Subject: SAB Review of EPA’s Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments

Dear Administrator Jackson:

EPA’s Office of Research and Development (ORD) requested that the Science Advisory Board (SAB) review the Agency’s draft report entitled EPA’s Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments (“Report”). The Report contains EPA’s technical response to key comments in the 2006 National Academy of Sciences (NAS) report, Health Risks from Dioxin and Related Compounds: Evaluation of the EPA Reassessment. The NAS reviewed EPA’s 2003 exposure and human health reassessment of dioxin and recommended that the Agency should: more thoroughly justify and communicate its approaches to dose-response modeling for the health effects of dioxin, taking into consideration both nonlinear and linear methods for characterizing cancer risk; improve the transparency and clarity of the selection of key data sets for the dioxin dose-response analysis; reevaluate its cancer weight-of-evidence determination for dioxin based on the Agency’s 2005 Cancer Guidelines; consider using physiologically-based pharmacokinetic (PBPK) modeling in the dioxin risk assessment; and improve transparency, thoroughness and clarity in quantitative uncertainty analysis. The NAS also encouraged EPA to calculate a reference dose (RfD), which had not been derived in the 2003 reassessment.

In response to EPA’s request, the SAB convened an expert panel to review the Agency’s Report. The SAB Panel was asked to comment on the scientific soundness of EPA’s responses to the NAS recommendations. The enclosed SAB report provides the consensus advice and recommendations of the Panel, with the exception of one member who offered a dissenting opinion mainly on 2,3,7,8 Tetrachlorodibenzo-p-dioxin (TCDD) carcinogenicity.

The SAB finds that EPA’s Report is clear, logical, and responsive to many but not all of the recommendations of the NAS. We have provided recommendations to further enhance the transparency, clarity, and scientific integrity of the Report. The SAB has identified deficiencies in EPA’s Report with respect to the completeness of its consideration of two critical elements of the TCDD assessment: 1) nonlinear dose-response for TCDD carcinogenicity, and 2) uncertainty analysis of TCDD toxicity. Our major comments and recommendations are provided below:
1. The SAB commends EPA for the comprehensive and rigorous process that was used to identify, review, and evaluate the TCDD literature. The criteria for study selection have been clearly articulated, well justified, and applied in a scientifically sound manner. To further improve clarity and transparency of the Report, we recommend that EPA include a better means of tracking and describing which studies did not satisfy inclusion criteria. Similarly, we recommend that EPA strengthen its justification for excluding studies of dioxin-like compounds. The Report can be enhanced by incorporating information from studies with dioxin-like compounds into a qualitative discussion of the weight-of-evidence for cancer and noncancer endpoints.

2. EPA used the Emond physiologically-based pharmacokinetic model to evaluate the internal dose of TCDD in human and rodent tissue, and to estimate the continuous daily TCDD intake over the relevant period of exposure. The SAB agrees with EPA that this model provides the best available basis for the dose metric calculations. We also support EPA’s use of blood TCDD concentrations as the relevant dose metric. However, we recommend that EPA expand the discussion of other published models, evaluate the impact of model selection on dose metric prediction, provide a more quantitative uncertainty analysis, and conduct an external peer review of the mouse model because it has not been published in the peer-reviewed literature.

3. The SAB agrees with EPA’s classification of TCDD as carcinogenic to humans in accordance with EPA’s 2005 Guidelines for Carcinogen Risk Assessment. The SAB recommends that in the weight-of-evidence characterization EPA build upon all available data to support its decision and clearly indicate how different types of data support each other. One Panel member, however, indicated that at best, there is equivocal evidence for TCDD classification as a human carcinogen.

4. The SAB agrees with EPA’s selection of the Cheng et al. (2006) study, which analyzed the National Institute for Occupational Safety and Health (NIOSH) occupational cohort, as the critical study for the quantitative cancer assessment. The SAB also agrees that it is appropriate to use all-cancer mortality as the basis of the oral slope factor because of the extensive dose-response information.

5. The SAB finds that the Report did not respond adequately to the NAS recommendation to adopt both linear and nonlinear methods of extrapolation in order to account for the uncertainty of the dose-response curve for TCDD. The Report states that only a linear approach could be justified. We recommend that EPA revise the Report to provide a discussion of evidence of possible modes of action that include both linear and nonlinear alternatives for the cancer endpoint. In the absence of a definitive nonlinear mode of action, estimates based on the linear option can serve as the baseline for comparison with other estimates.
The SAB supports EPA’s selection and use of two co-critical epidemiologic studies for the derivation of the RfD for TCDD. These studies evaluated the effects of human exposure to TCDD following accidental release at a chemical plant near Seveso, Italy. The SAB finds that the study endpoints used by EPA to determine the RfD (decrease in sperm count and motility and increased thyroid stimulating hormone in blood) are relevant to public health. The selection of these endpoints also resolves the critical issue of differing windows of susceptibility to environmental toxic agents over the course of the life cycle, with pre- and periconceptional exposures comprising the window of greatest susceptibility. We recommend, however, that EPA provide a discussion of the strengths and weaknesses of the studies and an indication of whether these weaknesses affect the RfD determination.

The SAB also agrees with the benchmark dose modeling approaches used by EPA in the Report and the decision to use human data as preferred to animal data for the RfD determination.

EPA’s Report discusses a broad range of philosophical and methodological issues to be considered in conducting an uncertainty analysis for TCDD toxicity. Although the SAB acknowledges the challenges of a unified quantitative uncertainty analysis, we do not agree with the position taken in the Report that such an analysis is unfeasible and we have suggested a number of methods that could be used for this purpose.

Finally, EPA’s Report could be improved by editing and restructuring to better integrate the material presented in various sections, eliminate redundancies, and move some material into appendices to provide more succinct responses to NAS concerns. In addition, we recommend including a glossary in the Report to help minimize confusion and misinterpretation among diverse users.

The SAB appreciates the opportunity to provide EPA with advice on this important subject. We support EPA in its effort to move in a proficient and expeditious manner to finalize the IRIS document for dioxin and look forward to receiving the Agency’s response.

Sincerely,

Dr. Deborah L. Swackhamer, Chair
EPA Science Advisory Board

Dr. Timothy J. Buckley, Chair
SAB Dioxin Review Panel
NOTICE

This report has been written as part of the activities of the EPA Science Advisory Board, a public advisory committee providing extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. The Board is structured to provide balanced, expert assessment of scientific matters related to problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the Environmental Protection Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use. Reports of the EPA Science Advisory Board are posted on the EPA Web site at: http://www.epa.gov/sab
U.S. Environmental Protection Agency
Science Advisory Board
Dioxin Review Panel

CHAIR
Dr. Timothy Buckley, Associate Professor and Chair, Division of Environmental Health Sciences, College of Public Health, The Ohio State University, Columbus, OH

MEMBERS
Dr. Harvey Clewell, Director of the Center for Human Health Assessment, The Hamner Institutes for Health Sciences, Research Triangle Park, NC
Dr. Louis Anthony (Tony) Cox, Jr., President, Cox Associates, Denver, CO
Dr. Elaine Faustman, Professor and Director, Institute for Risk Analysis and Risk Communication, School of Public Health, University of Washington, Seattle, WA
Dr. Scott Ferson, Senior Scientist, Applied Biomathematics, Setauket, NY
Dr. Jeffrey Fisher, Research Toxicologist, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, AR
Dr. Helen Håkansson, Professor of Toxicology, Unit of Environmental Health Risk Assessment, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
Dr. Russ Hauser, Frederick Lee Hisaw Professor, Department of Environmental Health, Harvard School of Public Health, Boston, MA
Dr. B. Paige Lawrence, Associate Professor, Departments of Environmental Medicine and Microbiology and Immunology, School of Medicine and Dentistry, University of Rochester School of Medicine and Dentistry, Rochester, NY
Dr. Michael I. Luster, Professor, Department of Community Medicine, West Virginia University Health Sciences Center, Morgantown, WV
Dr. Paolo Mocarelli, Professor of Clinical Biochemistry, Department of Clinical Laboratory, Hospital of Desio-Nuovo Monoblous, University of Milano Bicocca, Desio-Milano, Italy
Dr. Victoria Persky, Professor, Epidemiology and Biostatistics Program, School of Public Health, University of Illinois at Chicago, Chicago, IL
Dr. Sandra L. Petersen, Professor, Associate Graduate Dean, Department of Veterinary and Animal Sciences, College of Natural Sciences, University of Massachusetts-Amherst, Amherst, MA

Dr. Karl Rozman*, Professor, Pharmacology, Toxicology and Therapeutics, The University of Kansas Medical Center, Kansas City, KS

Dr. Arnold Schecter, Professor, Environmental and Occupational Health Sciences, School of Public Health-Dallas Campus, University of Texas, Dallas, TX

Dr. Allen E. Silverstone, Professor, Department of Microbiology and Immunology, Health Science Center, SUNY Upstate Medical University, Syracuse, NY and Adjunct Professor of Environmental Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY.

Dr. Mitchell J. Small, The H. John Heinz III Professor of Environmental Engineering, Department of Civil and Environmental Engineering and Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA

Dr. Anne Sweeney, Professor of Epidemiology, Department of Epidemiology and Biostatistics, School of Rural Public Health, Texas A&M Health Science Center, College Station, TX

Dr. Mary K. Walker, Professor, Division of Pharmaceutical Sciences, College of Pharmacy, University of New Mexico, Albuquerque, NM

SCIENCE ADVISORY BOARD STAFF
Dr. Thomas Armitage, Designated Federal Officer, U.S. Environmental Protection Agency, Washington, DC

Dr. Diana Wong, Designated Federal Officer, U.S. Environmental Protection Agency, Washington, DC

*Dissenting opinion in Appendix A
U.S. Environmental Protection Agency
Science Advisory Board
BOARD

CHAIR
Dr. Deborah L. Swackhamer, Professor and Charles M. Denny, Jr. Chair in Science, Technology and Public Policy - Hubert H. Humphrey School of Public Affairs and Co-Director of the Water Resources Center, University of Minnesota, St. Paul, MN

SAB MEMBERS
Dr. David T. Allen, Professor, Department of Chemical Engineering, University of Texas, Austin, TX

Dr. Claudia Benitez-Nelson, Full Professor and Director of the Marine Science Program, Department of Earth and Ocean Sciences, University of South Carolina, Columbia, SC

Dr. Timothy J. Buckley, Associate Professor and Chair, Division of Environmental Health Sciences, College of Public Health, The Ohio State University, Columbus, OH

Dr. Patricia Buffler, Professor of Epidemiology and Dean Emerita, Department of Epidemiology, School of Public Health, University of California, Berkeley, CA

Dr. Ingrid Burke, Director, Haub School and Ruckelshaus Institute of Environment and Natural Resources, University of Wyoming, Laramie, WY

Dr. Thomas Burke, Professor, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

Dr. Terry Daniel, Professor of Psychology and Natural Resources, Department of Psychology, School of Natural Resources, University of Arizona, Tucson, AZ

Dr. George Daston, Victor Mills Society Research Fellow, Product Safety and Regulatory Affairs, Procter & Gamble, Cincinnati, OH

Dr. Costel Denson, Managing Member, Costech Technologies, LLC, Newark, DE

Dr. Otto C. Doering III, Professor, Department of Agricultural Economics, Purdue University, W. Lafayette, IN
Dr. David A. Dzombak, Walter J. Blenko, Sr. Professor of Environmental Engineering, Department of Civil and Environmental Engineering, College of Engineering, Carnegie Mellon University, Pittsburgh, PA

Dr. T. Taylor Eighmy, Vice President for Research, Office of the Vice President for Research, Texas Tech University, Lubbock, TX

Dr. Elaine Faustman, Professor and Director, Institute for Risk Analysis and Risk Communication, School of Public Health, University of Washington, Seattle, WA

Dr. John P. Giesy, Professor and Canada Research Chair, Veterinary Biomedical Sciences and Toxicology Centre, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

Dr. Jeffrey K. Griffiths, Associate Professor, Department of Public Health and Community Medicine, School of Medicine, Tufts University, Boston, MA

Dr. James K. Hammit, Professor, Center for Risk Analysis, Harvard University, Boston, MA

Dr. Bernd Kahn, Professor Emeritus and Associate Director, Environmental Radiation Center, Georgia Institute of Technology, Atlanta, GA

Dr. Agnes Kane, Professor and Chair, Department of Pathology and Laboratory Medicine, Brown University, Providence, RI

Dr. Madhu Khanna, Professor, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Urbana, IL

Dr. Nancy K. Kim, Senior Executive, Health Research, Inc., Troy, NY

Dr. Catherine Kling, Professor, Department of Economics, Iowa State University, Ames, IA

Dr. Kai Lee, Program Officer, Conservation and Science Program, David & Lucile Packard Foundation, Los Altos, CA (affiliation listed for identification purposes only)

Dr. Cecil Lue-Hing, President, Cecil Lue-Hing & Assoc. Inc., Burr Ridge, IL

Dr. Floyd Malveaux, Executive Director, Merck Childhood Asthma Network, Inc., Washington, DC

Dr. Lee D. McMullen, Water Resources Practice Leader, Snyder & Associates, Inc., Ankeny, IA
Dr. Kathleen Segerson, Philip E. Austin Professor of Economics, Department of Economics, University of Connecticut, Storrs, CT

Dr. Herman Taylor, Director, Principal Investigator, Jackson Heart Study, University of Mississippi Medical Center, Jackson, MS

Dr. Barton H. (Buzz) Thompson, Jr., Robert E. Paradise Professor of Natural Resources Law at the Stanford Law School and Perry L. McCarty Director, Woods Institute for the Environment, Stanford University, Stanford, CA

Dr. Paige Tolbert, Professor and Chair, Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA

Dr. John Vena, Professor and Department Head, Department of Epidemiology and Biostatistics, College of Public Health, University of Georgia, Athens, GA

Dr. Thomas S. Wallsten, Professor and Chair, Department of Psychology, University of Maryland, College Park, MD

Dr. Robert Watts, Professor of Mechanical Engineering Emeritus, Tulane University, Annapolis, MD

Dr. R. Thomas Zoeller, Professor, Department of Biology, University of Massachusetts, Amherst, MA

SCIENCE ADVISORY BOARD STAFF

Dr. Angela Nugent, Designated Federal Officer, U.S. Environmental Protection Agency, Washington, DC

Ms. Stephanie Sanzone, Designated Federal Officer, U.S. Environmental Protection Agency, Washington, DC
TABLE OF CONTENTS

5 ABBREVIATIONS AND ACRONYMS

6 EXECUTIVE SUMMARY

7 INTRODUCTION

8 RESPONSES TO EPA’S CHARGE QUESTIONS

9 Charge Question 1. General Charge Questions

10 Charge Question 2. Transparency and Clarity in the Selection of Key Data Sets for Dose-Response Analysis

12 Charge Question 3. The Use of Toxicokinetics in the Dose-Response Modeling for Cancer and Noncancer Endpoints

14 Charge Question 4. Reference dose

15 Charge Question 5. Cancer assessment

16 Charge Question 6. Feasibility of Quantitative Uncertainty Analysis

17 REFERENCES

18 APPENDIX A: DISSENTING OPINION FROM KARL ROZMAN, PH.D.

19 APPENDIX B: VALUE OF INFORMATION

20 APPENDIX C: EDITORIAL COMMENTS AND CORRECTIONS

21 APPENDIX D: EPA’S CHARGE QUESTIONS

xix
ABBREVIATIONS AND ACRONYMS

1 AgR  aryl hydrocarbon receptor
2 BMD  benchmark dose
3 BMDL benchmark dose lower bound
4 BMR  benchmark response level
5 CYP  cytochrome P450
6 DLC  dioxin-like compound
7 ED   effective dose
8 EPA  U.S. Environmental Protection Agency
9 HED  human equivalent dose
10 IRIS integrated risk information system
11 LASC lipid-adjusted serum concentrations
12 LOAEL lowest-observed-adverse-effect level
13 MOA mode of action
14 NAS  National Academy of Sciences
15 NHEERL National Health and Environmental Effects Research Laboratory
16 NIOSH National Institute for Occupational Safety and Health
17 NOAEL no-observed-adverse-effect level
18 NRC  National Research Council
19 OSF  oral slope factor
20 PBPK physiologically-based pharmacokinetic
21 PCDDs polychlorinated dibenzo-p-dioxin
22 PCDFs polychlorinated dibenzofuran
23 POD  point of departure
24 Pfd reference dose
25 RR  relative risk
26 SAB Science Advisory Board
27 T3   triiodothyronine
28 T4   thyroxine
29 TCDD 2,3,7,8-Tetrachlorodibenzo-p-dioxin
30 TEF  toxicity equivalence factor
31 TEQ  toxicity equivalence
32 TSH  thyroid stimulating hormone
33 UF   uncertainty factor
34 WHO World Health Organization
EXECUTIVE SUMMARY

In 2003, EPA, along with other federal agencies, asked the National Academy of Sciences (NAS) to review aspects of the science in EPA’s draft dioxin reassessment entitled, Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and Related Compounds and, in 2004, EPA sent the 2003 draft dioxin reassessment to the NAS for review. In 2006, the NAS released the report of its review entitled, Health Risks from Dioxin and Related Compounds: Evaluation of the EPA Reassessment. The NAS recommended that EPA should: more thoroughly justify and communicate its approaches to dose-response modeling for the health effects of dioxin, taking into consideration both nonlinear and linear methods for characterizing cancer risk; improve the transparency and clarity of the selection of key data sets for the dioxin dose-response analysis; reevaluate its cancer weight-of-evidence determination for dioxin based on the Agency’s 2005 Cancer Guidelines; consider using physiologically-based pharmacokinetic (PBPK) modeling in the dioxin risk assessment; and improve transparency, thoroughness and clarity in quantitative uncertainty analysis. The NAS also encouraged EPA to calculate a reference dose (RfD), which had not been derived in the 2003 reassessment.

