EPA's Integrated Risk Information System (IRIS) Program

Progress Report and Report to Congress

U.S. Environmental Protection Agency: Office of Research and Development

February, 2015

Chapter 1 – Introduction

The U.S. Environmental Protection Agency's (EPA) Integrated Risk Information System (IRIS) Program develops human health assessments that provide health effects information on environmental chemicals to which the public may be exposed, providing a critical part of the scientific foundation for EPA's decisions to protect public health. In April 2011, the National Research Council (NRC), in their report *Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde*[1], made several recommendations to EPA for improving IRIS assessments and the IRIS Program. The NRC's recommendations were focused on the first step of the IRIS process, the development of draft assessments. Consistent with the advice of the NRC and as the Agency has reported previously, the IRIS Program is implementing these recommendations using a phased approach and is making the most extensive changes to assessments that are currently in the earlier stages of the IRIS process.

In April 2012, EPA delivered a report to Congress outlining EPA's progress toward implementing the NRC's recommendations. The 2014 Consolidated Appropriations Act further directs EPA to: "provide to the House and Senate Committees on Appropriations a progress report that describes the Agency's implementation of NAS Chapter 7 recommendations for fiscal years 2012 and 2013." Appendix A provides the exact language from the 2014 Consolidated Appropriations Act related to the IRIS Program. The purpose of this report is to update Congress, stakeholders, and the public on the status of the IRIS Program's implementation of the NRC recommendations since delivering the 2012 report to Congress. This report also provides specific examples demonstrating the application of scientific methods in IRIS assessments consistent with the NRC's 2011 recommendations.

Additionally, the 2014 Consolidated Appropriations Act directs EPA to "include a chapter on whether there are more appropriate scientific methods to assess, synthesize, and draw conclusions regarding likely human health effects associated with likely exposures to substances. The Agency also should discuss the current reevaluation of the formaldehyde and acrylonitrile assessments as well as any other assessments that may be relevant as case studies." The scientific methods the IRIS Program is using are described throughout this report, and examples are provided from several different assessments under development. In addition, updated information specific to the formaldehyde and acrylonitrile assessments is provided in Chapter 5.

Background on IRIS

IRIS human health assessments contain information that can be used to support the first two steps (hazard identification and dose-response analysis) of the risk assessment paradigm. IRIS assessments are scientific reports that provide information on a chemical's hazards and, when supported by available data, quantitative toxicity values for cancer and noncancer health effects. IRIS assessments are not regulations, but they provide a critical part of the scientific foundation for decisions to protect public health across EPA's programs and regions under an array of environmental laws (e.g., Clean Air Act, Safe Drinking Water Act, Comprehensive Environmental Response, Compensation, and Liability Act).

[1] National Research Council, 2011. Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde.

EPA's program and regional offices combine information from IRIS assessments with relevant exposure information for a chemical to assess the public health risks of environmental contaminants. EPA decision-makers use these risk assessments, along with other considerations (e.g., statutory and legal requirements that can include cost-benefit information, technological feasibility, and economic factors) to inform risk management decisions, which can include proposed regulations to protect public health. IRIS assessments also are a resource for risk assessors and environmental and health professionals from state and local governments and other countries. Figure 1 illustrates where IRIS assessments contribute information within the risk assessment and risk management paradigms.

