NCEA Proposed Draft Charge to the Science Advisory Board for the IRIS Toxicological Review of Benzo[a]pyrene

August 2013

Introduction

The U.S. Environmental Protection Agency (EPA) is seeking an external peer review of the scientific basis supporting the draft Toxicological Review of Benzo[a]pyrene that will appear on the Agency's online database, the Integrated Risk Information System (IRIS). IRIS is prepared and maintained by EPA's National Center for Environmental Assessment (NCEA) within the Office of Research and Development (ORD). An existing assessment for benzo[a]pyrene, which includes an oral slope factor (OSF) and a cancer weight of evidence descriptor was posted on IRIS in 1987.

IRIS is a human health assessment program that evaluates scientific information on effects that may result from exposure to specific chemical substances in the environment. Through IRIS, EPA provides high quality science-based human health assessments to support the Agency's regulatory activities and decisions to protect public health. IRIS assessments contain information for chemical substances that can be used to support the first two steps (hazard identification and dose-response assessment) of the human health risk assessment process. When supported by available data, IRIS provides health effects information and toxicity values for chronic health effects (including cancer and effects other than cancer). Government and others combine IRIS toxicity values with exposure information to characterize public health risks of chemical substances; this information is then used to support risk management decisions designed to protect public health.

The external review draft Toxicological Review of Benzo[a]pyrene is based on a comprehensive review of the available scientific literature on the noncancer and cancer health effects in humans and experimental animals exposed to benzo[a]pyrene. This draft IRIS assessment includes:

- a *Preamble* to describe the methods used to develop IRIS assessments;
- an *Executive Summary* to concisely summarize the major conclusions of the assessment;
- a *Literature Search Strategy/Study Selection* section to describe the process for identifying and evaluating the evidence for consideration in developing the assessment;
- a *Hazard Identification* chapter to systematically synthesize and integrate the available evidence of organ/system-specific hazards; and
- a *Dose-Response Analysis* chapter to describe the selection of studies for consideration in calculating toxicity values and to describe the analysis and methodology in deriving and selecting toxicity values.

Additionally, appendices for chemical and physical properties, toxicokinetic information, summaries of toxicity studies, and other supporting materials are provided as *Supplemental Information* (see Appendices A to E) to the draft Toxicological Review. The draft assessment was developed according to guidelines and technical reports published by EPA (see Preamble), and contains a qualitative characterization of the hazards for benzo[a]pyrene, including a cancer descriptor of the chemical's human carcinogenic potential, cancer risk estimates for oral, inhalation, and dermal exposure and noncancer toxicity values for chronic oral (reference dose, RfD) and inhalation (reference concentration, RfC) exposure.

Charge Questions

In April 2011, the National Research Council (NRC) released its *Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde*. In addition to offering comments specifically about EPA's draft formaldehyde assessment, the NRC included comments and recommendations for improving the development of IRIS assessments. The IRIS Program's implementation of the NRC recommendations is following a phased approach. Phase 1 of implementation has focused on a subset of the short-term recommendations, such as editing and streamlining documents, increasing transparency and clarity, and using more tables, figures, and appendices to present information and data in assessments. Phase 1 also focused on assessments that had been near the end of the development process and close to final posting. The IRIS Program is now in Phase 2 of implementation, which addresses all of the short-term NRC recommendations. The Program is implementing all of these recommendations but recognizes that achieving full and robust implementation of certain recommendations will be an evolving process with input and feedback from the public, stakeholders, and external peer review committees. This phased approach is consistent with the NRC's *Roadmap for Revision* as described in Chapter 7 of the formaldehyde review report. The NRC stated that "the committee recognizes that the changes suggested would involve a multi-year process and extensive effort by the staff at the National Center for Environmental Assessment and input and review by the EPA Science Advisory Board and others."

Below is a set of charge questions that address scientific issues in the draft IRIS Toxicological Review of Benzo[a]pyrene. The charge questions also seek feedback on whether the document is clear and concise, a central concern expressed in the NRC report. Please provide detailed explanations for responses to the charge questions. EPA will also consider the Science Advisory Board review panel's comments on other major scientific issues specific to the hazard identification and dose-response assessment of benzo[a]pyrene. Please consider the accuracy, objectivity, and transparency of EPA's analyses and conclusions in your review.

General Charge Questions:

1. NRC (2011) indicated that the introductory section of IRIS assessments needed to be expanded to describe more fully the methods of the assessment. NRC stated that they were "not recommending the addition of long descriptions of EPA guidelines to the introduction, but rather clear, concise statements of criteria used to exclude, include, and advance studies for derivation of [toxicity values]." Please comment on whether the new *Preamble* provides a clear and concise description of the guidance and methods that EPA uses in developing IRIS assessments.