EPA’s Office of Research and Development (ORD) prepared the draft report, entitled EPA’s Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments (EPA, 2010) (hereafter referred to as the Report), and requested that the EPA Science Advisory Board (SAB) conduct an independent external peer review of the Report. This Executive Summary highlights the findings and recommendations of the SAB Dioxin Review Panel (the “Panel”) in response to charge questions concerning each of the six sections of the Report.

General Charge

The SAB Panel was asked to comment on: whether the Report was clear and logical, whether the Agency had objectively and clearly presented the key National Academy of Sciences (NAS) recommendations, and whether there were other critical studies that would make a significant impact on the conclusions of the hazard characterization or dose-response assessment of the chronic noncancer and cancer health effects of TCDD.

As further discussed in the responses to Charge Question 1, the Panel found that EPA was effective in developing a report that was clear, logical, and responsive to many but not all of the recommendations of the NAS. The Panel has provided recommendations to further improve the clarity, organization, and responsiveness of various parts of the Report. The Panel was impressed with the process that EPA used to identify, review, and evaluate the relevant literature. The Panel found that EPA’s process was comprehensive and rigorous and included public participation. However, the Panel recommends that the Report be improved by: incorporating text to better integrate the material presented in the individual chapters, providing greater clarity and transparency.
in indicating which studies did not satisfy criteria for inclusion in EPA’s assessment of 
TCDD, and editing the *Report* to provide greater clarity in writing and make it more 
concise by moving some material into appendices.

During the course of its discussion, the Panel did not identify any additional 
studies that would make a significant impact on the conclusions of the hazard 
characterization and dose-response assessment. The Panel recommends that EPA 
provide an assessment of both the null studies and positive studies with more discussion 
and clarity concerning the exclusion of null epidemiologic studies. In addition, as further 
discussed in the responses to the relevant charge questions, the Panel has identified 
deficiencies in the *Report* with respect to the completeness of its consideration of two 
critical elements: 1) nonlinear dose-response for TCDD carcinogenicity, and 2) 
uncertainty analysis. As discussed below, the Panel has provided recommendations to 
improve the *Report* in these areas.

**Transparency and Clarity in the Selection of Key Data Sets for Dose-Response 
Analyses**

The NAS proposed that EPA develop a clear and readily understandable 
methodology for evaluating and including epidemiologic and animal bioassay data sets in 
dose-response evaluations. The SAB Panel was asked to comment on: whether EPA had 
been responsive to NAS concerns about transparency and clarity in data set selection, 
whether the epidemiology and animal bioassay study criteria and considerations had been 
scientifically justified and clearly described, and whether EPA had applied the 
epidemiology and animal bioassay study criteria considerations in a scientifically sound 
manner.

Section 2 of the *Report* contains a clear presentation of the process EPA used to 
select key data sets for dose-response analysis and is thus responsive to NAS 
recommendations in this area. The *Report* also clearly identifies the studies that were 
used for dose-response analysis. However, the Panel has provided recommendations to 
进一步增强总体清晰度和透明度的Section 2 of the *Report*. The Panel 
recommends careful and extensive editing to revise and consolidate Section 2. 
Specifically, editing should include aspects of grammar and syntax, minimizing 
redundancies, and use of more succinct language in responses to NAS concerns. The 
Panel also recommends restructuring Section 2 to improve its integration into the overall 
document and make it easier to follow the studies used by EPA from one section of the 
*Report* to another. In this regard, the Panel suggests that Section 2 could be used as the 
foundation for the entire document.

The Panel also found that EPA’s epidemiology and animal bioassay study criteria 
and considerations were scientifically justified, clearly described, and applied in a 
scientifically sound manner. The Panel has provided recommendations to improve and 
strengthen the scientific justification and clarity of description of EPA’s study criteria and 
considerations. The Panel recommends that EPA better justify the rationale for using
only studies where the exposure was primarily to TCDD for derivation of the reference dose. This justification should include both scientific and practical reasons. The Panel also recommends that EPA incorporate information from studies with dioxin-like compounds (DLCs) into a qualitative discussion of the weight-of-evidence for cancer and noncancer endpoints. In addition, the Panel has provided a number of specific recommendations to further clarify the justifications for some of the study inclusion and exclusion criteria.

**Use of Toxicokinetics in Dose-Response Modeling for Cancer and Noncancer Endpoints**

In the Section 3 of the Report, EPA discussed the use of a physiologically-based pharmacokinetic (PBPK) model (Emond et al., 2004, 2005, 2006) with blood concentration as the dose metric rather than first-order body burden. The Panel was asked to comment on the scientific justification for EPA’s application of this model, the model modifications that EPA implemented, and EPA’s characterization of uncertainty in the model. EPA also developed a PBPK model to estimate TCDD concentration in mouse tissues. The Panel was asked to comment on the scientific rationale for development of the mouse model, the performance of the mouse model, and whether model uncertainty had been adequately characterized. In addition, the Panel was asked to comment on the use of the Emond PBPK model to estimate human intake based on internal exposure measures, EPA’s sensitivity analysis of the kinetic modeling, and EPA’s estimate of lifetime average daily dose.

The Panel agrees with EPA’s use of blood TCDD concentration as a surrogate for tissue TCDD exposure. Blood TCDD concentration is a better choice than using body burden (as in the 2003 Reassessment) because it is more closely related to the biologically relevant dose metric: the free concentration of dioxin in the target tissues. The Panel further agrees that the PBPK model developed by Emond et al. (2004, 2005, 2006) provides the best available basis for the dose metric calculations in the assessment. However, the Panel recommends that EPA clarify how the model was used in studies that reported the concentrations of dioxin in plasma, serum, blood, or blood fat:blood measurements. The Panel also recommends additional discussion of: other published models, the intended use of the Emond model in the assessment, and the basis for selection of the Emond model. The Panel found that the EPA modifications to the published Emond model were minor and appropriate. However, the Panel notes that the use of 0.6 as the Hill coefficient in the Emond model for CYP1a2 induction is well outside the confidence interval of 0.78 and 1.14 reported by Walker et al. (1999). The use of a Hill coefficient value well below unity would lead to a nonlinear model behavior that is biologically implausible. As a result, when the human model was used for extrapolation to lower doses (in the calculation of risk-specific doses), the model would estimate a lower exposure level for a given blood concentration. The Panel suggests repeating the human Emond model calculations with multiple values for the Hill coefficient to characterize the resulting uncertainty in the exposure estimates.
The Panel also recommends that a more quantitative uncertainty analysis be conducted for the PBPK model. Methods that could be useful and informative for such analysis are suggested in the response to Charge Question 6.2. The sensitivity analysis in the Report left out the Hill coefficient, which is one of the most important parameters in the model for low-dose extrapolation. 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.

The Panel found that the mouse model developed by EPA based on the published rat model (Emond et al., 2004, 2005, 2006) was appropriate, but it is recommended that an external peer review of the mouse model be performed. The Panel agrees with the average daily dose calculation approaches described in the Report. However, the Panel recommends that EPA carefully explain how the early life stage internal doses were calculated because serum thyroid stimulating hormone (TSH) levels in newborns are used as a critical effect.

**Reference Dose**

In Section 4 of the Report EPA discussed the use of two co-critical studies (Mocarelli et al., 2008 and Baccarelli et al., 2008) for development of the reference dose for TCDD. The Panel was asked to comment on the scientific justification for selection and use of these studies to develop the reference dose.

**a. Selection of Critical Studies and Effects**

The Panel supports EPA’s selection of the Mocarelli et al. (2008) and Baccarelli et al. (2008) studies for identifying “co-critical” effects for the derivation of the reference dose (RfD). The Panel found that these two human epidemiological studies were well designed. The studies provided sufficient exposure information, including biological concentrations that could be used to establish acceptable lifetime daily exposure levels. The rationale for selecting these two studies over numerous other available studies for determining the RfD was clearly described but study weaknesses were not clearly delineated. The Panel recommends that EPA provide a discussion of the strengths and weaknesses of these studies with an indication of whether the weaknesses affect determination of the RfD. In addition, the Panel recommends that the comprehensive data base of both animal and human epidemiological studies be used to demonstrate a consistent and integrative signal of toxicity across species and endpoints for TCDD. The collective impact of the studies should be made stronger in the Report by including discussion of both human and experimental animal studies that have examined the effects of dioxin and DLCs on other reproductive and endocrine endpoints. In this regard, dose-response relationships as well as comparisons of no-observed-adverse-effect levels (NOAELs) and lowest-observed-adverse-effect levels (LOAELs) should be discussed.

The Panel agrees with EPA’s assertion that traditional (e.g., immune, endocrine, reproductive) endpoints are more appropriate than biochemical endpoints for establishing...
points of departure (PODs). The associations of traditional endpoints with health outcomes have been well studied and they are more tightly associated with adverse outcomes than biochemical endpoints. However, EPA should discuss biochemical endpoints, particularly P450s, relevant to establishing and strengthening the proposed reference dose.

b. Estimation of Continuous Exposure for Mocarelli et al. (2008)

Mocarelli et al. (2008) reported male reproductive effects (decrease in sperm count and motility) observed later in life for boys with high acute exposure to TCDD between the ages of 1 and 9 (average age 5 years), followed by low level background dietary exposure. EPA identified a 10 year critical exposure window and estimated the continuous TCDD intake as the average of the high acute exposure and the 5 year average exposure during the critical exposure window. The Panel found that the pattern of exposure from Seveso posed some extrapolation issues for the EPA, particularly whether the same endpoints and dose-response from high acute exposures would be expected when extrapolating to low-dose chronic exposures. It would be useful for EPA to provide a discussion of published examples in which dioxin studies were conducted using both high-dose acute and low-dose chronic exposures in animals for the same endpoint and how the outcomes compare both qualitatively and quantitatively. The life stage-specific approach to hazard and dose-response characterization for children’s health risk assessment in EPA’s Framework for Assessing Health Risks of Environmental Exposures to Children (EPA, 2006) is also relevant to addressing this issue and should be discussed.

c. Designation of a 20% Decrease in Sperm Count as a LOAEL for Mocarelli et al. (2008)

The Panel supports the use of the change from normal sperm counts and sperm motility for determining an RfD. While the shifts observed in sperm counts may or may not pose a significant health effect in a single individual, such shifts on a population basis could presumably lead to an increased incidence of adverse health outcomes. The Panel recommends that World Health Organization (WHO) reference values for male reproductive parameters and life stage differences in sperm counts in humans be discussed in the Report.

d. Determination of Effective Exposure Estimate for the Baccarelli et al. (2008) Study

EPA determined the maternal intake at the LOAEL from the maternal serum-TCDD vs. neonatal TSH regression model by finding the maternal TCDD lipid-adjusted serum concentrations (LASC) at which neonatal TSH exceeded 5 µU/ml. EPA then used the Emond PBPK model under the human gestational scenario to estimate the continuous daily oral TCDD intake that would result in a TCDD LASC corresponding to a neonatal TSH of 5 µU/ml at the end of gestation. EPA estimated the effective maternal intake as
0.024 ng/kg-day. The Panel supports EPA’s decision to use the Baccarelli et al. (2008) estimates of the relevant effective doses. The Panel also suggests that since the bulk of the calculations were based on zonal averages of exposed individuals in Baccarelli et al. (2008), EPA should clarify how these measurements relate to ranges and variations in exposure in utero.

e. Designation of 5µ-units TSH per ml blood as a LOAEL for Baccarelli et al. (2008)

EPA selected a LOAEL of 5µ-units TSH per ml blood in neonates. The Panel supports EPA’s designation of the TSH endpoint within the context of the broader dioxin literature. While the shift observed in TSH levels may or may not pose a significant health effect in a single individual, such a shift on a population basis could presumably lead to an increased incidence of adverse health outcomes. There is a need to better describe the potential adverse health outcomes related to altered neonatal TSH levels. For example, in addition to effects on growth, both cognitive and motor deficits have been found in young adults with congenital hypothyroidism. The Report could better describe the consequences of transient hypothyroidism on reproductive outcomes.

f. Selection of Uncertainty Factors

A composite uncertainty factor of 30 (an uncertainty factor of 10 for the lack of a NOAEL, and an uncertainty factor of 3 for human interindividual variability) was applied to the LOAEL of 0.020 ng/kg-day from Mocarelli et al. (2008) to obtain the RfD. The Panel agrees that EPA has used the appropriate uncertainty factors for the derivation of the RfD. However, a short discussion of the decision not to include an uncertainty factor for data quality is needed.

g. Benchmark Dose (BMD) Modeling of animal bioassay data and EPA’s Choice of POD from These Studies

The Panel agrees with the BMD modeling approaches used in the Report. In addition, the Panel agrees that the animal data have sufficient limitations that preclude their use to establish a RfD.

Cancer Assessment

In Section 5 of the Report EPA has provided: a weight-of-evidence characterization of TCDD as a known human carcinogen, conclusions regarding the mode of carcinogenic action for TCDD, EPA’s selection of data sets for cancer dose-response modeling, and consideration of approaches for assessment of TCDD carcinogenicity. The Panel was asked to comment on the scientific soundness of these aspects of EPA’s cancer assessment.
a. Weight-of-Evidence Cancer Descriptor

The Panel agrees with EPA’s conclusion that TCDD is “Carcinogenic to Humans.” The Panel recommends that the Agency provide more discussion of the power of the studies used and the difficulties involved when assessing rare tumors. The Panel also recommends that EPA consider including studies with substantial DLC exposure where toxicity equivalence factors (TEFs) can be calculated in the weight-of-evidence discussion. EPA should also attempt to characterize the uncertainty regarding the carcinogenicity of TCDD at low human exposures, since the minimum dose at which carcinogenic effects would be expected to occur cannot be clearly delineated from the current epidemiological human data.

b. Mode of Action

The Panel believes the mode of action for TCDD toxicity should be “reasonably well known” rather than “largely unknown,” although the Panel agrees that the exact mechanism of action has not been fully delineated for any distinct TCDD toxicity endpoint. The Panel recommends that EPA provide a discussion of the evidence for possible modes of action that include both linear and nonlinear alternatives; and that the description of the nature of a receptor mediated dose-response be expanded by including more evidence regarding the nonlinearity of the receptor mediated dose-response for dioxin.

c. Selection of Critical Study for Cancer Endpoint

The Panel agrees with the inclusion of the Cheng et al. (2006) study in the cancer assessment. This study incorporated information on gradation of exposure. However, expanded discussion of several other studies would support the weight-of-evidence for carcinogenicity in less common cancers such as lymphomas and soft tissue sarcoma. The Panel agrees that Cheng et al. (2006) was the appropriate study for quantitative cancer assessment, and that it was appropriate to use all-cancer mortality in this case, because of the extensive dose-response information. The Panel also agrees that the use of the Emond model to estimate risk-specific doses from Cheng et al. (2006) dose-response modeling results was scientifically justified and clearly described but, as previously discussed, the value of the Hill coefficient used in the model is problematic. The Panel found that Cheng et al. (2006) study did not provide completely clear information regarding risks below current background exposure levels. The Panel therefore suggests that EPA expand the discussion to consider the possibility that mode of action considerations could help indicate whether linear extrapolation of the Cheng et al. (2006) data is appropriate to obtain risk estimates in this range of exposures.

d. Nonlinear Approach for Assessment of TCDD Carcinogenicity

The Panel found that the Report did not respond adequately to the NAS recommendation to adopt “both linear and nonlinear methods of risk characterization to
account for the uncertainty of dose-response relationship shape below the ED01.” The Panel recommends that EPA present both linear and nonlinear risk assessment approaches. The nonlinear examples in the document should be formalized and extended. In the absence of a definitive nonlinear mode of action, the linear option results can serve as the baseline for comparison with these other estimates.