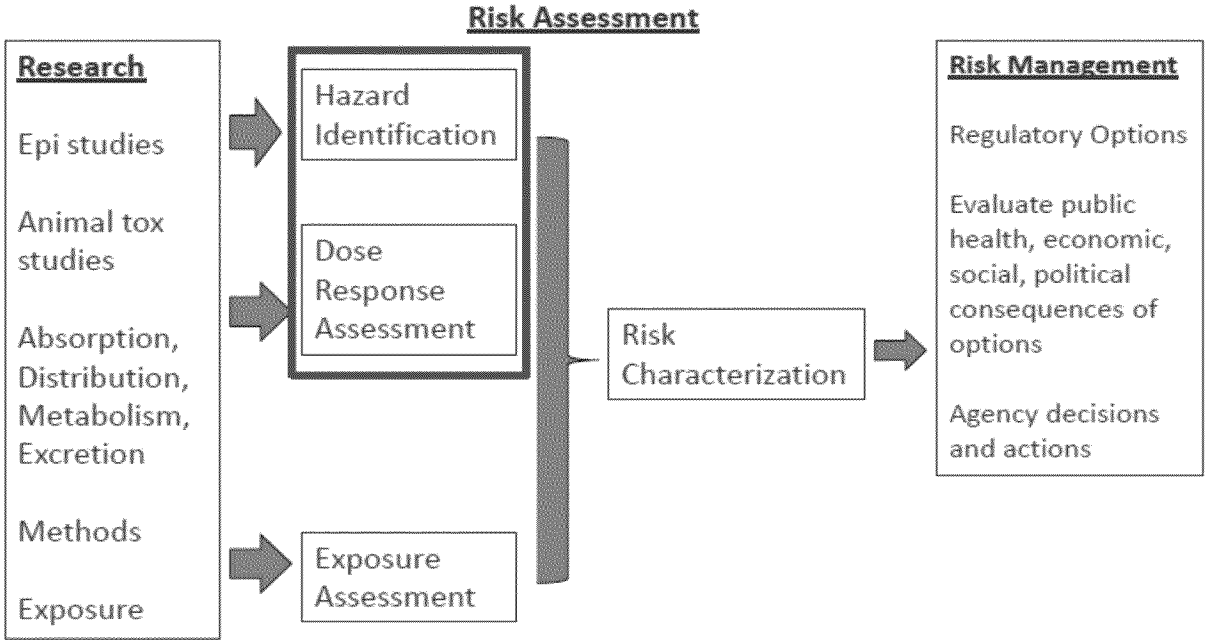

Figure 1. Risk Assessment Risk Management Paradigm (adapted from the National Research Council's paradigm, 1983). The red box shows the information included in IRIS assessments.

Overview of EPA's Implementation of NRC's Recommendations

EPA agrees with the NRC's 2011 recommendations for the development of IRIS assessments and is fully implementing them consistent with the Panel's "Roadmap for Revision," which viewed the full implementation of their recommendations by the IRIS Program as a multi-year process. In response to the NRC's 2011 recommendations, the IRIS Program has made changes to streamline the assessment development process, improve transparency, and create efficiencies in the Program. The NRC has acknowledged EPA's successes in this area. Their May 2014 report *"Review of the Integrated Risk Information System (IRIS) Process,"* finds that EPA has made substantial improvements to the IRIS Program in a short amount of time. They also provide several recommendations which they say should be seen as building on the progress that EPA already has made.

The following chapters outline the NRC's 2011 recommendations and provide details regarding changes that the IRIS Program has made and will make in response to the recommendations. Chapter 2 – *NRC's Report, "Review of the Integrated Risk Information System (IRIS) Process"* – provides an overview of the NRC's recent report reviewing the IRIS assessment development

process. This chapter includes highlights from the NRC report and a summary of the committee's recommendations. Chapters 3 and 4 provide details about how the IRIS Program is implementing the NRC 2011 recommendations, and include examples that illustrate the work the program is doing to improve IRIS assessments consistent with those recommendations. These chapters also provide information on the additional activities and initiatives being undertaken to continue to transform the IRIS Program. More specifically, Chapter 3 – *NRC's General Recommendations and Guidance* – provides information on how the IRIS Program is implementing the NRC's general recommendations, including details of the July 2013 enhancements to the IRIS Program. The focus of Chapter 4 – *NRC's Specific Recommendations and Guidance* – is on scientific approaches to identify, evaluate, and synthesize evidence (including weight of evidence evaluation) for hazard identification, select studies for deriving toxicity values, and calculate toxicity values. Both chapters include chemical-specific examples demonstrating the application of scientific methods in IRIS assessments consistent with the NRC's 2011 recommendations. The examples are not to be construed as final Agency conclusions and are provided for the sole purpose of demonstrating how the IRIS Program is currently implementing the NRC recommendations.

Summary

EPA is committed to a strong, vital, and scientifically sound IRIS Program. Over the past three years, EPA has worked to strengthen and streamline the IRIS Program, improve transparency, and create efficiencies. Significant changes have been made in response to the NRC recommendations, and further efforts are underway to fully implement the recommendations. The NRC has acknowledged that EPA has made substantial improvements to the IRIS Program in a short amount of time. As the IRIS Program continues to evolve, EPA is committed to evaluating how well our approaches promote constructive public discussion with our stakeholders, as well as reviewing how our approaches can more effectively facilitate subsequent assessment development.

Chapter 2 – NRC's 2014 Report, *"Review of the Integrated Risk Information System (IRIS) Process"*

In 2012, EPA contracted with the NRC to conduct a comprehensive review of the IRIS assessment development process and changes that are in progress or planned by EPA as a result of the NRC's 2011 review. The NRC also reviewed current methods for integrating and weighing scientific evidence for chemical hazard identification. The NRC convened two public meetings related to this project: 1) a September 2012 meeting to kick off their review of the IRIS assessment development process and 2) a December 2012 meeting to discuss EPA's current and future process for developing IRIS assessments, including topics on identifying, evaluating, and integrating evidence, selecting studies, and calculating toxicity values. EPA also asked the NRC to convene a public workshop to obtain input from the scientific community, as well as stakeholders and the public, on weight-of-evidence considerations. This workshop was held in March 2013.

