research* 33(4): 817-830.
- 24 Clemen, R.T. 1996. *Making Hard Decisions: An Introduction to Decision Analysis* (2nd ed.), Belmont,
25 CA: Duxbury Press.
- 26 Cooke, R.M. 2009. (ed) *Uncertainty Modeling in Dose Response, Bench-testing Environmental Models*,
27 Wiley, New York.
- 28 EPA (U. S. Environmental Protection Agency) 2007. Region 6 Superfund Program: Calcasieu Estuary
29 Initiative. <http://www.epa.gov/earth1r6/6sf/sfsites/calcnit.htm>, accessed July 2009
- 30 EPA (U.S. Environmental Protection Agency) 2002-2005. Several documents available at
31 <http://www.epa.gov/region1/ge/thesite/restofriver-reports.html> describing the U.S. EPA's use of
32 probability bounds analysis in the "rest of river" component of the assessment for the [GE/Housatonic](#)
33 [River Site in New England](#), including both the Ecological Risk Assessment and the Human Health
34 Risk Assessment.
- 35 Ferson, S. 2002. *RAMAS Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers*. Lewis
36 Publishers, Boca Raton, Florida.
- 37 Ferson, S. and T.F. Long. 1995. Conservative uncertainty propagation in environmental risk assessments.
38 *Environmental Toxicology and Risk Assessment, Third Volume, ASTM STP 1218*, J.S. Hughes, G.R.
39 Biddinger and E. Mones (eds.), American Society for Testing and Materials, Philadelphia, pp. 97-
40 110.
- 41 Ferson, S., V. Kreinovich, L. Ginzburg, D.S. Myers, and K. Sentz. 2003. *Constructing Probability Boxes*
42 *and Dempster-Shafer Structures*. SAND2002-4015. Sandia National Laboratories, Albuquerque,
43 NM.
- 44 Ferson, S., C.A. Joslyn, J.C. Helton, W.L. Oberkampf and K. Sentz. 2004. Summary from the epistemic
45 uncertainty workshop: consensus amid diversity. *Reliability and Engineering and System Safety* 85:
46 355-370.
- 47 Finkel, A.M. and Evans, J.S. 1987. Evaluating the benefits of uncertainty reduction in environmental
48 health risk management, *Journal Air Pollution Control Association*, 37:1164-1171.
- 49 Davidovitch, L., R. Stoklosa, J. Majer, A. Nietrzeba, P. Whittle, K. Mengersen, Y. Ben-Haim. 2009. Info-
50 gap theory and robust design of surveillance for invasive species: The case study of Barrow Island.
51 *Journal of Environmental Management* 90(8): 2785-2793.

- 1 Dixon, W.J. 2007. *The use of Probability Bounds Analysis for Characterising and Propagating*
2 *Uncertainty in Species Sensitivity Distributions*. Arthur Rylah Institute for Environmental Research,
3 Technical Report Series No. 163, Department of Sustainability and Environment. Victoria, Australia
4 [http://www.land.vic.gov.au/CA256F310024B628/0/F11A9B711262E4C3CA2573840014660E/\\$File](http://www.land.vic.gov.au/CA256F310024B628/0/F11A9B711262E4C3CA2573840014660E/$File/ARI+Technical+Report+163+-+The+use+of+probability+bounds+analysis+for+characterising+and+propagating+uncertainty+in+species+sensitivity+distributions.pdf)
5 [/ARI+Technical+Report+163+-](http://www.land.vic.gov.au/CA256F310024B628/0/F11A9B711262E4C3CA2573840014660E/$File/ARI+Technical+Report+163+-+The+use+of+probability+bounds+analysis+for+characterising+and+propagating+uncertainty+in+species+sensitivity+distributions.pdf)
6 [+The+use+of+probability+bounds+analysis+for+characterising+and+propagating+uncertainty+in+s](http://www.land.vic.gov.au/CA256F310024B628/0/F11A9B711262E4C3CA2573840014660E/$File/ARI+Technical+Report+163+-+The+use+of+probability+bounds+analysis+for+characterising+and+propagating+uncertainty+in+species+sensitivity+distributions.pdf)
7 [pecies+sensitivity+distributions.pdf](http://www.land.vic.gov.au/CA256F310024B628/0/F11A9B711262E4C3CA2573840014660E/$File/ARI+Technical+Report+163+-+The+use+of+probability+bounds+analysis+for+characterising+and+propagating+uncertainty+in+species+sensitivity+distributions.pdf)
- 8 Glimcher, P.W. 2003. *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*. The MIT
9 Press.
- 10 Hall, J.W. and H. Harvey 2009. Decision making under severe uncertainty for flood risk management: a
11 case study of info-gap robustness analysis. *Proceedings of the Eighth International Conference on*
12 *Hydroinformatics*, January 12-16, 2009, Concepcion, Chile.