Quantitative Uncertainty Analysis

Section 6 of the Report discusses a broad range of philosophical and methodological issues to be considered in conducting an uncertainty analysis for TCDD toxicity. The Panel was asked to comment on: whether the discussion in this section of the Report was clearly presented and scientifically justified, the conclusion that a comprehensive quantitative uncertainty analysis (QUA) is not feasible, the discussion regarding volitional uncertainty and how it limits the ability to conduct a QUA, and approaches that EPA used to conduct sensitivity analyses.

The Panel found that Section 6 of EPA’s Report was clearly presented and provided many useful insights for EPA’s dioxin reassessment, but it was not scientifically justified. As further discussed in the responses to Charge Question 6, the Panel does not agree with EPA’s argument that conducting a unified QUA for TCDD toxicity is unfeasible. EPA’s decision to not conduct an integrated QUA may be based primarily on grounds of practicality or timeliness. In particular, EPA argues that a complete quantitative uncertainty analysis would require data and resources not available. We disagree with this logic. More limited evaluations can, and should, be implemented to inform critical issues in the dioxin reassessment. In the response to Charge Question 6.2 we suggest a number of methods that could be used. The Panel recommends that EPA revise its argument that QUA for dioxin toxicity is unfeasible.

EPA’s document contrasted volitional uncertainty with cognitive uncertainty. The Panel recommends that the term “volitional uncertainty,” which might also have been called “decisional uncertainty,” be dropped from the Agency’s document. The Panel recommends that EPA focus on uncertainties about the state of the world and display different modeling choices and the consequences of making them. The Panel recommends that EPA apply standard tools and techniques for analysis of model uncertainty.

In addition, the Panel found that the sensitivity studies EPA has already completed are useful. The Panel is mindful of the need to minimize further delay of the finalization of EPA’s already protracted dioxin assessment and we recommend that sensitivity studies that EPA has already completed be integrated into whatever overall uncertainty analysis the Agency elects to undertake.
INTRODUCTION

EPA has been preparing an assessment of the potential health impacts of 2,3,7,8-
Tetrachlorodibenzo-p-Dioxin (TCDD) for many years. In 2003, EPA released an
external review draft report entitled, Exposure and Human Health Reassessment of
2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and Related Compounds (U.S. EPA, 2003)
(hereafter referred to as the 2003 Reassessment) that was reviewed by the EPA Science
Advisory Board (SAB), and then by the National Academy of Sciences (NAS). In 2006,
the National Research Council (NRC) of the National Academies published their
evaluation of EPA’s reassessment, Health Risks from Dioxin and Related Compounds:
Evaluation of the EPA Reassessment (NRC, 2006).

The NAS identified key recommendations that they believed would result in
substantial improvement to the 2003 Reassessment and thus support a scientifically
robust characterization of human responses to exposures to TCDD. The NAS
recommended that EPA should: more thoroughly justify and communicate its approaches
to dose-response modeling for the health effects of dioxin, taking into consideration both
nonlinear and linear methods for characterizing cancer risk; improve the transparency and
clarity of the selection of key data sets for the dioxin dose-response analysis; reevaluate
its cancer weight-of-evidence determination for dioxin based on the Agency’s 2005
Cancer Guidelines; consider using physiologically-based pharmacokinetic (PBPK)
modeling in the dioxin risk assessment; and improve transparency, thoroughness and
clarity in quantitative uncertainty analysis. The NAS also encouraged EPA to calculate a
reference dose (RfD), which had not been derived in the 2003 Reassessment.

In 2010, EPA’s Office of Research and Development (ORD) prepared the draft
report, EPA’s Reanalysis of Key Issues Related to Dioxin Toxicity and Responses to NAS
Comments (EPA, 2010) (hereafter referred to as the Report). The Report includes new
analyses completed in response to the NAS recommendations and recently published
literature, as well as a discussion of topics where EPA’s views differed from those of the
NAS. The Report is not an assessment per se; it was designed to supplement the
information provided in EPA’s 2003 Reassessment. However, the Report provides a
noncancer reference dose and updated cancer values. Detailed discussions of many of the
issues addressed in the Report are available in the 2003 Reassessment and were not
reproduced in the Report.

ORD requested that the EPA Science Advisory Board (SAB) conduct an
independent external peer review of the Report. In its review, the SAB was asked to
consider the accuracy, objectivity, and transparency of EPA’s reanalysis and responses.

In response to ORD’s request, the SAB convened an expert panel to conduct the
review. The Panel held an initial public teleconference on June 24, 2010 to receive an
orientation to EPA’s Report. The Panel then held two public face-to-face meetings (July
13 – 15, 2010 and October 27 – 29, 2010) to deliberate on the charge questions (see
Appendix D) and two public teleconferences (March 1 and 2, 2011) to discuss its report.
Public comments were provided in oral and written form to the Panel at the
teleconferences and face-to-face meetings. There were charge questions on the 6 sections
of EPA’s Report. The questions focused on: transparency and clarity in the selection of
key data sets for dose-response analysis, the use of physiologically-based
pharmacokinetic (PBPK) modeling in dose-response modeling for cancer and noncancer
endpoints, derivation of a proposed oral reference dose (RfD) for noncancer endpoints,
cancer weight-of-evidence classification, mode of action of dioxin carcinogenicity,
derivation of oral slope factor (OSF) for dioxin, and quantitative uncertainty analysis.
This report provides the consensus advice and recommendations of the Panel, with the
exception of one member who offered a dissenting opinion mainly on the TCDD
carcinogenicity (see Appendix A).
RESPONSES TO EPA’S CHARGE QUESTIONS

1

Charge Question 1. General Charge Questions

The Panel was asked to comment on: whether the EPA’s Report was clear and logical, whether the Agency had objectively and clearly presented the key National Academy of Sciences (NAS) recommendations, and whether there were other critical studies that would make a significant impact on the conclusions of the hazard characterization or dose-response assessment of the chronic noncancer and cancer health effects of TCDD.

1.1. Is the draft Response to Comments clear and logical? Has EPA objectively and clearly presented the three key NRC recommendations?

Response:

The Panel found that EPA has developed a report that is clear, logical and responsive to many but not all of the recommendations of the NAS. While we provide a general assessment of opportunities for improvement in the context of this overview charge question, most of the issues related to clarity, organization, and responsiveness are addressed more completely and specifically in the context of the subsequent, more specific charge questions.

With respect to the first question, the Panel found that the EPA was effective in developing a clear, transparent, and logical response. The Panel was particularly impressed with the process that EPA used for identifying, reviewing, and evaluating the relevant literature. EPA’s process was comprehensive, rigorous, and included public participation. The Agency’s report, EPA’s Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments, consists of two volumes. The first volume contains the main text of the Report and is 690 pages long (including 37 pages of references). The second volume contains appendices and is 1,159 pages long. Because of the size and complexity of EPA’s Report, the Panel found the Executive Summary to be particularly important and valuable in providing a concise and accurate summary. As described in detail in the response to Charge Question 2, the Panel identified the need for better integration across chapters and greater clarity and transparency in indicating which studies did not satisfy inclusion criteria, and therefore were not carried forward for further consideration in subsequent chapters of the Report to meet particular needs. In contrast, the Panel was satisfied that the inclusion criteria and the associated retained studies were well described and transparent. The Panel acknowledges that, given the enormity of the dioxin published literature, it is not a trivial matter to characterize what has been “left on the cutting room floor.” We therefore suggest that EPA do this in a way that provides only a general consideration of this issue.
The Report is long and dense with a considerable amount of jargon and in some places it is quite repetitive. These features, while a necessity of this type of document, at times detract from clarity or make the EPA’s logic difficult to discern. The Panel found some instances where the Report would benefit from greater clarity in writing. For example, topic sentences are sometimes not easily connected to paragraph content. A specific example of this is in the second paragraph on page xxvii of volume 1 where the text does not clearly identify separate EPA activities to address NAS comments. Another example of a Report section that could be edited to improve clarity is the qualitative discussion of the uncertainty in the RfD (Section 4.4 of volume 1). The clarity of this section could be improved by including bullet points to highlight and separate key points and/or provide links to information in other sections of the document (e.g., Section 6 – Feasibility of Quantitative Uncertainty Analysis). The Panel suggests a careful review by a qualified technical editor. Similarly, the Panel suggests that the clarity and accessibility of the Report could be enhanced by the inclusion of a glossary to help minimize confusion and misinterpretation among the diverse users of the document. At 690 pages, volume 1 is a formidable report. The Panel appreciates the dilemma of preparing a report that is both complete and rigorous and at the same time succinct and efficient, but it is suggested that EPA find additional efficiencies (e.g., greater use of appendices and elimination of redundancies) that would yield a more approachable document.

With respect to the second part of Charge Question 1.1 (i.e., objectivity and clarity of presentation of the three key NAS recommendations), the Panel found that EPA has been successful. The Panel found EPA’s Report to be clear in presentation of the key NAS recommendations. However, as described more fully in responses to the relevant specific charge questions below, the Panel identified deficiencies in the Report with respect to the completeness of its consideration of two critical elements: 1) nonlinear dose-response for TCDD carcinogenicity and 2) uncertainty analysis.

**Recommendations**

- As further discussed in the response to Charge Question 2, the Panel recommends that the Report be revised to provide greater clarity and transparency in the discussion of studies that did not satisfy inclusion criteria for use in the dioxin assessment. Given the enormity of the dioxin published literature, the Panel recognizes that it is not a trivial matter to characterize the studies that were not considered, and therefore the Panel suggests that the Report be revised to generally indicate how this issue was considered.

- The Report is long and dense and contains a considerable amount of jargon. It would benefit from greater clarity in writing. The Panel therefore recommends that the Report be carefully reviewed by a qualified technical editor and revised to incorporate such improvements as better integration across chapters, better
connection between topic sentences and paragraph content, and elimination of repetition.

- The Panel recommends that the clarity and accessibility of EPA’s Report be enhanced by the inclusion of a glossary to help minimize confusion and misinterpretation among the diverse users of the document.

- The Panel recommends that EPA find additional efficiencies (e.g., greater use of appendices and elimination of redundancies) to yield more succinct and approachable document.

- As discussed in the responses to other charge questions in this report, the Panel identified deficiencies in EPA’s Report with respect to the completeness of its consideration of two critical elements: 1) nonlinear dose-response for TCDD carcinogenicity, and 2) uncertainty analysis. In the relevant charge question responses below, the Panel has provided recommendations to improve the Report in these areas.

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?

Response:

During the course of its discussion, the Panel did not identify any additional studies that would impact the hazard characterization or the dose-response assessment. However, the Panel found that EPA’s Report should provide more clarity on the exclusion of null epidemiologic studies.

Recommendations

- The Panel recommends that EPA’s Report provide more discussion and clarity on the exclusion of null epidemiologic studies.

Charge Question 2. Transparency and Clarity in the Selection of Key Data Sets for Dose-Response Analysis

General Comments:

The NAS committee proposed that EPA develop a clear and readily understandable methodology for evaluating and including epidemiologic and animal bioassay data sets in dose-response evaluations. Section 2 of EPA’s Report describes the Agency’s approach to ensuring transparency and clarity in the selection of the studies for...
dose-response analyses. The Panel was asked to comment on: whether EPA had been responsive to NAS concerns about transparency and clarity in data set selection, whether the epidemiology and animal bioassay study criteria and considerations had been scientifically justified and clearly described, and whether EPA had applied the epidemiology and animal bioassay study criteria considerations in a scientifically sound manner.

EPA developed and applied two sets of criteria for the animal bioassays and epidemiologic data. The Agency collected and evaluated these studies, including studies from the 2003 Reassessment and newer studies found through literature searches and through public submissions. The Panel viewed with favor all of the efforts made by EPA to develop this section of the document. The Panel compliments the Agency for its efforts to present the nuanced differences and complicating issues surrounding this subject in a comprehensive and logical manner. The intention of the comments and recommendations provided below is to assist the EPA in further improvement of Section 2.

2.1. Is this section responsive to the NAS concerns about transparency and clarity in data set selection for dose-response analysis?

Response:

Members of the Panel found that Section 2 of the Report was responsive to NAS concerns about transparency and clarity. Moreover, it was perceived and appreciated that, in addressing these concerns, EPA had improved the approach in the original 2003 Reassessment. The EPA’s collaboration with Argonne National Laboratory, and invitation to the public to engage in updating the literature search to identify all appropriate studies for evaluation, as well as the conduct of the Dioxin Workshop in February of 2009, were instrumental in enhancing the transparency and clarity regarding the process of selection of studies for the dose-response analysis. The development of clear criteria for study evaluation and inclusion was crucial in addressing the concerns raised by the NAS.

EPA’s Report presents a clear identification of the study selection process and the studies that were used for dose-response analysis. For example, the process and criteria used to select key data sets for dose-response analyses is described in Section 2.3 of the Report and in the Executive Summary. Flow diagrams (e.g., ES-1 and ES-2) clearly demonstrate how studies were chosen for inclusion. Likewise, Appendix B, which includes a point-by-point evaluation of which epidemiological studies were included and excluded, was useful and provides a detailed rationale explaining why the EPA used the particular studies selected in the Report. In addition, the results of the literature search performed by EPA are available online. Clarity could be improved by providing search words used for the MedLine searches. A clear case for including high-quality human studies over animal studies is also made.
While Section 2 of the Report is deemed responsive to NAS concerns, the Panel found that overall clarity and transparency regarding dataset selection would be further and markedly enhanced if EPA were to make Section 2 (and the document as a whole) more concise. In its present form, Section 2 was viewed by the Panel as overly verbose, to the detriment of overall clarity and we provide the following recommendations to improve the Report.

**Recommendations**

- The Panel strongly recommends careful and extensive editing to revise and consolidate Section 2 and the Report as a whole. Specifically, editing should include aspects of English grammar and syntax, minimizing redundancies, and efforts to provide more succinct responses to NAS concerns.

- The Panel recommends restructuring Section 2 to make it easier to follow a study used by EPA from one section of the Report to another. In other words, EPA should improve overall document integration using Section 2 as the foundation for this integration.

**Charge Questions 2.2 and 2.3**

2.2. *Are the epidemiology and animal bioassay study criteria/considerations scientifically justified and clearly described?*

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

**Response:**

The Panel’s discussion of Charge Questions 2.2 and 2.3 was highly integrated. Therefore, comments and specific recommendations that stem from these two questions are presented together.

The Panel found that EPA’s study criteria and considerations were scientifically justified and clearly described, and that these were presented in a scientifically sound manner. Thus, Section 2 was deemed responsive to NAS concerns regarding the scientific justification and clarity of description for epidemiology and animal bioassay study criteria/considerations. However, several concerns were discussed by the Panel, and are summarized here.

The Panel’s major concern pertains to improving clarity with regard to the decision to include or exclude particular studies and groups of studies from the data sets to be used. The rationale for distinct criteria for epidemiological and animal studies should be made stronger, and data set selection for noncancer and cancer endpoints has...
room for further clarification and justification. There was discussion, with differences of opinion among members of the Panel, regarding EPA’s scientific justification and clarity of description concerning the Agency’s decision to exclude dioxin-like compounds.

There was consensus among Panel members that the following recommended improvements would strengthen this section, and thus the document as a whole.

**Recommendations**

**Rationale for excluding dioxin-like compounds**

- EPA should better justify the rationale for using studies where the exposure is primarily to TCDD (or for animal studies only to TCDD) to calculate the reference dose. This justification should include scientific and practical reasons.

- EPA should incorporate information from studies with dioxin-like chemicals into a qualitative discussion of the weight-of-evidence for cancer and noncancer endpoints.