2. NRC (2011) provided comments on ways to improve the presentation of steps used to generate IRIS assessments and indicated key outcomes at each step, including systematic review of evidence, hazard identification, and dose-response assessment. Please comment on the new IRIS document structure and whether it will increase the ability for the assessments to be more clear, concise, and easy to follow.

3. NRC (2011) states that "all critical studies need to be thoroughly evaluated with standardized approaches that are clearly formulated" and that "strengthened, more integrative, and more transparent discussions of weight of evidence are needed." NRC also indicated that the changes suggested would involve a multiyear process. Please comment on EPA's success thus far in

implementing these recommendations.

Chemical-Specific Charge Questions:

A. Executive Summary

1. The major conclusions of the assessment pertaining to the hazard identification and dose-response analysis have been summarized in the *Executive Summary*. Please comment on whether the conclusions have been clearly and sufficiently described for purposes of condensing the Toxicological Review information into a concise summary.

B. Literature Search Strategy/Study Selection

1. The process for identifying and selecting pertinent studies for consideration in developing the assessment is detailed in the *Literature Search Strategy/Study Selection* section. Please comment on whether the literature search approach, screening, evaluation, and selection of studies for inclusion in the assessment are clearly described and supported. Please identify any additional peer-reviewed studies from the primary literature that should be considered in the assessment of noncancer and cancer health effects of benzo[a]pyrene.

C. Hazard Identification

Synthesis of Evidence

1. A synthesis of the evidence for benzo[a]pyrene toxicity is provided in Chapter 1, *Hazard Identification*. Please comment on whether the available data have been clearly and appropriately synthesized for each toxicological effect. Please comment on whether the weight of evidence for hazard identification has been clearly described and scientifically justified.

Summary and Evaluation

2. Does EPA's hazard assessment of noncancer human health effects of benzo[a]pyrene clearly integrate the available scientific evidence (i.e., human, experimental animal, and mechanistic evidence) to support the conclusion that benzo[a]pyrene poses a potential hazard to the developing fetus; the nervous system in the developing fetus; the male and female reproductive systems; and the immune system?

3. Does EPA's hazard assessment of the carcinogenicity of benzo[a]pyrene clearly integrate the available scientific evidence to support the conclusion that under EPA's *Guidelines for Carcinogen Risk Assessment* (U.S. EPA, 2005), benzo[a]pyrene is "carcinogenic to humans" by all routes of exposure?

4. Does EPA's hazard assessment of the mode of action for carcinogenicity of benzo[a]pyrene clearly integrate the available scientific evidence to support the conclusion that a mutagenic mode of action is the primary mode of action of benzo[a]pyrene-induced carcinogenicity?

D. Dose-Response Analysis

Oral Reference Dose (RfD)

Several hazards were identified for oral exposure to benzo[a]pyrene. Studies and effects within each hazard (i.e., developmental, reproductive, and immunotoxicity) were evaluated and the most relevant, informative studies and effects were selected for dose-response analysis, where data were amenable, for consideration in deriving an RfD.

1. Please comment on whether the evaluation and selection of studies and effects for the derivation of candidate values to consider for the RfD are scientifically supported and clearly described. Specifically, please comment on the selection of the following studies and effects for dose-response analysis. Please identify and provide the rationale for any other studies or effects that should be considered.

    a. Developmental toxicity: Chen et al. (2012) [neurodevelopmental changes]; Jules et al. (2012) [cardiovascular effects]
    b. Reproductive toxicity: Xu et al. (2010) [decreased ovary weights]; Zheng et al. (2010) [decreased intratesticular testosterone]; Mohamed et al. (2010) [decreased sperm count and motility]; and Gao et al. (2011) [increased cervical hyperplasia]
    c. Immunotoxicity: Kroese et al. (2001) [decreased thymus weights]; DeJong et al. (1999) [decreased serum IgM and IgA levels and number of B cells]

2. Benchmark dose (BMD) modeling was applied to derive points of departure (POD) for the candidate values when possible. Has the BMD modeling been appropriately conducted? Are the benchmark responses (BMR) selected for use in deriving the PODs scientifically supported and clearly described? When BMD modeling was not possible a NOAEL/LOAEL approach was used to calculate candidate values. Please comment on whether these approaches are scientifically supported and clearly described.

3. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the PODs for the derivation of the candidate values. Are the UFs appropriate, based on the recommendations described in Section 4.4.5 of *A Review of the Reference Dose and Reference Concentration Processes* (U.S. EPA, 2002), and clearly described? If changes to the selected UFs are proposed, please identify and provide scientific support for the proposed changes.