13 http://www.floodsite.net/html/partner_area/project_docs/T20-08-05-hic-infogap.pdf
- 14 Hsu, M., M. Bhatt, R. Adolphs, D. Tranel, and C.F. Camerer. 2005. Neural systems responding to degrees
15 of uncertainty in human decision-making. *Science* 310 (5754): 1680–1683.
16 <http://www.sciencemag.org/cgi/content/abstract/310/5754/1680>
- 17 Karanki, D.R., H.S. Kushwaha, A.K. Verma, and S. Ajit. 2009. Uncertainty analysis based on probability
18 bounds (p-box) approach in probabilistic safety assessment. *Risk Analysis* 29(5):662-75.
- 19 Keeney, R.L. 1982. Decision analysis: an overview, *Operations Research* 39, 803–838.
- 20 Kreinovich and Trejo
- 21 Minnery, J.G., J.G. Jacangelo, L.I. Boden, D.J. Vorhees and W. Heiger-Bernays. 2009. Sensitivity
22 analysis of the pressure-based direct integrity test for membranes used in drinking water treatment.
23 *Environmental Science and Technology* 43(24): 9419–9424.
- 24 Montgomery, V. 2009. *New Statistical Methods in Risk Assessment by Probability Bounds*. Ph.D.
25 dissertation, Durham University, UK. <http://maths.dur.ac.uk/stats/people/fc/thesis-VM.pdf>
- 26 Moore, R.E. 1966. *Interval Analysis*. Prentice Hall, Englewood Cliffs, New Jersey.
- 27 Neumaier, A. 1990. *Interval Methods for Systems of Equations*. Cambridge University Press, Cambridge.
- 28 Oberkampf, W.L., J.C. Helton, C.A. Joslyn, S.F. Wojtkiewicz and S. Ferson. 2004. Challenge problems:
29 uncertainty in system response given uncertain parameters. *Reliability and Engineering and System*
30 *Safety* 85: 11-20.
- 31 Patwardhan, A., Small, M.J.. 1992. Bayesian methods for model uncertainty analysis with application to
32 future sea level rise. *Risk Analysis* 4, 513-523.
- 33 Raiffa, H. 1968. *Decision Analysis: Introductory Lectures on Choice under Uncertainty*, Reading, MA:
34 Addison-Wesley.
- 35 Regan, H.M., B.K. Hope and S. Ferson. 2002a. Analysis and portrayal of uncertainty in a food-web
36 exposure model. *Human and Ecological Risk Assessment* 8: 1757 - 1777.
- 37 Regan, H.M., B.E. Sample and S. Ferson. 2002b. Comparison of deterministic and probabilistic
38 calculation of ecological soil screening levels. *Environmental Toxicology and Chemistry* 21: 882-
39 890.
- 40 Regan, H.M., Y. Ben-Haim, B. Langford, W.G. Wilson, P. Lundberg, S.J. Andelman, and M.A. Burgman
41 2005. Robust decision-making under severe uncertainty for conservation management. *Ecological*
42 *Applications* 15(4):1471-1477.
- 43 Rout, T.M., C.J. Thompson and M.A. McCarthy. 2009. Robust decisions for declaring eradication of
44 invasive species. *Journal of Applied Ecology* 46: 782–786.
- 45 Small, M.J. 2008. Methods for assessing uncertainty in fundamental assumptions and associated models
46 for cancer risk assessment. *Risk Analysis* 28(5): 1289-1307.
- 47 Sohn, M.D., Small, M.J. and Pantazidou, M. 2000. Reducing uncertainty in groundwater site
48 characterization using Bayes Monte Carlo methods. *Journal of Environmental Engineering* 126(10):
49 893-902.
- 50 Taylor, A.C., Evans, J. and McKone, T. 1993. The value of animal test information in environmental
51 control decisions, *Risk Analysis* 13:403-412.

1 Walley, P. 1991. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall.
2 Winkler, R.L. and Murphy, A.H. 1985. Decision analysis, in A.H. Murphy and R.W. Katz (eds),
3 *Probability, Statistics, and Decision Making in the Atmospheric Sciences*, pp. 493–524, Boulder,
4 CO: Westview Press.
5 Yokomizo, H., H.P. Possingham, M.B. Thomas, and Y.M. Buckley. 2009. Managing the impact of
6 invasive species: the value of knowing the density-impact curve. *Ecological Applications* 19(2): 376-
7 386.
8
9
10
11
12
13
14
15
16
17
18