**Study inclusion and exclusion criteria and considerations**

- EPA should further clarify the justifications for study inclusion and exclusion criteria/considerations. To be clear, this recommendation does not indicate that the Panel suggests that a different approach to data set selection is needed. However, the approach used should be explained more effectively and clearly. In this regard, the following specific recommendations are provided to address points of concern raised by Panel members about the study inclusion and exclusion criteria:
  
  o EPA should remove the criterion that studies must contain an explicit statement of TCDD purity. For research purposes, TCDD is available from a limited set of vendors, and all sell it as a highly purified compound. Thus, for the animal studies, it is highly unlikely that any study would be conducted using impure TCDD. Therefore, excluding a study simply due to absence of statements regarding TCDD purity runs the risk of excluding high quality studies because the author or journal editorial staff did not elect to include this piece of information.

  o EPA should revise the explanation of the in vivo mammalian bioassay evaluation indicating that the “study design is consistent with standard toxicological practices.” This is too vague as it likely has different meaning to readers from different backgrounds. In addition to defining this more clearly, it is recommended that, if possible, a reference should be provided to an EPA document in which these practices are described in detail.
EPA should consider eliminating use of the phrase “outside the range of normal variability,” especially when discussing animal studies.

EPA should define the phrase “common practices,” and if possible cite appropriate Agency documents to which the reader can refer for further detail. To provide further context, this recommendation refers specifically to statements such as the following one on page 2-5: “The study criteria shown below and in Figure 2-3 for animal bioassay data reflect EPA’s preferences for TCDD-specific study inclusion, some of which are based on common practices and guidance for POD selection and RfD and OSF derivation.”

EPA should provide a more thorough (albeit concise) discussion of data set limitations to educate the reader regarding Agency decisions about study inclusion/exclusion criteria. For instance, consider adding an expanded discussion on suitability of studies of immunological effects and/or thyroid and diabetes (e.g., Baccarelli et al., 2002, 2004; Calvert, 1999; Steenland, 2001).

Considerations concerning selection of epidemiology studies

The Panel recommends that EPA better justify and explain considerations relating to the selection of epidemiology studies. The following specific recommendations are provided. Many of these specifically address the use of more standard epidemiology vocabulary and descriptors.

EPA evaluated the available epidemiologic cohorts and studies based on five considerations presented on pages 2-6 and 2-7 of the Report. The Panel found that Consideration #2 (page 2-6) was worded awkwardly and that epidemiologic terms are misspecified. The Panel therefore recommends that EPA revise Consideration #2 as follows:

- Define “susceptible to important biases.” This is a non-specific term and the biases should be explained.

- Clarify what is meant by “control for potential confounding exposures.” Does this refer to only exposure to dioxin-like compound exposures or was it meant to more broadly refer to other exposures as well (NIOSH cohort studies)? Does the text “bias arising from study design” refer to selection bias or is this phrase used more broadly to describe how exposure and outcome are measured and covariate data collected?

- Define what is meant by the phrase “bias arising from statistical analyses.” It is unclear if bias is the correct term, rather this may refer to model misspecification.
With regard to scientific justification and application of Consideration #3 (listed on page 2-7), the Panel recommends that EPA provide more discussion and clarity on the exclusion of null epidemiologic studies.

In Exclusion Criterion #3 (listed on page 2-7) EPA should define “reported dose.”

The Panel recommends that the discussion in Section 2 of the consideration of “confounding and other potential sources of bias” be clarified. The differences between males and females with regard to TCDD half-life are discussed, but the description of the number of males and females in each study population were often missing or very difficult to determine. Also, in the occupational cohort studies, the possibility of men and women performing different job tasks also increased the possibility that the men and women were exposed at different levels. However, when the job categories with assigned TCDD exposure levels were presented, there was often no discussion of the numbers by gender in the categories. For example, the Manz et al. study (1991) of the Hamburg cohort (1,583 men and 399 women) does not describe the TCDD categories by gender. In addition, the validity of the TCDD exposure levels assigned to the categories was examined “in a group of 48 workers who provided adipose tissue samples” (Page 2-41, lines 18-19). How were these workers selected? How many were approached but refused to provide a sample? Assessment of selection bias in this and other similar circumstances was lacking in some of the studies. This is particularly notable in the lack of overall response rates reported for several of these studies. Inclusion of these factors in the study review would be very helpful.

The Panel recommends that discussion of the consideration that “statistical precision, power, and study follow-up are sufficient” be clarified. These metrics can be difficult to determine with the smaller sample size populations, but there are studies that can be very useful even given the small samples. For example, the relative risks calculated for increasing TCDD exposure and risk of breast cancer in the Seveso study were greatly increased in the 3rd and 4th highest exposure categories, but the relative risks were not statistically significant (page 2-56, lines 1-8).

Charge Question 3. The Use of Toxicokinetics in the Dose-Response Modeling for Cancer and Noncancer Endpoints

In its Report, EPA used a physiologically-based pharmacokinetic (PBPK) model (Emond et al., 2004, 2005, 2006) with blood concentration as the dose metric rather than first-order body burden. The Panel was asked to comment on the scientific justification for EPA’s application of this model, the model modifications that EPA implemented, and
EPA’s characterization of uncertainty in the model. EPA also developed a PBPK model to estimate TCDD concentration in mouse tissues. The Panel was asked to comment on the scientific rationale for development of the mouse model, the performance of the mouse model, and whether model uncertainty had been adequately characterized. In addition, the Panel was asked to comment on the use of the Emond PBPK model to estimate human intake based on internal exposure measures, EPA’s sensitivity analysis of the kinetic modeling, and EPA’s estimate of lifetime average daily dose.

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

Please comment on the following:

3.1.a. Please comment on the justification of applying a PBPK model with whole blood TCDD concentration as a surrogate for tissue TCDD exposure in lieu of using first-order body burden for the dose-response assessment of TCDD.

Response:

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

The rationale for the use of blood concentration rather than lipid adjusted serum concentration (LASC) should not be based on the Emond model structure. It would be trivial to change the model so that LASC could be predicted. Indeed, the model is apparently used to estimate LASCs in the RfD calculations (e.g., page xli, line 21 in the Executive Summary of the Report). The question that should be addressed is only whether blood concentrations or LASCs provide better surrogates for cross-species and
cross-study comparisons of free dioxin concentration in the target tissues. LASC is the preferred measure for reporting dioxin biomonitoring data, and is the measurement reported in most of the human epidemiological studies. A metric that considers blood lipid content is also more likely to reflect free dioxin concentration in the plasma, and hence free concentration in the target tissue. The EPA pointed out (page xxxiv in the Executive Summary of the Report) that the LASC was related to the blood concentration by a scalar; however, EPA incorrectly concluded that the metrics are equivalent and later (page 3-511, line 6 of the Report) discussed the fact that the relationship between them was subject to inter-individual and inter-species variation. If the LASC were used to drive the distribution of TCDD to tissues, the pharmacokinetic outcome would be different from using blood as the driver because the tissue: blood ratio would differ. If the blood fat: blood and tissue: blood values were accounted for in the model, the use of blood and LASC would be similar. It’s not clear at this point how this issue was addressed in the dose metric calculations. Consideration of this issue is unlikely to drastically affect the outcome of the risk calculations, but it would be important for a quantitative uncertainty analysis.

**Recommendations**

- The use of the blood metric is acceptable for the PBPK model. EPA should clarify how the model deals with studies that report the concentration of dioxin in plasma, serum, blood or blood fat: blood measurements.

3.1.b.  *Please comment on the scientific justification for using the Emond et al. model as opposed to other available TCDD kinetic models.*

**Response:**

The Emond model provided the best available basis for the dose metric calculations in the assessment. It is the product of a high-caliber, multi-year research effort at EPA’s National Health and Environmental Effects Research Laboratory, and represents a significant effort in terms of data collection. This model builds on prior PBPK modeling efforts conducted by Andersen et al. (1997). However, additional discussion of other published models and quantitative evaluation of the impact of model selection on dose metric predictions should also be provided.

**Recommendations**

- The *Report* should discuss how the model was intended to be used in the assessment, which would then dictate why a particular model was selected. That is, for the intended purposes, was the Emond model more robust and/or simpler than other models, and did it contain sufficient details for biological determinants deemed important by the Agency?
3.1.c. Please comment on the modifications implemented by EPA to the published Emond et al. model.

Response:

The EPA modifications to the published Emond model (modifications described on page 3-44 of the Report account for volume of plasma and describe urinary clearance using blood concentration and not a lumped compartment) are minor and appropriate. The model changes are scientifically appropriate and well supported.

3.1.d. Please comment on whether EPA adequately characterized the uncertainty in the kinetic models.

Response:

The Report presents a reasonably thorough qualitative characterization of the uncertainty in the kinetic models that is sufficient to support their use in the assessment. A more quantitative uncertainty analysis is needed. Methods that could be useful and informative for such an analysis are suggested in the response to Charge Question 6.2. It is critical to demonstrate the dependence of human equivalent dose (HED) and risk predictions on uncertainty and variability in the model parameters, particularly those with high sensitivity (Evans and Andersen, 2000). Moreover, dose metric uncertainty needs 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.

The Hill coefficients for CYP1a1 and CYP1a2 induction used in the Emond model were 1.0 and 0.6, respectively, based on fitting of kinetic data from single doses of dioxin (Wang et al., 1997; Santostefano et al., 1998). However, Walker et al. (1999) subsequently estimated a Hill coefficient of 0.94 for both CYP1a1 and CYP1a2 induction using chronic exposures which were more relevant to the use of the Emond model in the dioxin risk assessment. The value of 0.6 used in the Emond model was well outside the confidence interval of 0.78 to 1.14 reported by Walker et al. (1999). The use of a Hill coefficient value well below unity would lead to a nonlinear model behavior that is biologically implausible (hypersensitivity to induction at doses near zero). As a result, when the human model was used for extrapolation to lower doses (as in the calculation of risk-specific doses) the model would tend to estimate a lower exposure level for a given blood concentration. This effect could be seen in Table ES-1 of the Report, where a 5 order-of-magnitude change in risk was associated with a 6 order-of-magnitude change in risk specific dose. That is, the model-estimated risk specific doses in the vicinity of $10^{-6}$ risk were about a factor of 10 lower (more conservative) than linear extrapolation. The evidence for this parameter needs to be carefully reviewed and the reasonable range of values determined. At the least, the human Emond model calculations will need to be repeated with multiple values to characterize the resulting uncertainty in the estimates.
When this is done, the Agency should also consider increasing the fat:blood partition in the human model from 100 to 200 to be more consistent with the human data (Patterson et al., 1988; Schecter and Ryan, 1989; Schecter et al., 1989; Iida et al., 1999; Maruyama et al., 2002). The Hill coefficient is not likely to have as significant an effect on calculations with the animal models, since low-dose extrapolation was not performed in the animals, but this should also be verified by sensitivity/uncertainty analysis of the animal models. Public comments were submitted to the Panel recommending consideration of a Hill coefficient value of 1.0 and pointing out why lower values are inappropriate (comments from Drs. Thomas Starr, July 7, 2010 and October 26, 2010 and Melvin E. Andersen, November 4, 2010).

**Recommendations**

- The Panel recommends additional efforts to fully characterize the uncertainty in the models with special consideration of the Hill coefficient value.

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

*Please comment on the following:*

3.2.a. Please comment on the scientific rationale for the development of EPA’s mouse model based on the published rat model (Emond et al., 2004, 2005, 2006).

**Response:**

The Panel agrees that an appropriate approach was used to develop the mouse model on the basis of the published rat model and the available mouse kinetic data. It should be noted that the NAS recommendation to use human data for dose metric could be accomplished because dose-dependent elimination of TCDD has been described in humans, albeit in just a few cases. Dose-dependent elimination has been reported repeatedly in animals and the PBPK model reflected this dose-dependence. Using CYP1A2 data from humans (caffeine metabolism) and mice would offer an opportunity to validate and/or adjust the mouse model.

**Recommendations**

- An external peer review of the mouse model should be conducted because this model has not been published in the peer-reviewed literature. This is typically a requirement for models to be used by the Agency.
3.2.b. Please comment on the performance of the mouse model in reference to the available data.

Response:

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

The EPA’s suggestion in the RfD chapter that the clustering of mouse points of departure (PODs) at the lowest doses was due to mouse model failure was inappropriate and should be rewritten.

Recommendations

- EPA should use the mouse model. The scientific credibility of the model will be enhanced by its publication in an appropriate peer reviewed journal.

3.2.c. Please comment on 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.

Response:

EPA provided 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 as Markov chain Monte Carlo analysis was not considered necessary or particularly useful. However, a more quantitative uncertainty analysis is needed.

Recommendations

- A more quantitative uncertainty analysis is recommended. Methods that could be useful and informative for such an analysis are suggested in the response to charge question 6.2 in this report.

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

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).

Recommendations

- The modeling of the Cheng et al. (2006), Mocarelli et al. (2008), and Bacarelli et al. (2008) 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).

Response:

The Report only presented the sensitivity analysis published by Emond et al. (2006), which was not entirely adequate for the purposes of this assessment. The analysis left out the Hill coefficient, which was one of the most important parameters in the model for low dose extrapolation (Evans and Andersen, 2000). Moreover, model sensitivities were species, dose, and dose-scenario dependent, so they need to be determined under the same exposure conditions as those for which dose metrics were 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. This represents the most pragmatic path forward for an evaluation of model sensitivity as it relates to potential environmental regulation.

Recommendations

- EPA should provide a sensitivity analysis of the model to authenticate the model for its intended purpose.

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

Response:

The Panel agrees with the average daily dose calculation approaches described in the Report. It was not clear to some Panel members how the computational estimates of
internal dose for newborns were carried out since a lactation model was not used. This is important because of the use of TSH in newborns as a critical effect. EPA, and Baccarelli et al. (2008), developed an empirical description of the relationship between maternal TCDD levels (lipid adjusted) in serum at birth of neonate and the measured serum TSH in the newborns up to 3 days of age. The Emond et al. model was run in an iterative fashion by adjusting chronic daily intake (ng/kg/day) in the human gestation model to predict maternal serum level of TCDD at term that was associated with infant serum thyroid stimulating hormone (TSH) concentration of 5 uU/ml (by using the regression equation). The result was 0.024 ng/kg bw/day.

**Recommendations**

- EPA should carefully explain how the early life stage internal doses are calculated.

**Charge Question 4. Reference dose**

EPA selected two co-critical studies (Mocarelli et al., 2008 and Baccarelli et al., 2008) for development of the reference dose for TCDD. The Panel was asked to comment on the scientific justification for selection and use of these studies to develop the reference dose.

4.1. *The Mocarelli et al. (2008) and Baccarelli et al. (2008) studies were selected as co-critical studies for the derivation of the RfD. Is the rationale for the choice of Mocarelli and Baccarelli scientifically justified and clearly described? Please identify and provide the rationale for any other studies that should be selected, including the rationale for why the study would be considered a superior candidate for the derivation of the RfD. Also comment on whether the selection of male reproductive effects and changes in neonatal thyroid hormone levels was scientifically justified and clearly described.*

**Response:**

The Panel found that use of the Mocarelli et al. (2008) and Baccarelli et al. (2008) studies was appropriate for identifying “co-critical” effects for the RfD calculation. These are human epidemiological studies that were well designed and executed. The studies provided sufficient exposure information, including biological concentrations that could be used to help establish acceptable life-time daily exposure levels. Some of the strengths of the human studies included the use of a well-characterized human cohort, conducted by dioxin epidemiology experts, and the fact that similar PODs were found across a broad spectrum of other reported dioxin toxicities in multiple species. The rationale for selecting these two studies over numerous other available studies was clearly described and the Panel believed that, overall, EPA provided a well-considered and rational discussion of why these two human studies were selected for determining the
RFID. However, one issue discussed by the Panel was that, while the strengths of the two
human studies were well-described, the study weaknesses were not always clearly
delineated. For example, in the Baccarelli (2008) study there was limited discussion of
how the presence of polychlorinated dibenzo-\textit{p}-dioxins (PCDDs), polychlorinated
dibenzofurans (PCDFs) and coplanar polychlorinated biphenyls (PCBs) that were also
found in the blood might confound the interpretation of TCDD association with elevated
TSH levels. In addition there was no discussion of the potential impact of residential
histories (e.g., individuals who may have moved in and out of Zone A after the accident).
The Panel believes that more discussion of the strengths and weaknesses of these two
studies is needed.