4. From the candidate values, an organ/system-specific reference value was selected for each hazard (developmental, reproductive, and immunotoxicity). EPA concluded that these values best represented the hazards based on considerations of weight of evidence, uncertainty, and sensitivity. Please comment on whether the selection of the organ/system-specific reference value is scientifically supported, appropriate for development of a chronic RfD, and is clearly described. Please identify and provide the rationale for any other values that should be considered.

5. The proposed overall RfD was based on neurodevelopmental changes observed by Chen et al. (2012). This value was selected based on the confidence in and sensitivity of the reference value. Please comment on whether the selection of this RfD is scientifically supported and clearly described. Please identify and provide the rationale for any other values that should be considered.

Inhalation Reference Concentration (RfC)

Several hazards were identified for inhalation exposure to benzo[a]pyrene. Studies and effects within each hazard (i.e., developmental and reproductive) were evaluated and the most relevant, informative studies were selected for dose-response analysis, where data were amenable, for consideration in deriving an RfC.

6. Please comment on whether the evaluation and selection of studies and effects for the derivation of candidate values to consider for the RfC are scientifically supported and clearly described. Specifically, please comment on the selection of the following studies and effects for dose-response analysis. Please identify and provide the rationale for any other studies or effects that should be considered.

    a. Developmental toxicity: Archibong et al. (2002) [decreased fetal survival]
    b. Reproductive toxicity: Archibong et al. (2008) [decreased testes weight and decreased sperm count and motility]

7. The NOAEL/LOAEL approach was used to derive the PODs for the candidate values. Please comment on whether this approach is scientifically supported and clearly described.

8. Please comment on the rationale for the selection of the uncertainty factors (UFs) applied to the PODs for the derivation of the candidate values. Are the UFs appropriate, based on the recommendations described in Section 4.4.5 of *A Review of the Reference Dose and Reference Concentration Processes* (U.S. EPA, 2002), and clearly described? If changes to the selected UFs are proposed, please identify and provide scientific support for the proposed changes.

9. From the candidate values, an organ/system-specific reference value was selected for the developmental hazard. A reference value for the reproductive hazard was not selected due to significant uncertainty in deriving the reproductive candidate values. EPA concluded that the developmental value best represented the hazard considering the effect on fetal survival is the most sensitive noncancer developmental effect observed following inhalation exposure to benzo[a]pyrene. Please comment on whether the selection of the organ/system-specific reference value is scientifically supported, appropriate for development of a chronic RfC, and is clearly described. Please identify and provide the rationale for any other values that should be considered.

10. The proposed overall RfC was based on decreased fetal survival observed by Archibong et al., (2002). Please comment on whether the selection of this RfC is scientifically supported and clearly described. Please identify and provide the rationale for any other values that should be considered.

Cancer Risk Estimates

Oral Slope Factor (OSF)

Carcinogenicity studies examining oral exposure to benzo[a]pyrene were evaluated and the most relevant, informative studies and endpoints were selected for dose-response analysis, where data were amenable, for consideration in deriving an OSF.

11. The Kroese et al. (2001) and Beland and Culp (1998) studies were selected as the best available studies for dose-response analysis. The incidence data for forestomach and oral cavity, liver, jejunum/duodenum, kidney, and skin tumors in male and female rats reported by Kroese et al. (2001) and forestomach, esophagus, tongue, larynx tumors (alimentary tract) in female mice reported by Beland and Culp (1998) were selected for dose-response analysis. Please comment on whether the evaluation, selection, and relevance of studies and endpoints for dose-response analysis is scientifically supported and clearly described. Please identify and provide the rationale for any other studies or endpoints that should be considered.

12. BMD modeling was conducted using the incidence of the individual tumor types reported in Kroese et al. (2001) and Beland and Culp (1998) in conjunction with dosimetric adjustments for calculating the human equivalent doses to estimate the PODs. The candidate OSFs were calculated by linear extrapolation from the PODs (i.e., the lower 95% confidence limit on the dose associated with 10% extra risk). Please comment on whether this approach is scientifically supported and clearly described.

13. The OSF associated with alimentary tract tumors in female mice as reported by Beland and Culp (1998) was selected as the recommended slope factor for assessing human cancer risk following oral exposure to benzo[a]pyrene. Please comment on whether this selection is scientifically supported and clearly described. Please identify and provide the rationale for any other studies or endpoints that should be selected to serve as the basis for the OSF.