Additionally, the NRC reviewed documents submitted by the IRIS Program which provide information about the changes that have been or are being made in the program, along with chemical-specific examples of how the Program is implementing NRC recommendations. The NRC considered these materials as they reviewed the IRIS assessment development process.

In May 2014, the NRC released their report reviewing the IRIS assessment development process. In this report, the NRC applauds EPA's efforts to improve IRIS and finds that the Program has moved forward steadily in planning for, and implementing changes in, each element of the assessment process. The report also notes that EPA has made substantial improvements to the IRIS Program in a short time. They specifically note that, "overall, the committee finds that substantial improvements in the IRIS process have been made and it is clear that EPA has embraced and is acting on the recommendations in the NRC formaldehyde report."[2]

The 2014 NRC report also provides several recommendations the Committee says should be seen as building on the progress that EPA already has made.

EPA is grateful to the NRC for their thorough and thoughtful review. The NRC reviewed materials that EPA submitted in the first half of 2013. Since that time, we have continued to evolve, and we have made further changes that are in line with the recommendations in this report. We embrace the recommendations in the NRC report, and we will implement them.

The remaining portion of this chapter provides highlights of the NRC's findings and recommendations. The full list of findings and recommendations is in Appendix B.

General Process Issues

The 2014 NRC report notes that, "overall, the changes that EPA has proposed and implemented to various degrees constitute substantial improvement in the IRIS process" and that "if current

[2] National Research Council, 2014. Review of EPA's Integrated Risk Information System (IRIS) Process.

trajectories are maintained, inconsistencies identified in the present report are addressed, and objectives still to be implemented are successfully completed, the IRIS process will become much more effective and efficient in achieving the program's basic goal of developing assessments that provide an evidence-based foundation for ensuring that chemical hazards are assessed and managed optimally."[3] Of note, the committee agrees that the new document structure for IRIS assessments improves the organization of and streamlines the assessments, and the evidence tables and graphic displays of study findings increases clarity and transparency. The report states that this approach brings IRIS assessments more in line with the state of practice for systematic reviews. The NRC report also comments on stakeholder engagement, for which the IRIS Program has expanded opportunities, agreeing that early and continuing stakeholder involvement not only will increase the likelihood that EPA will address the concerns of diverse stakeholders but should strengthen the quality of IRIS assessments. The NRC also offers several recommendations related to general process issues. Highlights include:

- When extracting data for evidence tables, EPA should use at least two reviewers to assess each study independently for risk of bias;
- EPA should consider ways to provide technical assistance to under-resourced stakeholders to help them to develop and provide input into the IRIS process; and
- IRIS stopping rules should be explicit and transparent, describe when and why the window for evidence inclusion should be expanded, and should be sufficiently flexible to accommodate truly pivotal studies.

Problem Formulation and Protocol Development

While the NRC 2011 report that reviewed the formaldehyde assessment did not provide any specific recommendations regarding problem formulation and protocol development, the NRC 2014 report reviewing the IRIS assessment development process felt that some discussion of these topics was warranted. In their report, they note up front that problem formulation in the IRIS process is restricted to scientific questions that pertain to the first two elements of the risk assessment paradigm: hazard identification and dose-response assessment. The NRC committee suggests a three-step process for conducting problem formulation in which EPA would:

1. Perform a broad literature search designed to identify possible health outcomes associated with the chemical under investigation;
2. Construct a table to guide the formulation of specific questions that would be the subjects of specific systematic reviews; and
3. Examine the table to determine which outcomes warrant a systematic review and how to define the systematic-review question, such as, "Does exposure to chemical X result in neurotoxic effects?"

The NRC recommends that, after the systematic-review questions are specified, EPA develop protocols for conducting the systematic reviews to address the questions, and that any changes made after the protocol is in place should be transparent, with a rationale for the change stated.

Evidence Identification

In their report, the NRC notes that EPA has been responsive to recommendations regarding evidence identification and is well on the way to adopting a more rigorous approach to evidence

[3] National Research Council, 2014. Review of EPA's Integrated Risk Information System (IRIS) Process.

identification that would meet standards for systematic reviews. They recommend that EPA maintain the trajectory for change in this area. They also make several recommendations. Highlights include:

- Protocols for IRIS assessments should include a line-by-line description of the search strategy for each systematic-review question addressed in the assessment that is written in collaboration with information specialists trained in systematic-review methodology;
- Protocols should explicitly state the inclusion and exclusion criteria for studies; and
- EPA should engage information specialists trained in systematic reviews in the process of evidence identification, for example, by having an information specialist peer review the proposed evidence-identification strategy in the protocol for the systematic review.