As indicated above, the Panel agreed that the major strengths of the human studies
were the use of a well-characterized dioxin-exposed human cohort, conducted by dioxin
epidemiology experts, and the fact that similar PODs were found across a broad spectrum
of other reported dioxin toxicities in multiple species. However, in isolation from each
other, and lacking a description of supportive animal and epidemiological studies, the
studies were less useful for setting the RFID. The Panel emphasizes the need to consider
these other supportive studies within the context of the weight of the dioxin and dioxin-
like compound (DLC) database. The strength of the RFID should not be based solely on
these two human epidemiology studies, but rather should be supported by integration
with other similar supporting dioxin and DLC studies. A strong voice from the
committee was given for looking at the comprehensive data base of both animal and
human epidemiological studies together to demonstrate a consistent and integrative signal
of toxicity across species and endpoints for TCDD. It was suggested that similar studies
with DLCs should also be included as these would be supportive, at least for a semi-
quantitative comparative analysis. This “collective” impact of the studies was stated in
the Report but needs to be made stronger as it represents the contextual framing for
understanding dioxin health impacts. This response would include discussions of both
human and experimental animal studies that have examined the effects of dioxin or DLCs
on other reproductive and endocrine endpoints and should, for example, include
discussion of dose-response relationships as well as comparisons of no-observed-adverse
effect levels (NOAELs) and lowest-observed-adverse effect levels (LOAELs).

The Panel notes that Figures 4.3 and 4.4 in the Report show quantitative
comparisons across the RFIDs and benchmark dose lower bounds (BMDLs) calculated
from the animal and epidemiological studies. These figures are useful in understanding
the quantitative similarities (to the PODs in the chosen studies) in these calculations. The
Panel also notes that since the figures did not have an indication of endpoints being
measured, just the reference to the publications, the consistency in signal (i.e., the
similarities in PODs determined) was not as readily apparent as it could be.

Although it has been addressed in the Report, the Panel recommends expanding
the discussion of the known human age-specific variability in endpoints such as sperm
counts, though the data from Moccarelli et al. (2008) do show ranges and variance (in
Figure 3/Table 2), and neonatal TSH levels.


**Recommendations**

- EPA should provide a discussion of the strengths and weaknesses of the Mocarelli et al. (2008) and Baccarelli et al. (2008) studies with an indication of whether the weaknesses affect determination of the RfD.

- EPA should label the endpoints for studies included in Figures 4.3 and 4.4.

- The comprehensive data base of both animal and human epidemiological studies, including studies with DLCs (e.g., studies cited in Goodman et al., 2010), should be discussed together to demonstrate a consistent and integrative signal of toxicity across species and endpoints for TCDD.

4.2. *In the Seveso cohort, the pattern of exposure to TCDD is different from the average daily exposure experienced by the general population. The explosion in Seveso created a high dose pulse of TCDD followed by low level background dietary exposure in the exposed population. In the population, this high dose pulse of TCDD was slowly eliminated from body tissues over time. There is uncertainty regarding the influence of the high-dose pulse exposure on the effects observed later in life.*

4.2.a. *Mocarelli et al. (2008) reported male reproductive effects observed later in life for boys exposed to the high dose pulse of TCDD between the ages of 1 and 10. EPA identified a 10 year critical exposure window. In the development of the candidate RfD, EPA used an exposure averaging approach that differs from the typical approach utilized for animal bioassays. EPA determined that the relevant exposure should be calculated as the mean of the pulse exposure and the 10-year critical exposure window average. Please comment on the following:*

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

**Response:**

The Panel discussed extensively, both as part of the deliberations on Section 4 of the Report and also as part of the discussion on Section 3, extrapolation issues posed by the pattern of exposure from Seveso. Issues raised included the question of whether the same endpoints and or dose-response would be expected from such exposure scenarios with high acute exposures when extrapolating to low-dose chronic exposures. It would be useful for EPA to provide a discussion of published examples in which dioxin studies were conducted using both high-dose acute and low-dose chronic exposures in animals for the same endpoint and how the outcomes compare both qualitatively and quantitatively. It would be important to determine whether similar results were observed
for similar endpoints. Several Panel members thought there were sufficient data in the immunological or reproductive areas that may allow such a comparison. Several chronic dioxin animal studies may be useful in this regard (Yoshizawa et al., 2009; Sand et al., 2010; Yoshizawa et al., 2010). The life stage-specific approach to hazard and dose-response characterization for children’s health risk assessment found in EPA’s Framework for Assessing Health Risks of Environmental Exposures to Children (EPA, 2006), is also relevant to addressing this issue and should be discussed. The Panel also recommends that the publication of Bell et al., (2010), which summarized and presented data on some differences about chronic vs. acute exposure in maternal transfer, be considered in this discussion.

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

Response:

The Panel found that changes from normal sperm counts and sperm motility are of public health relevance and therefore of interest for determining an RfD. Collectively, there was support for these endpoints within the context of the broader dioxin literature. The Panel discussed whether the magnitude of these changes would represent an adverse health effect. While the shifts observed in sperm counts may or may not pose a significant health effect in a single individual, such shifts on a population basis could presumably lead to an increased incidence of adverse health outcomes. Although there was concern expressed about the sample size used for sperm number and known variability in the biological endpoint, the Panel found that sample collection was conducted consistently across subjects and the differences in groups were apparent.

The Panel supports EPA’s approach of using the World Health Organization (WHO) reference value for determining relevant TSH levels, and the Panel strongly suggests that further discussion of WHO reference values for male reproductive parameters be included in the Report. Several references were available which provided background information and current values recommended by WHO regarding sperm counts (e.g., Skakkebaek, 2010). The Panel suggests that the standard deviations or range of changes from Mocarelli et al. (2008) be discussed in the Report because this provides a better understanding of the potential magnitude of effect.

Life stage differences in sperm counts were discussed by the Panel. Members of the public also provided comments on this issue. It would be appropriate to indicate in the Report that life stage differences clearly exist in sperm counts in humans and to cite and discuss the EPA life stage document (EPA, 2006).
Recommendations

- Discussion on WHO reference values for male reproductive parameters should be included in the Report (e.g., Skakkebaek, 2010).

- The standard deviations or range of changes from the Mocarelli (2008) study should be discussed in the Report to provide a better understanding of the potential magnitude of effect.

4.2.b. For Baccarelli et al. (2008), the critical exposure window occurs long after the high-dose pulse exposure. Therefore, the variability in the exposure over the critical exposure window is likely to be less than the variability in the Mocarelli et al. subjects. EPA concluded that the reported maternal exposures from the regression model developed by Baccarelli et al. provide an appropriate estimate of the relevant effective dose as opposed to extrapolating from the measured infant TCDD concentrations to maternal exposure. Additionally, EPA selected a LOAEL of 5 µ-units TSH per ml blood in neonates; as this was established by World Health Organization (WHO) as a level above which there was concern about abnormal thyroid development later in life. Please comment on the following:

4.2.b.i. Please comment on EPA’s decision to use the reported maternal levels and the appropriateness of this exposure estimate for the Baccarelli et al. study.

Response:

The Panel discussed and supports EPA’s decision to use the Baccarelli et al. (2008) estimates of the relevant effective doses. Since the bulk of the calculations were based on zonal averages, it should be made clearer how these measurements relate to ranges and variations in exposure in utero.

4.2.b.ii. Please comment on EPA’s designation of 5 µ-units TSH per ml blood as a LOAEL for Baccarelli et al.,( 2008.)

Response:

The change in TSH levels reported by Baccarelli et al. (2008) was of public health relevance and therefore of interest for determining an RfD. Collectively, there was support for this endpoint within the context of the broader dioxin literature. There was discussion on whether the magnitude of these changes would represent an adverse health effect. The Panel notes that the shift observed in TSH levels may or may not pose a significant health effect in a single individual, but such a shift on a population basis could presumably lead to an increased incidence of adverse health outcomes. The Panel also discussed the variability in neonatal TSH levels but concerns about this issue were
minimized by the fact that samples were all collected on the same postnatal day. The Panel suggests that if any follow-up data on thyroid hormone levels, such as T3, T4 or TSH levels, are available from the population studied, then these results should be discussed in the Report. The Panel discussed several studies describing health effects associated with elevated neonatal TSH levels not always recognized as associated with congenital hyperthyroidism (CH). There is a need to better describe the potential adverse health outcomes related to altered neonatal TSH levels. For example, in addition to effects on growth, both cognitive and motor deficits have been found in young adults with congenital hypothyroidism (Oerbeck et al., 2003; Oerbeck et al., 2007). The Report could better describe the consequences of transient hypothyroidism on reproductive outcomes e.g., see Anbalagan et al. (2010). Other references that relate to this question include: Chevrier et al. (2007), Dimitropoulos et al. (2009), and Yr (2008).

Recommendations

- EPA should better describe the potential adverse health outcomes related to altered neonatal TSH levels (e.g., effects on both cognitive and motor deficits).

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

Response:

A composite uncertainty factor of 30 (an uncertainty factor of 10 for the lack of a NOAEL, and an uncertainty factor of 3 for human interindividual variability) was applied to the LOAEL of 0.020 ng/kg-day from Mocarelli et al. (2008) to obtain the RfD. The Panel agrees that the appropriate uncertainty factors (UFs) were included. The exclusion or inclusion of the UFs in the Report is obvious, clearly discussed, and adequately rationalized. The Report would be more transparent if EPA included a short discussion of the basis for the decision not to include a UF for data quality.

4.4. EPA did not consider biochemical endpoints (such as CYP induction, oxidative stress, etc.) as potential critical effects for derivation of the RfD for TCDD due to the uncertainties in the qualitative determination of adversity associated with such endpoints and quantitative determination of adversity associated with such endpoints and quantitative determination of appropriate response levels for these types of endpoints in relation to TCDD exposure. Please comment on whether the decision not to consider biochemical endpoints is scientifically justified and clearly described.

Response:

Biochemical endpoints such as P450 activation, increased oxidative stress, etc. may be acceptable endpoints to establish PODs, particularly when the quantitative
relationship between the biochemical endpoint and an adverse health outcome is clearly evident. However, with respect to TCDD, the Panel agrees that more traditional endpoints (e.g., immune, endocrine, reproductive) are more appropriate because associations of these endpoints with health outcomes are well studied and provide a stronger association to an adverse outcome than biochemical endpoints. However, because of the wealth of data on P450s and their importance in disease development, normal development, and chemical response to exogenous agents, EPA should discuss biochemical endpoints, particularly P450s, relevant to establishing and strengthening the proposed reference dose.

4.5. *In using the animal bioassays, EPA averaged internal blood TCDD concentrations over the entire dosing period, including 24 hours following the last exposure. Please comment on EPA’s approach for averaging exposures including intermittent and one-day gestation exposure protocols.*

**Response:**

For animal studies it has been shown that for some effects acute exposure could give different results than chronic exposure. For TCDD, however, its persistence might suggest that such differences would be partly negated. In Baccarelli et al., (2008), there was extensive discussion regarding the use of the exposure average time for the TCDD concentrations. This is of biological significance as several papers have indicated the unique aspects of high peak exposure of TCDD as occurred in Seveso and in several of the animal studies. The endpoints affected as a result of these peaks do not always translate to impacts from lower chronic exposures. As stated earlier in this section, it would be helpful to discuss any available animal studies comparing high-dose acute vs. low-dose chronic effects on similar endpoints for dioxin or DLCs. By returning to the broader animal literature and using time and dose-response studies from the dioxin and DLC studies, biological support for the two critical endpoints might be found.

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

**Response:**

The Panel agrees with the BMD modeling approaches used in Section 4 of the *Report.* EPA conclusions that the animal data had sufficient limitations that precluded their use to establish a RfD are adequately justified. The reasons provided, however, are quite diverse, (e.g., no NOAEL, not considered an adverse effect, the effect at the LOAEL is too divergent from the control group, insufficient dose groups at the low-end of the dose-response curve, monotonic responses) and there is no way for the reader to determine which study has particular deficiencies without going back to the original paper. To help address this gap, the Panel suggests that several of the best animal studies be discussed in some detail so these limitations are more apparent to the reader. As
indicated previously, the EPA authors need to better cite the endpoint guidance that is present within EPA documents for defending these approaches and application of BMD models for the critical effects. This is especially necessary given public comments that EPA was not following its own guidelines.

4.7. For the animal bioassay modeling, EPA applied the kinetic extrapolation at the level of the POD prior to applying the uncertainty factors because EPA has less confidence in the kinetic model output at lower doses reflective of the RfD. Please comment on whether the kinetic extrapolation at the level of the POD prior to applying the uncertainty factors was scientifically justified and clearly described.

Response:

The EPA approach of applying the kinetics on the actual data present at the POD is preferred in this assessment (see additional discussion in the response to Charge Question 3 - The use of toxicokinetics in the dose-response modeling for cancer and noncancer endpoints).

4.8. Please comment as to whether EPA’s qualitative discussion of uncertainty in the RfD is justified and clearly described.

Response:

The Panel agrees that EPA provided a clear and justified discussion of the uncertainties in deriving the RfD using the Seveso cohort. Section 4 of the Report discussed study limitations regarding the need to adjust from acute exposure to average daily dose, the issue of critical windows, co-exposure to DLCs, and the strength/weaknesses of the animal data. The Panel agrees with EPA that the major limitation of the Seveso cohort is the uncertainty arising from how well the effects resulting from high-dose acute exposure translate to low-dose daily exposures. Again, it might be useful to re-review the animal studies to identify whether there are any studies where dioxin or DLCs were administered by acute as well as chronic (or even subchronic) exposure and comparable endpoints were examined. If so, the information can be used to help confirm or refute the accuracy of the “average daily dose” adjustment. This is of particular concern in the Mocarelli study as “time periods of susceptibility” appear in male reproductive development and these periods (windows) may be very short. Again, animal studies, particularly those involving male reproduction, may be helpful.

It would also be useful to include a discussion of potential uncertainty in the exposure estimates from the Baccarelli study. Serum dioxin levels were only established in a subset of the cohort (approximately 51) at the time of the study while dioxin levels from the main cohort were estimated from data collected from zone of residence (A or B) at a much earlier time.
The discussion in the Report of whether the background DLC exposure may have a significant impact, particularly at the lower TCDD exposure levels, is important. While the Panel agrees that the true DLC impact can’t be determined, it might be helpful to provide some general estimates of the variability that may occur at the proposed RfD.

5. Charge Question 5. Cancer assessment

In the Report EPA has provided: a weight-of-evidence characterization of TCDD as a known human carcinogen, conclusions regarding the mode of carcinogenic action for TCDD, EPA’s selection of data sets for cancer dose-response modeling, and consideration of approaches for assessment of TCDD carcinogenicity. The Panel was asked to comment on the scientific soundness of these aspects of EPA’s cancer assessment.

General Comment:

The Panel was impressed by the extensive work performed by EPA in its response to the NAS comments on cancer assessment. The comments below are intended to support the Agency in further developing Section 5 of the Report.

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

Response:

The Panel agrees on the classification that “TCDD is carcinogenic to humans” under EPA’s 2005 Guidelines for Carcinogen Risk Assessment. Available occupational epidemiologic studies provide convincing evidence of an association between TCDD and human cancer that cannot be reasonably attributed to chance or confounding and other types of bias, and with a demonstration of temporality, strength of association, consistency, biological plausibility, and a biological gradient. Additional evidence from animal studies and from mechanistic studies provides additional support for the classification of TCDD as carcinogenic to humans. A dissenting opinion (see Appendix A of this report) was expressed by one Panel member who indicated that at best, there is equivocal evidence for the carcinogenicity of TCDD in the occupational setting where body burdens were much higher than current or previous background levels. The Panel provides the following recommendations to strengthen the Report.
**Recommendations**

- EPA should provide more discussion of the power of studies used and the difficulties involved when assessing rare tumors. Thoroughly addressing these aspects will make the weight-of-evidence characterization in Section 5 of the Report more clear and transparent.

- In the weight-of-evidence characterization, the Agency should build on all the available data to support the decision. It needs to be made clear how different types of data (in vitro, in vivo, human) support each other; or not.

- EPA should consider including studies with substantial DLC exposure where TEFs can be calculated. Specific experimental studies include Li and Rozman (1995), Rozman et al. (1993, 2005), and Viluksela et al. (1994, 1997a,b, 1998a,b).

- EPA should attempt to characterize the uncertainty regarding the carcinogenicity of TCDD at low human exposures, since the minimum dose at which carcinogenic effects would be expected to occur cannot be clearly delineated from the current epidemiological human data. The agency has concluded that AhR activation is a necessary but not sufficient precursor event in the carcinogenic activity of TCDD. Therefore, it would be beneficial if the Agency could evaluate available data on AhR activation and related effects in human cells and animal models to help inform the doses at which these precursor events are observed for comparison with the epidemiological data.