Inhalation Unit Risk (IUR)

The benzo[a]pyrene inhalation database for carcinogenicity consists of a lifetime inhalation bioassay and several intratracheal instillation studies. The instillation studies were not considered for dose-response analysis because use of this exposure method alters the deposition, clearance, and retention of substances, and therefore, is less relevant and informative for the quantitative estimation of inhalation cancer risk compared with inhalation bioassays.

14. The Thyssen et al. (1981) study was selected as the best available study for dose-response analysis as it represents the only lifetime inhalation cancer bioassay available for describing exposure-response relationships for cancer from inhaled benzo[a]pyrene. The incidence data for tumors of the upper respiratory and digestive tracts (pharynx, larynx, trachea, esophagus, nasal cavity, and forestomach) reported by Thyssen et al. (1981) were selected for dose-response analysis. Please comment on whether the evaluation, selection, and relevance of studies and endpoints for dose-response analysis is scientifically supported and clearly described. Please identify and provide the rationale for any other studies or endpoints that should be considered.

15. BMD modeling was conducted using the overall incidence of the tumors of the upper

respiratory and digestive tracts reported by Thyssen et al. (1981) to estimate the PODs. Dosimetric adjustments for calculating the human equivalent concentrations were not conducted due to the lack of data to inform a basis for extrapolation to humans. It was assumed that equal risk for all species would be associated with equal concentrations in air; thus, the continuous time-weighted group average concentrations in male hamsters were used for the dose-response analysis under the assumption that these are representative across species. The candidate IURs were calculated by linear extrapolation from the PODs (i.e., the lower 95% confidence limit on the concentration associated with 10% extra risk of tumors of the upper respiratory and digestive tracts). Please comment on whether this approach is scientifically supported and clearly described.

16. The IUR associated with tumors of the upper respiratory and digestive tracts in male hamsters (in which the tumors were considered incidental to the death of an animal) as reported by Thyssen et al. (1981) was selected as the recommended unit risk for assessing human cancer risk following inhalation exposure to benzo[a]pyrene. Please comment on whether this selection is scientifically supported and clearly described. Please identify and provide the rationale for any other studies or endpoints that should be selected to serve as the basis for the IUR.

Dermal Slope Factor (DSF)

Carcinogenicity studies examining dermal exposure to benzo[a]pyrene were evaluated and the most relevant, informative studies and endpoints were selected for dose-response analysis, where data were amenable, for consideration in deriving a DSF.

17. The Roe et al. (1970), Sivak et al. (1997), and Poel (1959) studies were selected as the best available studies for dose-response analysis. Several other studies provided supportive information but were considered less informative due to incomplete exposure duration information or greater uncertainty associated with extrapolating to lower doses. These studies were included in the dose-response analysis to help characterize similarities among the studies on a quantitative basis. The incidence data for skin tumors in male and female mice were selected for dose-response analysis. Please comment on whether the evaluation, selection, and relevance of studies and endpoints for dose-response analysis is scientifically supported and clearly described. Please identify and provide the rationale for any other studies and endpoints that should be considered.

18. BMD modeling was conducted using the incidence of skin tumors reported in the chronic mouse bioassays to estimate the PODs. The candidate DSFs were calculated by linear extrapolation from the PODs (i.e., the lower 95% confidence limit on the concentration associated with 10% extra risk of skin tumors). Please comment on whether this approach is scientifically supported and clearly described.

19. Among the three studies considered the most relevant and informative for the DSF, the male mouse data (reported by Sivak et al., 1997 and Poel, 1959) were more sensitive than the female mouse data (reported by Roe et al., 1970). Therefore, the DSF associated with skin tumors in male mice was calculated by linear extrapolation from the average of the PODs from the Sivak et al. (1997) and Poel (1959) studies. The resulting DSF was selected as the recommended slope factor for assessing human cancer risk following dermal exposure to benzo[a]pyrene. Please comment on whether this selection is scientifically supported and clearly described. Please

identify and provide the rationale for any other studies or endpoints that should be selected to serve as the basis for the DSF.

20. The DSF was adjusted to account for interspecies scaling between mice and humans. This cross-species adjustment was based on allometric scaling using body weight to the 3/4 power. Under this approach, rodents and humans exposed to the same daily dose of a carcinogen, adjusted for $BW^{3/4}$, would be expected to have equal lifetime risks of cancer. However, because there is no established methodology for cross-species extrapolation of dermal toxicity, several alternative approaches were evaluated (see Appendix E of the Toxicological Review). Please comment on whether the selected interspecies scaling approach is scientifically supported and clearly described. Also, please comment on whether the alternative approaches presented are clearly described and whether any of these approaches should be selected as the recommended approach. Please identify and provide the rationale for any alternative approach that should be selected.