**5.2. Mode of Action:** The mode of action of a carcinogen can inform identification of hazards and approaches used for a dose-response assessment. The mode of carcinogenic action for TCDD has not been elucidated for any tumor type. EPA concluded that, while interaction with the Ah receptor is likely to be a necessary early event in TCDD carcinogenicity in experimental animals, the downstream events involved are unknown.

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

**Response:**

The Panel appreciates the attempts by the Agency to further develop cancer mode of action concepts based on available dioxin liver, lung, and thyroid toxicity data. Such innovative and explorative work is clearly fundamental to the continued need to further develop risk assessment sciences and make more detailed and integrated use of already existing and published data.
The Panel complements the Agency for providing an up-to-date dioxin cancer mode of action section in its response to NAS comments. It could, however, be improved by incorporating additional data on linear and nonlinear modes of action in different target tissues and life stages. A large amount of data related to the mode of action for the carcinogenicity of TCDD is described in the Report, but the focus appears to be on presenting evidence that supports the use of a default linear approach rather than providing a balanced evaluation of alternative mode of action hypotheses.

The discussion of the likely dose-response for receptor mediated processes focuses only on the first step, binding of the agonist to the receptor, which is ultimately linear at low concentrations. However, no discussion is given to the nature of the dose-response for the down-stream sequelae of receptor activation, for which there is evidence of nonlinearity. It is, in fact, the fundamentally nonlinear nature of the dose-response for receptor mediated processes that underlies the conviction of a large segment of the scientific community that a nonlinear approach should be preferred for the risk assessment for dioxin.

**Recommendations**

- EPA should further expand the discussion of mode of action data available to delineate linear versus nonlinear modes of action and effects in different target tissues at different life stages.

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

**Response:**

The Panel notes that much is known about TCDD toxicity and mode of action. Some Panel members indicated that the characterization of the mode of action should be “reasonably well known” rather than “largely unknown.” Nevertheless, the Panel agrees that the exact mechanism-of-action has not been fully delineated for any distinct TCDD-toxicity end-point. For example, it was pointed out that most TCDD toxicities are mediated by activation of the AhR. Many studies have demonstrated that TCDD can activate or interfere with the activity of estrogen receptors, as well as other steroid receptors. Such interference can disrupt the regulation of cell proliferation, cell death and tissue differentiation. By disrupting these cell functions, TCDD can have profound and lasting effects as demonstrated by studies showing that TCDD exposure during development produces adult neural dysfunctions.
Recommendations

- EPA should provide a discussion of the evidence for possible modes of action that include both linear and nonlinear alternatives.

- EPA should describe the receptor mediated nonlinear mode of action for dioxin (e.g., Van den Heuvel et al., 1994; Li and Rozman 1995; Andersen et al., 1997; Bhattacharya et al 2010; Gim et al., 2010) and DLCs (e.g., Rozman et al., 1993; 2005, Stahl et al., 1994; Viluksela et al., 1994, 1997a,b, 1998a,b), as well as evidence regarding the fundamentally nonlinear nature of receptor mediated cellular responses (e.g., Andersen et al., 1999; Louis and Becskei, 2002; Zhang et al., 2010).

5.3. Is EPA’s approach for selecting data sets from the key epidemiologic studies and animal bioassays identified for cancer dose response modeling scientifically justified and clearly described?

Response:

The Panel agrees with the inclusion of the Cheng study in the cancer assessment. The study incorporated information on gradation of exposure. Expanded discussion of several other studies would support the weight-of-evidence for carcinogenicity in less common cancers such as lymphomas and soft tissue sarcoma. The Panel discussed the possible value of including studies with DLCs in the evaluation of the weight-of-evidence, in light of the small number of studies involving primarily exposure to TCDD. There are a number of studies in the literature involving the health effects of DLCs. These include rice oil poisoning incidents in Japan and Taiwan. These incidents have been described, and additional references have been provided, in Schecter and Gasiewicz (2003).

Recommendations

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

5.4. For the animal bioassay data, potential cancer oral slope factors (OSFs) were calculated by linear extrapolation (using a linear, non threshold cancer approach) from the point of departure (POD). EPA also estimated the composite risk of the occurrence of several tumor types from the animal cancer bioassay data.
5.4.a. Please comment on whether the approach for estimating cancer risk, including the use of tumor modeling of the TCDD animal cancer bioassay data, is scientifically justified and clearly described.

Response:

The Panel agrees 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.

Response:

The Panel notes the consistency of the selection of the BMDL01 as the POD with Agency guidelines and has no further comments.

5.5. 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.

Response:

The Panel agrees that Cheng et al (2006) is the appropriate study for OSF development. 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.

Response:

The Panel agrees 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.
Response:

The Panel agrees that the use of the Emond model to estimate risk-specific doses from the Cheng et al. (2006) dose-response modeling results is scientifically justified and clearly described. This is because the “concentration-and-age-dependent elimination model” (CADM) used in Cheng et al. (2006) did not facilitate this process. Also, the dose conversions were consistent with those used in the derivation of the RfD. However, as discussed in the response to charge question 3.1.d, the Panel is concerned about the value of the Hill coefficient used.

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 risk itself at a single risk level (BMR) for extrapolation to the origin. Please comment on EPA’s choice of extrapolation approach.

Response:

Since the fat concentrations generated by CADM were not linear with the oral exposure at higher doses, a single oral slope factor to be used for all risk levels could not be obtained. EPA used the upper 95% bound on the slope (from Cheng et al., 2006) of the linear relationship between the natural logarithm of the rate ratio and the cumulative fat TCDD concentration (fat-AUC) to estimate risk-specific doses for TCDD at all risk levels. The Panel agrees that the Agency has chosen the appropriate extrapolation approach.

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.

Response:

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

Recommendations

- EPA should expand the discussion in the Report 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. Please comment on whether EPA has clearly described the major qualitative uncertainties in the derivation of the OSF.
Response:

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

5.7. *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.*

Response:

While the Panel found that it was important to include DLC studies in the weight-of-evidence analysis, we are conflicted on their use as a source of dose-response estimates for TCDD. The Panel notes the scientific importance and regulatory relevance of including a coordinated TEQ/DLC discussion in the Report. Including TEQ/DLC aspects in the evaluation would allow for the use of additional studies with dose-response information that more closely mirror environmental exposures. On the other hand, the Panel recognizes the complications associated with developing a TCDD risk estimate that is dependent on current TEF values.

Recommendations

- DLC studies should be considered in the weight-of-evidence discussion.

5.8. *The NRC suggested that EPA consider nonlinear approaches for the assessment of TCDD carcinogenicity. In the Response to Comments, EPA presents two illustrative nonlinear approaches for cancer, but considers both inappropriate to use because lack of MOA information.*

5.8.a. *Please comment on these two illustrative nonlinear approaches including EPA’s conclusions regarding the limitations of these approaches.*

Response:

EPA’s *Report* did not respond adequately to the NAS recommendation to adopt “both linear and nonlinear methods of risk characterization to account for the uncertainty of dose-response relationship shape below the ED01.” Instead of adopting both linear and nonlinear methods, the EPA argued that only a linear approach could be justified, and derived two examples of RfD development using a nonlinear approach that they characterized as an illustrative exercise only. The choice not to include both linear and
nonlinear risk assessment approaches for TCDD was inconsistent with the EPA (2005) cancer guidelines (page 3-23/24):

“Nonlinear extrapolation having a significant biological support may be presented in addition to a linear approach when the available data and a weight-of-evidence evaluation support a nonlinear approach, but the data are not strong enough to ascertain the mode of action applying the Agency’s mode of action framework.”

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

**Recommendations**

- EPA should present both linear and nonlinear risk assessment approaches. In the absence of a definitive nonlinear mode of action, the linear option results can serve as the baseline for comparison with other estimates. The examples in the current document should be formalized and extended to allow for such a comparison.

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

**Recommendations**

- Since the EPA nonlinear analysis only used studies in S-D rats that were identified in Section 2 of the Report for potential noncancer dose-response modeling, additional alternative PODs should be added. For example, Simon et al. (2010), which was cited in EPA’s Report, provided a number of alternative PODs for a nonlinear approach that should be included in the EPA risk assessment.

**Charge Question 6. Feasibility of Quantitative Uncertainty Analysis**

In its evaluation of EPA’s 2003 Reassessment, the NAS committee recommended that EPA improve the transparency, thoroughness, and clarity in quantitative uncertainty analysis (QUA). Section 6 of EPA’s Response to NAS Comments document addresses NAS comments regarding QUA. The Panel was asked to comment on: whether Section 6 of EPA’s Report was clearly presented and scientifically justified; EPA’s conclusion that
a QUA is not feasible; the discussion of volitional uncertainty, and the utility of the
limited sensitivity studies presented by EPA.

6.1. Please comment on the discussion in this Section. Is the response clearly
presented and scientifically justified?

Response:

As discussed below, the Panel found that Section 6 of EPA’s Report is clearly
presented, but it is not scientifically justified. In particular, the Panel disagrees with
EPA’s argument that, since the most detailed and complete available methods for
conducting a quantitative uncertainty analysis (QUA) are unfeasible, no analysis can be
conducted at all. There are a number of approaches for conducting a QUA, some of
which are feasible given current knowledge and data. Specific methods that can be
implemented in a timely manner using available data and knowledge are suggested.

Clarity of the EPA response to the NAS presented in Section 6 of the Report

The EPA response is clearly presented. 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. Section 6 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.

Section 6 of the Report provides many useful insights for the Agency’s dioxin
reassessment. However, in its discussion of available methods, the Report is somewhat
biased in its treatment of certain statistical methods (discussed below) which could
address some of these issues (though the Report does note the potential contribution of
the methods at the end of Section 6, as part of ongoing or future studies) and overly
pessimistic regarding our ability provide improved quantitative estimates for certain
portions of the toxicity assessment.

Some Panel members indicated that the whole section should be rewritten to make
it more accessible to non-statisticians. As further discussed in the editorial comments on
Section 6 in Appendix C of this report, some phrasing and word choices in the text
should be reconsidered, in particular “exotic methods,” “volitional uncertainty,” and
“epistemic uncertainty.” The Panel found that the definition of “quantitative uncertainty
analysis” was overly narrow and should be expanded to embrace other common and
useful methods discussed below. In a few other places, the Report’s wording in Section 6
is strongly at variance with the literature on uncertainty analysis (see editorial comments
in Appendix C of this report).
Scientific justification of the arguments presented in Section 6

The Panel found that the arguments in Section 6 are not scientifically justified. In Section 6, EPA’s decision to not do an integrated quantitative uncertainty analysis is presented and a variety of theoretical issues are discussed, but EPA’s decision may be based primarily on grounds of practicality or timeliness. EPA indicates that a complete quantitative uncertainty analysis would require data and resources not available. We disagree with this logic. More limited evaluations can, and should, still be implemented to inform critical issues in the dioxin reassessment. EPA should be methodical in considering what variables and components of the assessment would be included in the analysis. The Panel found that the uncertainty narratives and sensitivity analyses already in the document are an excellent beginning and may constitute the lion’s share of the work necessary to implement quantitative uncertain analysis based on simple bounding.

The Panel does not concur with the specific argument EPA used to justify not doing a unified QUA. If the answer to the question of why EPA did not undertake one is that it was not possible to specify precise marginal distributions and dependence functions from existing data, then the conclusion would be that EPA has not been responsive to the NAS criticism, because there are many possible approaches that could be used that do not depend on such specifications. If the argument is that EPA guidance doesn’t require a QUA, then one might agree that the NAS criticism is perhaps itself unreasonable. If EPA had asserted that it actually had done an uncertainty analysis in the form of uncertainty factors (UFs) and the limited sensitivity studies that were performed, then that might be understandable, though not consistent with the current state-of-the-art in risk and uncertainty analysis. Even if the argument had been that mounting a QUA is a significant and controversial undertaking itself and that doing one shouldn’t delay the finalization of the Report, then such a practicality argument would be understandable given the protracted delay in completing the dioxin reassessment.

Instead, EPA asserts that “Data are the ultimate arbiter of whether quantitative uncertainty analysis … has sufficient evidentiary support.” This flies in the face of how uncertainty analyses are normally conceived. Of course, the absence of data is never a substantive reason not to conduct an uncertainty analysis; it is the reason to do one.

In its Report, EPA indicates that it needs an “underlying distribution from which to sample” in order to conduct a quantitative uncertainty analysis. The Panel notes that this is not necessarily true, and it is facile to shrug off a call to characterize and account for important uncertainties in the assessment process on these grounds alone. If one can estimate the value of a quantity, then one should be able to express the uncertainty about the value, otherwise one does not really have a scientific measurement in the first place. One is not forced to identify precise probability distributions and dependence functions for everything that is to be characterized as uncertain. Even when the uncertainty is volitional (or decisional or just model uncertainty), there can be relevant ranges that are interesting to decision makers and stakeholders. In some cases, the analysis may be
formally closer to a sensitivity analysis, but some appropriate response is usually possible, if not always practicable. To its credit, EPA has acknowledged the legitimacy of the call for QUA by NAS and undertaken some efforts in this direction.

In the Report, EPA calls uncertainty analysis an “emerging area in science” and this is inarguably true, but it does not seem reasonable to hold that methodological research is necessary for EPA to do anything more comprehensive to respond to NAS’s criticism, even if we disallow the use of expert elicitation. Under a commitment to the idea that analyses be data-driven, it is possible to do something that’s useful, even if it is not predicated on precise distributions. There are a variety of ways to conduct a quantitative uncertainty analysis, even an entirely probabilistic one that obeys the Kolmogorov axioms (Gillies, 2000) that require neither extensive data nor expert elicitation. The response to Charge Question 6.2 below provides a list of various ways (with references) to accomplish this. The list includes probability trees or model choice trees that articulate the structure of the model and dependencies, sensitivity analyses, simple interval analysis that just propagates the plausible ranges, and the supervaluation approach that uses nested inner and outer intervals (with the inner range representing the values that most everyone considers to be plausible values and the outer range representing conservatively broad ranges). There is also a continuous and unbounded version of nesting intervals in an approach known as info-gap analysis that would be useful if one cannot develop finite bounds on some of the inputs. One can also propagate bounds on distribution functions, so whatever imperfect information about each input variable’s distribution is available, one can fashion bounds on distribution functions and propagate them through the calculations, with or without assumptions or information about the dependencies among variables.

The Panel notes that the approaches mentioned above require EPA to make certain modeling judgments, in the same way that developing any analysis requires judgments. However, this does not mean that analysts would be required to make up numbers or elicit any expert opinion. Such an analysis does not necessarily require a lot of extra work by EPA. These methods can be simple to develop, and they are mostly computationally trivial. Of course, the more comprehensive the analysis is, the harder it is to complete. But the analysis does not have to be fully comprehensive to provide useful insights.

We note that there was not perfect consensus among Panel members about the value of a quantitative uncertainty analysis. Some on the Panel agree that an uncertainty analysis is not an absolute good. For instance, if the final answer is already clear, an uncertainty analysis can be a waste of time and resources. It would not be reasonable to insist on another analysis which would merely waste time and resources. Likewise, if the analysis is done poorly, or without appeal to available evidence from the real world, it can be misleading. For instance, the idea, mentioned in footnote 66 on page 6-20 of EPA’s Report, of arbitrarily converting uncertainty factors to independent lognormal random variables in a scattered attempt to mount a QUA would entail a suite of unjustified and probably untenable assumptions rendering the exercise nearly pointless.
Finally, if the analysis is used *strategically* to avoid rendering or finalizing a decision that is proper, it can be counterproductive. However, most members of the Panel felt that quantitative uncertainty analysis is an integral part of any good assessment, and that one is essential to address the many empirically unresolved questions and issues that have arisen in this assessment which beg for explicit consideration in the context of an uncertainty analysis. In its discussion of the other charge questions, the Panel has identified a number of important issues that should be addressed in an eventual uncertainty analysis.

**Other methods to be considered**

The Panel found that relevant Bayesian methods have been inadequately addressed and improperly dismissed in Section 6. In particular, methods that should be given a more extensive and balanced discussion with more citations to the literature include: 1) Bayesian hierarchical modeling (Axelrad et al., 2007; Choi et al., 2010; Coull et al., 2003; Ryan, 2008) which is used for combining information from multiple studies, and 2) Bayesian model averaging (Morales et al., 2006; Viallefont et al., 2011; Wheeler and Bailer 2007, 2009) which would be useful for considering more than one dose-response equation, while allowing the data to weight their *relative* likelihood and contribution to the estimate. These Bayesian methods should not be referred to as “exotic.” For example, in agreeing with the Section 6 authors that these methods should be pursued in ongoing and future case studies, White et al. (2009) refer to them as “advanced,” rather than exotic. Specifically, they recommend that health scientists should explore statistical approaches to model selection and suggest that “improvements to statistical approaches for model selection, such as model averaging, should be pursued. Case study applications of these advanced statistical approaches will identify potential strengths and weaknesses of the approaches and their significance for risk characterization” (White et al., 2009).

**Recommendations**

- The Panel recommends that EPA revise Section 6 of the *Report* because, as discussed above, the arguments in this section are not scientifically justified. In particular, EPA should consider revising its argument that quantitative uncertainty analysis is unfeasible for the dioxin assessment. Specific suggestions regarding feasible methods for quantitative uncertainty analysis are provided herein.

**6.2. Please comment on EPA’s overall conclusion that a comprehensive quantitative uncertainty analysis is not feasible.**

**Response:**

The Panel rejects EPA’s argument that a quantitative uncertainty analysis is unfeasible. Although a quantitative uncertainty analysis is challenging, the Panel does not agree that it is impossible or even impractical to undertake one. While it may well be
true that we lack an adequate empirical basis for full Monte-Carlo propagation of input
distributions, there are many other options available. Many on the Panel indicated that
the present circumstances warrant a compromise approach that would be simple and
achievable with modest effort by the Agency. Various bounding approaches, sensitivity
studies, uncertainty set analyses, and event trees (probability trees without the
probabilities) are suggested as possible approaches that could be used. With such
methods, legitimate and comprehensive uncertainty analyses (including even fully
probabilistic analyses) are possible. They would be useful and sufficient to respond to
NAS’ criticism.

The Panel agrees with EPA’s assertion that expert elicitation would be
problematic and should be “off the table.” However, many on the Panel further
suggested that value-of-information methods would also be very useful, although
feedback from EPA included reservations about this idea. A discussion of value of
information methods is provided in Appendix B of this report.

The Panel considered the use of bounding approaches for quantitative uncertainty
analysis and asked EPA to provide information about the limitations of bounding
approaches. In response, EPA asked Dr. Roger Cooke to send the Panel a document on
bounding analysis. The short bounding analysis document provided to the Panel by Dr.
Cooke focused on the features of interval analysis, although this is not by any means the
only approach that might be useful in the context of the dioxin assessment. The bounding
analysis document mentions one issue that could be construed as a disadvantage of this
simplest bounding approach. It is the idea that the ranges are supposed to be absolute
bounds on the possible values of each input variable. So, for instance, the only thing one
can say about a percentage is that it is between zero and 100%, or the only thing one can
say about a dispersal distance is that it is between zero and the circumference of the Earth
(these are Dr. Cooke’s examples). But the Panel finds that this criticism seems to
represent a misunderstanding of the word “absolute.” Vacuous (e.g., physically limiting)
bounds are not the only bounds that can be used in interval analysis. In fact, they are
meant to be informed by observed study results. Furthermore, one is not necessarily
limited to interval ranges and interval analysis.

The Panel suggests that there are in fact a variety of methods that, with proper
application, could be useful and informative, including:

- **Sensitivity analysis studies** (even if not completely comprehensive) (Saltelli et
  al., 2000a,b; Frey and Patil, 2002),
- **Interval analysis** (Moore 1966; Neumaier, 1990) which has been widely used for
decades and can be applied to complex models and even blackbox models (Trejo
and Kreinovich, 2001),
- **Nesting of intervals**, e.g., two levels, wide and narrow can give conservative and
  optimistic characterizations of overall uncertainty (van Frassen, 1966, 1980),
There are, of course, many significant benefits to undertaking a quantitative uncertainty analysis. Although a completely comprehensive analysis might indeed be too much to expect, it is possible and practical to provide readers with much more useful
information about uncertainty. A policy maker might reasonably expect the Report to provide insight into major uncertainties and questions such as the following:

- How likely is it that TCDD is not a human carcinogen at current exposure levels? Full discussion of this uncertainty may help to overcome probability neglect and action bias (Patt and Zeckhauser, 2000).
- How likely is it that TCDD at current exposure levels has health effects that have not yet been identified in the toxicological or epidemiological literature (Diamanti-Kandarakis et al., 2009; Soto and Sonnenschein, 2010)?
- What is the probability that reducing TCDD exposures would not reduce cancer risk at all, or only by amounts that would not be measurable, based on recent epidemiological studies and updates such as Pesatori et al. (2009)?
- What is the probability that reducing TCDD exposures would reduce cancer risk in the whole U.S. population, or targeted subpopulations, by amounts significantly greater than a prediction derived from the cancer slope factor estimated by EPA?
- What is the probability that reducing TCDD exposures would increase cancer risk (e.g., if the dose-response relation is J-shaped or U-shaped)?
- What is the decision-analytic value of information (VoI) from collecting more information on AhR kinetics and dose-response before making risk management decisions?
- What is the probability that TCDD interacts with other compounds to which U.S. or targeted subpopulations are exposed, increasing cumulative risk for cancer or other health effects (Carpenter et al., 2002)?

Although many members of the public believe that it is imprudent or even morally wrong to delay tighter regulation of TCDD exposures (perhaps reflecting beliefs that TCDD is a potent carcinogen, developmental toxin, etc.) many on the Panel felt that EPA should provide a thorough quantitative decision analysis that makes explicit the current uncertainties and trade-offs and that shows the conditions under which acting now or postponing action are the optimal actions. Without such quantitative analysis, risk management decisions for TCDD will not be adequately informed, and principles other than those of rational decision making (e.g., the biases discussed in Sunstein and Zeckhauser, 2010) may dominate risk management decisions for TCDD. EPA’s uncertainty analysis should provide the scientific basis for improved decision making. The current decision, in effect, to “punt” on quantitative uncertainty analysis is not adequate for informing responsible risk management decision and policy-making, and is not justified.

**Recommendations**

- The Panel recommends that EPA reconsider the argument for not doing a quantitative uncertainty analysis, or undertake one. EPA could follow the recommendation of the NAS on this point by using one or more of the techniques suggested above.
6.2a. Please comment on the discussion in Section 6 regarding volitional uncertainty and how this type of uncertainty limits the ability to conduct a quantitative uncertainty analysis.

Response:

In the Report, EPA contrasts volitional uncertainty with cognitive uncertainty. The Panel recommends that the term “volitional uncertainty,” which might also have been called “decisional uncertainty,” should be dropped from the Report. EPA should focus instead on uncertainties about the state of world and display the different modeling choices and the consequences of making them. The decisions mentioned in the discussion in Section 6 of volitional uncertainty are modeling choices, and they should be dealt with using techniques for model uncertainty. Standard tools and techniques for analysis of model uncertainty can be applied.

Recommendations

- The Panel recommends that EPA delete from the Report the notion of “volitional uncertainty.” EPA should display the different modeling choices and the consequences of making them.

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

Response:

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

The Panel is mindful of the need to minimize further delay of the finalization of this already protracted dioxin assessment. The work the EPA has already done in the sensitivity studies should be leveraged to hasten the completion of whatever uncertainty analysis EPA elects to undertake.
Recommendations

- The Panel recommends that sensitivity studies that have already been completed be integrated into whatever overall uncertainty analysis EPA elects to undertake.
REFERENCES


8 Baccarelli, A., A.C. Pesatori, S.A. Masten, D.G. Patterson, Jr., L.L. Needham, P.


10 Baccarelli, A., S.M. Giacomini, C. Corbetta., M.T. Landi, M. Bonzini, D. Consonni, P


Emond, C, L.S. Birnbaum, and M.J. DeVito. 2006. Use of a physiologically based pharmacokinetic model for rats to study the influence of body fat mass and induction of CYP1A2 on the pharmacokinetics of TCDD. Environmental Health Perspectives 114:1394-1400.


Rozman, K.K., M. Lebofsky, and D.M. Pinson. 2005. Chronic toxicity and carcinogenicity of 1,2,3,4,6,7,8-heptachlorodibenzo-p-dioxin displays a distinct dose/time threshold (c x t = k) and a profound subthreshold hormetic effect: prolongation of life. *Food and Chemical Toxicology* 43:729-40.


Appendix A: Dissenting Opinion from Karl Rozman, Ph.D.

The University of Kansas Medical Center
Karl K. Rozman, Ph.D.
Professor
School of Medicine
Department of Pharmacology,
Toxicology and Therapeutics

December 9, 2010

Thomas Armitage, Ph.D.
Designated Federal Officer
USEPA Science Advisory Board (1400R)
1200 Pennsylvania Ave., N.W.
Washington, D.C. 20460

RE: A Dissenting Opinion

Dear Tom,

As I have indicated in my previous written and oral opinions to this panel, I disagree with the panel conclusions regarding the carcinogenicity of TCDD and the adequacy of the EPA response to the criticisms of the NAS report.

There is at best equivocal evidence (statistically not significant) for the carcinogenicity of TCDD (or DLCs) in the occupational setting where the body burdens were at least 100 or 1000 times higher than the current or previous background levels. Therefore, the consideration of a practical threshold for any defined population requires acceptance of the compelling scientific conclusion that there is negligible (essentially zero) carcinogenic risk at current background levels which are much lower than past levels. Any other conclusion is incompatible with sound science and no amount of modeling or data manipulation will transform a non-existing effect at occupational exposure levels into a risk at current background levels other than the non-scientific, policy-driven, non-threshold extrapolation by EPA.

Further, it is my opinion that the EPA document (2010 Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments, 600/-10/038A) is deliberately non-responsive to the recommendations of the NAS report.

Respectfully,

Karl K. Rozman, Ph.D., D.A.B.T.
Professor

3901 Rainbow Blvd., MS1018, Kansas City, KS 66110 • (913) 588-7717 • FAX (913) 588-7501
krozman@kumc.edu
Appendix B: Value of Information

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

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

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

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

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.
Appendix C: Editorial Comments and Corrections

a. Section 2

Minor suggestions to further improve clarity regarding data set inclusion/exclusion criteria

Page xxxvii, Lines 16-19. The sentence needs clarification. It currently gives the impression that studies that were eliminated for further analysis would have NOAELs available.

Pages 2-234 – 2-247. EPA should consider adding information to the appendices and/or tables to provide readers with clarification regarding the exclusion of particular studies. For example, an extra column in Table 2-7 listing, by numbered reference, the criteria that were or were not met for each study would be helpful.

b. Section 6


Page 6-3, bottom: The word “margins” should be “marginals.”

Page 6-3, line 26: If EPA wants to use the adverb “always”, the phrase “as a joint distribution” should be “as some characterization of a joint distribution” to be correct.

Page 6-4, lines 9-12: This text is strange and off-putting. A reader might ask who wrote this and why. It seems opinionated and unnecessary.

Page 6-4, line 9: The tone is too pedagogical (“This is not the place . . .”).

Footnote 54: The discussion of alternatives to strict, single-measure probability theory is ham-handed. Neither interval probabilities nor imprecise probabilities (sensu Walley, 1991) depart from probability theory; they follow the Kolmogorov axioms. They are motivationally and essentially equivalent to sensitivity analyses, except they do not make use of sampling strategies and can be more comprehensive.

Pages 29-30: It is simply untrue that sensitivity analyses have to be systematic. The word “systematic” might better be “comprehensive” and the word “essential” should be weakened, perhaps to “advantageous”.

Page 6-5, lines 4-7 and footnote 55: There seem to be only two axioms mentioned in the text, but Kolmogorov needs three to make probability theory.

Page 6-5: The meaning of the phrase “epistemic uncertainty” given on this page is plainly incorrect. Epistemic uncertainty is the uncertainty that arises from imperfect
knowledge such as from limitations on the amount or quality of data available or
deficiencies in our causal understanding about a system. It is not true that a quantity
about which there is epistemic uncertainty is necessarily fixed. Although it is perhaps
clear how one might come to this mistaken impression, no researchers use the phrase to
imply that the underlying quantity has no variability (although all would admit that this
could be the case given our ignorance about it). Indeed, a variable can have both forms
of uncertainty. For example, when body weight varies across a population, but with a
distribution that is unknown, the variable has both aleatory and epistemic uncertainty.
This mistake echoes in a couple of other places throughout this section.

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

Page 6-5: The words “aleatoric” and “aleatory” are both used on this page as
(synonymous) adjectives of uncertainty. Actually, in the engineering literature, only
“aleatory” is preferred for this use. In any case, please pick one to use.

Page 6-5, line 10: The assertion that the frequentist and Bayesian interpretations are not
mutually exclusive may be misleading. They are mutually exclusive in the sense that it
would be improper to mix and match components of each into an analysis. It would be
appropriate to omit the clause with the phrase “mutually exclusive,” although it is surely
fair to say that subjective probabilities can and do track relative frequencies.

Page 6-5, lines 30-32: The text on the subject of dependence is strange here, and also in
section 6.1.3.3. It is incorrect that the “[i]ssues involving…epistemic and aleatory
uncertainty translate into issues of dependence.” This is just wrong (even under their
unusual definition of “epistemic”).

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

Page 6-6, line 20: Maybe the last word of the header should be plural.

Line 21: Modern practice has replaced “error” with “uncertainty” in this context.

Footnote 56: EPA could add “or subtracting” after “adding.”

Page 6-7, line 14: “The role of dependence modeling” should be replaced with “Dependence among variables.”

Page 6-7. More examples of use of expert judgment for health assessment are available and should be cited.

Page 6-7, last paragraph: This paragraph extending onto the next page should be rewritten. The example is reasonable and important, but the discussion about it is confused. The first sentence is incorrect. The uncertainty mentioned in the second sentence may be epistemic, but the sentence is erroneous in its claim. In the following sentences, the words “variable” and “fixed” (or “constant”) should be used rather than “aleatoric” and “epistemic.” It is nonsense to say that a kinetic constant is “completely correlated across individuals.” It’s not correlated; it is invariant. This case is not an example of a dependence issue. There is no correlation between a distribution and a fixed quantity (even if it’s uncertain). Correlation is defined between varying quantities.

If the number is fixed, whether or not we know what it is, then one cannot say it is correlated with anything. The authors may have come to this twisted language because they’re thinking of the uncertainties in terms of how they might plan to quantitatively characterize them in a Monte Carlo simulation (repeatedly selecting a random deviate for the kinetic constant but assigning it to every individual). Of course, variables such as body fat, age, and smoking, on the other hand, can and do exhibit correlations that definitely should be accounted for in the quantitative assessments. Likewise, it is also important to keep track of the constancy of particular quantities about which we may not know the precise value. These two issues should be untangled and discussed in a less confusing way.

Page 6-8, line 12: The first paragraph of section 6.1.3.4 seems to be saying that one can sometimes express model uncertainty as parametric uncertainty, which simplifies its handling. This could be said more plainly. It would be helpful to mention that this trick cannot always be used (as when the possible models cannot be listed). It might also be especially helpful to mention that this trick is not so much a way to propagate model uncertainty as a way to sweep it under the rug. Model averaging, including Bayesian model averaging (which is mentioned in several places, including 6-36, lines 3ff), erases...
model uncertainty in the same way that averaging variable quantities erases their variation.

Page 6-8, line 13: Omit the unnecessary fancy after the semicolon.

Lines 15-17: This sentence is nonsense, if we understand what a linear low-dose model is. Parsing the sentence, it seems to say “uncertainty over a…slope…may be quantified, but uncertainty…in slope…cannot be captured” which is self-contradictory. Perhaps what the authors mean to say is that the linearity assumption is not itself subject to uncertainty quantification.

Page 6-9, line 1: The mathematical symbol \( x \) should be italicized, as should all Roman letters throughout the document that represent unknown quantities, i.e., are symbols representing something else rather than names like “e” the base of the natural logarithms.

Lines 14 and 16: The prefixes “pseudo” and “quasi” are not words. Hyphens are needed.

Page 6-9, line 18: Provide citations for dependence modeling.

Page 6-9, Section 6.1.3.6 might also mention graphs, and other traditional communication tools other than correlation indices.

Page 6-10, line 4: Add mention of methods that identify uncertain assumptions or parameters that are important for determining whether the model is consistent with observed data (Hornberger and Spear, 1983) and for affecting a decision that is made as a result of the model (Merz et al., 2009).

Page 6-10, lines 29-30: Do the authors mean “this probabilistic language,” referring to the word “likely” in the quoted text?

Page 6-11, line 19: Of course there is no guarantee that linear will be protective.

Page 6-13, line 18: Of course it isn’t really apodictic knowledge at all, but rather only an opinion or an assumption. We see the authors’ point and agree with it entirely, but perhaps they should use a word other than “apodictic” here since it’s not technically correct.

Page 6-14, lines 33-34: The parenthetical phrase “volitional uncertainty” should be expanded into a sentence that says what the authors mean to express. The phrase “cognitive uncertainty” does not mean anything in this context. Perhaps if the authors expanded it into a sentence too, maybe making it “epistemic uncertainty” along the way, it would be possible to understand what they are trying to say here.

Footnote 62: “Effective” is misspelled, as is “cancer.”
Page 6-16, line 5: We note that it’s not really a guarantee of course.

Line 8: The word “common” should be “predominant”.

Page 6-16, line 20: Perhaps we can say that variability (and uncertainty) in the factors that are used to determine a particular UF can be considered in choosing the particular value of the UF.

Page 6-17, lines 3-14: This problem can be addressed using a Bayesian analysis with a beta conjugate for the uncertain response probability, $p$, with uninformative (uniform) prior for $p$. The probability that “an experiment with a null response might have yielded a positive response” can be estimated from the predictive distribution (which will depend on the number of test animals in the original study that yielded zero responses) for the next experiment (with any number of exposed animals).

Page 6-17, line 28: The word “band” should be “limit”.

Page 6-20, footnote 66: The text starting “each have an error factor” should be followed by “of” rather than “or”.

Page 6-21, line 6: It would be helpful to say something about what the concerns are.

Page 6-21, lines 12-14: NAS was not suggesting that EPA use the uncertainty factors approach to mount an uncertainty analysis, but rather a more modern approach.

Page 6-22, line 19: Would it be appropriate to note “and establishes a concomitant reduction in some UFs?”

Line 29: The word “invokes” should perhaps be “would require”.

Page 6-23, line 33 and passim: The word “exotic” is a poor choice that is unnecessarily and transparently loaded.

Page 6-25, line 29: This sentence is ungrammatical.

Page 6-26, line 24 and Figure 6-1: Would it be helpful to draw the 45-degree line on the graph?

Page 6-27, line 10: The word “epistemic” here is acceptable.

Line 14: The word “epistemic” here should be replaced by “fixed across individuals,” and “is estimated from” should be replace by “varies with.” How does half life’s estimability from data imply that it is variable?

Page 6-28, lines 1-2: One would need the dependence between the variables to proceed.
Line 9: We suggest that “and” should be “although.”

Page 6-29, line 1-2: There are bounding techniques based on the classical Fréchet inequality that do not require any knowledge of or any assumptions about dependencies.

Line 32: Omit “to.”

Page 6-31, line 24: The pessimistic conclusion is a bit strong. Any estimate made from data is amenable to a quantitative uncertainty analysis so, if one is measuring anything, one can propagate uncertainties such as mensurational uncertainty, sampling uncertainty, and perhaps even surrogacy uncertainty. It’s not quite as hard to get quantitative models as the text here seems to suggest.

Page 6-32, lines 13-14: The dour conclusion is confusing. One could do a sensitivity analysis in this case, couldn’t one? If so, it seems that some kind of uncertainty analysis is clearly possible.

Page 6-33: The example in the text box is great, but the second table seems to say the log-likelihood for LLD is 2.46 and for Hill is 2.16, which would make LLD’s larger than Hill’s, which contradicts what’s said in the text.

Page 6-34, line 4: Shouldn’t “Delivered dose” be a new bullet?

Line 8: We don’t think this statement is true. Perhaps “statistically more powerful” should be “typically yield more sensitive”.

Lines 24-25: We don’t think it is necessary or helpful to persist with Box’s platitude. Model uncertainty is the uncertainty about a model’s predictions that arises from doubt about the relevance of that model for making such predictions.

Page 6-37, line 29: The caveat is overwrought. Exploring relevant alternative values in a sensitivity analysis could constitute a quantitative uncertainty analysis, even if the exploration is limited.

Page 6-37, line 30: This sentence is false. Analytical methods of propagation (convolution) don’t “sample” anything, and analyses based on intervals or imprecise probabilities don’t depend on uncertainty “distributions” (i.e., precise probability distributions).

It is important to keep in mind that, in general, we are not necessarily limited to identifying precise probability distributions for everything that is to be characterized as uncertain (as seems to be suggested here). Simple intervals about uncertain quantities can support a straightforward, albeit crude, interval analysis that propagates uncertainty about parameters and other model choices to statements about the range of possible
results. Similarly, an approach based on interval probabilities, probability boxes, or
general imprecise probabilities (Walley, 1991) can combine such intervals with precise
distributions if they are known for some other inputs, and with structures that are
intermediate between coarse intervals and delicate probability distributions when some
but incomplete knowledge is available. If the inputs are profoundly uncertain, the results
from such analyses are likely to be wide in reflection of these uncertainties. In almost all
cases, it is possible to be entirely rigorous without necessarily being precise and without
completely specifying each probability distribution.

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

Page 6-38, line 30 and passim: The adjective “data driven” needs a hyphen, as it has
elsewhere in the document.

Line 23-24: We think this sentence is true, but, again, sampling from a distribution is not
the only way to conduct a quantitative uncertainty analysis.

Line 26: What is “(2.a)?”

Page 6-41, line 23: Omitting the word “extra” would make the sentence more easily
understandable.

Line 31: What does “How Forward?” mean? Is this idiomatic?

Section 6.5.2: The assertions in this section are rather surprising and questionable. EPA
says that uncertainty quantification is an “emerging area in science” and that it is “an area
where research could be focused” because “the requisite knowledge does not yet exist” to
apply quantitative uncertainty analysis in assessments such as this one for dioxin. The
document peremptorily dismisses the utility of “convening a blue-ribbon panel” to
identify the proper approach and suggests instead that “multiple approaches should be
encouraged.” Is the inference that the present review panel shouldn’t try to say what the
proper approaches to uncertainty quantification are, even if we think the area is more
mature than emerging? It is hard to understand what these statements are suggesting.
Will the Agency support intramural and extramural research efforts in this direction? If
not, what do the statements mean? Is it impossible that EPA could benefit from some
tech transfer efforts as well as basic research on uncertainty quantification? The
paragraph beginning on page 6-42 (line 3) mentions a European idea of bench-test
exercises to compare different approaches. It may be worth mentioning that this idea has
been implemented in the United States as well (Oberkampf et al., 2004; Ferson et al.,
2004).

The document’s reference list is alphabetically arranged, but seems to go from Z back to
A again on page R-33.
1 Appendix D: EPA’s Charge Questions

MEMORANDUM


FROM: Becki Clark, Deputy Director
National Center for Environmental Assessment (8601P)
Office of Research and Development

TO: Vanessa T. Vu, Ph.D., Director
EPA Science Advisory Board (1400F)

This is to request a review by the Science Advisory Board of the draft report entitled “EPA’s Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments” (EPA/600/R-10/038A). This draft report details the Environmental Protection Agency’s (EPA) response to key comments and recommendations included in the 2006 NAS report ("Health Risks from Dioxin and Related Compounds: Evaluation of the EPA Reassessment") on their review of the (EPA) 2003 draft dioxin reassessment. This draft report also includes significant new analyses on both the potential cancer and noncancer human health effects that may result from chronic exposures to dioxins.

Attached is the Charge that provides background information as well as questions that are to be the focus of the Science Advisory Board review of this draft report.

Please let me know if you have any questions. Thank you.


cc: Peter W. Preuss
Annette Gatchett
Glenn Rice
Cheryl Itkin
Proposed Charge to the Science Advisory Board for Peer Review Of Draft Report

“EPA’s Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments”

May, 2010

EPA has been preparing an assessment of the potential health impacts of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) for many years. In 2003, EPA released an external review draft report entitled, Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and Related Compounds (U.S. EPA, 2003) (herein referred to as “2003 Reassessment”) that was reviewed by the EPA Science Advisory Board (SAB), and then by the National Academy of Sciences (NAS). In 2006, the National Research Council (NRC) of the National Academies published their report of EPA’s reassessment, Health Risks from Dioxin and Related Compounds: Evaluation of the EPA Reassessment (NRC, 2006).

The current Report EPA’s Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments (“Response to Comments”) before the SAB is a response to the review by the NRC, and includes new analyses completed in response to the NRC recommendations and recently published literature, as well as a discussion of topics where our views differed. The draft Response to Comments document is not an assessment per se; it is designed to supplement the information provided in the 2003 Reassessment. However, the draft Response to Comments provides a noncancer reference dose and updated cancer values. Detailed discussions of many of the issues addressed in the draft Response to Comments are available in the 2003 Reassessment and have not been reproduced in the current Report – whenever appropriate; the reader is directed to the pertinent chapters of the 2003 Reassessment.

The NRC identified three key recommendations that they believed would result in substantial improvement to the EPA 2003 Reassessment and thus support a scientifically robust characterization of human responses to exposures to TCDD. These three key areas are (1) improved transparency and clarity in the selection of key data sets for dose-response analysis, (2) further justification of approaches to dose-response modeling for cancer and noncancer endpoints, and (3) improved transparency, thoroughness, and clarity in quantitative uncertainty analysis. The NRC Report also encouraged EPA to calculate a reference dose (RfD), which had not been derived in the 2003 Reassessment.

The draft Response to Comments document addresses each of these issues. Please consider the accuracy, objectivity, and transparency of EPA’s reanalysis and responses in your review.

General Charge Questions

1.1 Is the draft Response to Comments clear and logical? Has EPA objectively and clearly presented the three key NRC recommendations?
1.2 Are there other critical studies that would make a significant impact on the conclusions of the hazard characterization or dose-response assessment of the chronic noncancer and cancer health effects of TCDD?

Specific Charge Questions

Section 2. Transparency and Clarity in the Selection of Key Data Sets for Dose-Response Analysis

2.1. Is this Section responsive to the NAS concern about transparency and clarity in data-set selection for dose-response analysis?

2.2. Are the epidemiology and animal bioassay study criteria/considerations scientifically justified and clearly described?

2.3. Has EPA applied the epidemiology and animal bioassay study criteria/considerations in a scientifically sound manner? If not, please identify and provide a rationale for alternative approaches.

Section 3. The Use of Toxicokinetics in Dose-Response Modeling for Cancer and Noncancer Endpoints

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

Please comment on:

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

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

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

D-3
3.1.d. Whether EPA adequately characterized the uncertainty in the kinetic models.

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

Please comment on:

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

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

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.

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

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

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

Section 4. Reference Dose

4.1. The Mocarelli et al. (2008) and Baccarelli et al. (2008) studies were selected as co-critical studies for the derivation of the RfD. Is the rationale for this selection scientifically justified and clearly described? Please identify and provide the rationale for any other studies that should be selected, including the rationale for why the study would be considered a superior candidate for the derivation of the RfD. In addition, male reproductive effects and changes in neonatal thyroid hormone levels, respectively, were selected as the co-critical effects for the RfD. Please comment on whether the selection of these critical effects is scientifically justified and clearly described. Please identify and provide the rationale for any other endpoints that should be selected as the critical effect.

4.2. In the Seveso cohort, the pattern of exposure to TCDD is different from the average
daily exposure experienced by the general population. The explosion in Seveso
created a high dose pulse of TCDD followed by low level background dietary
exposure in the exposed population. In the population, this high dose pulse of
TCDD was slowly eliminated from body tissues over time. There is uncertainty
regarding the influence of the high-dose pulse exposure on the effects observed
later in life.

4.2.a. Mocarelli et al. (2008), reported male reproductive effects observed later
in life for boys exposed to the high dose pulse of TCDD between the ages
of 1 and 10. EPA identified a 10 year critical exposure window. In the
development of the candidate RfD, EPA used an exposure averaging
approach that differs from the typical approach utilized for animal
bioassays. EPA determined that the relevant exposure should be
calculated as the mean of the pulse exposure and the 10-year critical
exposure window average. Please comment on the following:

4.2.a.i. EPA’s approach for identifying the exposure window and
calculating average exposure for this study.

4.2.a.ii. EPA’s designation of a 20% decrease in sperm count (and an
11% decrease in sperm motility) as a LOAEL for Mocarelli et al.
(2008).

4.2.b. For Baccarelli et al. (2008), the critical exposure window occurs long
after the high-dose pulse exposure. Therefore, the variability in the
exposure over the critical exposure window is likely to be less than the
variability in the Mocarelli et al. subjects. EPA concluded that the
reported maternal exposures from the regression model developed by
Baccarelli et al. provide an appropriate estimate of the relevant effective
dose as opposed to extrapolating from the measured infant TCDD
concentrations to maternal exposure. Additionally, EPA selected a
LOAEL of 5 µ-units TSH per ml blood in neonates; as this was
established by World Health Organization (WHO) as a level above which
there was concern about abnormal thyroid development later in life.
Please comment on the following:

4.2.b.i. EPA’s decision to use the reported maternal levels and the
appropriateness of this exposure estimate for the Baccarelli et al.
study.

4.2.b.ii. EPA’s designation of 5 µ-units TSH per ml blood as a LOAEL
for Baccarelli et al. (2008).

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

4.4. EPA did not consider biochemical endpoints (such as CYP induction, oxidative stress, etc.) as potential critical effects for derivation of the RfD for TCDD due to the uncertainties in the qualitative determination of adversity associated with such endpoints and quantitative determination of appropriate response levels for these types of endpoints in relation to TCDD exposure. Please comment on whether this decision is scientifically justified and clearly described.

4.5. In using the animal bioassays, EPA averaged internal blood TCDD concentrations over the entire dosing period, including 24 hours following the last exposure. Please comment on EPA’s approach for averaging exposures including intermittent and one day gestation exposure protocols.

4.6. Please comment on the benchmark dose (BMD) modeling conducted by EPA to analyze the animal bioassay data and EPA’s choice of points of departure (PODs) from these studies.

4.7. For the animal bioassay modeling, EPA applied the kinetic extrapolation at the level of the POD prior to applying the uncertainty factors because EPA has less confidence in the kinetic model output at lower doses reflective of the RfD. Please comment on whether this approach was scientifically justified and clearly described.

4.8. Please comment as to whether EPA’s qualitative discussion of uncertainty in the RfD is justified and clearly described.

Section 5. Cancer Assessment

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

5.2. Mode of Action: The mode of action of a carcinogen can inform identification of hazards and approaches used for a dose-response assessment. The mode of carcinogenic action for TCDD has not been elucidated for any tumor type. EPA concluded that, while interaction with the Ah receptor is likely to be a necessary early event in TCDD carcinogenicity in experimental animals, the downstream events involved are unknown.

5.2.a. Are the available data related to mode(s) of action for the carcinogenicity of TCDD appropriately characterized and clearly presented?
5.2.b. Do the available data support EPA’s conclusion that the overall mode(s) of action for TCDD-induced carcinogenesis is largely unknown? Please comment on whether this evaluation is clearly described.

5.3. Is EPA’s approach for selecting data sets from the key epidemiologic studies and animal bioassays identified for cancer dose response modeling scientifically justified and clearly described?

5.4. For the animal bioassay data, potential cancer oral slope factors (OSFs) were calculated by linear extrapolation (using a linear, nonthreshold cancer approach) from the point of departure (POD). EPA also estimated the composite risk of the occurrence of several tumor types from the animal cancer bioassay data.

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

5.4.b. Please comment on the choice of using a BMDL_{01} as the POD for the development of candidate oral slope factors derived from the TCDD animal cancer bioassays.

5.5. 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 quantitatively cancer assessment.

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.

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.

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 risk itself at a single risk level (BMR) for extrapolation to the origin. Please comment on EPA’s choice of extrapolation approach.
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.

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

5.7. 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.

5.8. The NRC suggested that EPA consider nonlinear approaches for the assessment of TCDD carcinogenicity. In the Response to Comments, EPA presents two illustrative nonlinear approaches for cancer, but considers both inappropriate to use because of the lack of MOA information.

5.8.a. Please comment on these two illustrative nonlinear approaches including EPA’s conclusions regarding the limitations of these approaches.

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

Section 6. Feasibility of Quantitative Uncertainty Analysis from NAS Evaluation of the 2003 Reassessment

6.1. Please comment on the discussion in this Section. Is the response clearly presented and scientifically justified?

6.2. Please comment on EPA’s overall conclusion that a comprehensive quantitative uncertainty analysis is not feasible.

6.2.a. Please comment on the discussion in Section 6 regarding volitional uncertainty and how this type of uncertainty limits the ability to conduct a quantitative uncertainty analysis.

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