Evidence Evaluation

The NRC notes in their report that the checklist developed by EPA that is presented in the IRIS preamble successfully addresses many of the concerns raised by the NRC formaldehyde report. The report also notes that EPA has developed broad guidance for assessing the quality of observational studies of exposed human populations and, to a smaller extent, animal toxicology studies. The NRC discusses the concept of "risk of bias," noting that it is related to the internal validity of a study and reflects study design characteristics that can introduce a systematic error that might affect the magnitude and the direction of the apparent effect. They note that incorporating risk-of-bias assessments into the IRIS assessment process might take some time and approaches will depend on the complexity and extent of data on a chemical and the resources available to EPA. Highlights of the NRC's recommendations in the area of evidence evaluation include:

- EPA should explicitly identify factors that can lead to bias in animal studies so that these factors are consistently evaluated for experimental studies and should consider a tool for assessing risk of bias in in vitro studies;
- EPA should conduct a risk of bias assessment on studies used as a primary source of data for the hazard identification and dose-response assessment;
- EPA should specify the empirically based criteria it will use to assess risk of bias for each type of study design in each type of data stream; and
- EPA should consider funding source in the risk of bias assessments conducted for systematic reviews that are part of an IRIS assessment.

Evidence Integration for Hazard Identification

The NRC report discusses the phrase "weight of evidence" and notes that it has become too vague as used in practice today and thus is of little scientific use. The committee found that the phrase "evidence integration" is more useful and more descriptive of what is done in an IRIS assessment. The NRC report notes that the committee appreciates that EPA's improvement for evidence integration are still being developed, and they offer some options for moving forward. The NRC recommends that EPA continue to improve its evidence-integration process incrementally and enhance the transparency of its process. It should either maintain its current guided-expert-judgment process, but make its application more transparent, or adopt a structured (or GRADE[4]-like) process for evaluating evidence and rating recommendations along the lines that National

[4] The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Working Group has developed an approach to grading the quality of evidence and strength of recommendations in health care. More information is available at http://www.gradeworkinggroup.org/index.htm.

Toxicology Program (NTP) has taken. If EPA does move to a structured evidence-integration process, it should combine resources with NTP to leverage the intellectual resources and scientific experience in both organizations. The committee does not offer a preference but suggests that EPA consider which approach best fits its plans for the IRIS process. Highlights of other recommendations in this area include:

- EPA should expand its ability to perform quantitative modeling of evidence integration (e.g., Bayesian modeling of chemical hazards);
- EPA should develop templates for structured narrative justification of the evidence integration process and conclusion; and
- Guidelines for evidence integration for cancer and noncancer endpoints should be more uniform.

Calculation of Toxicity Values

The NRC report noted that the IRIS Program has made a number of responsive changes related to deriving toxicity values, including: 1) developing a process for study selection that requires transparent documentation of study quality, credibility of the evidence of hazard, and adequacy of quantitative dose-response data for determining a point of departure; 2) deriving multiple toxicity values and presenting them graphically; and 3) documenting the approach for conducting dose-response modeling output and considering organ-specific or system-specific and overall toxicity values. The NRC report provides several recommendations related to this area. Highlights include that EPA should:

- Develop criteria for determining when evidence is sufficient to derive toxicity values;
- Continue its shift toward the use of multiple studies rather than single studies for dose-response assessment but with increased attention to risk of bias, study quality, and relevance in assessing human dose-response relationships; and
- Conduct uncertainty analysis systematically and coherently in IRIS assessments and develop IRIS-specific guidelines to frame uncertainty analysis and uncertainty communication.

Future Directions

In their overall evaluation, the 2014 NRC report states that "the committee expects that EPA will complete its planned revisions in a timely way and that the revisions will transform the IRIS program."[5] The committee found that appropriate revisions of all elements of the IRIS assessment process were underway or planned. The report offers three lessons learned that are critical for ensuring the IRIS Program provides the best possible assessment:

1. Assessment methods should be updated in a continuing, strategic fashion;
2. Inefficiencies in the IRIS Program need to be systematically identified and addressed; and
3. Evolving competences that reflect new scientific directions are needed.

The following chapters provide details about how the IRIS Program is implementing the NRC 2011 recommendations. The examples illustrate the work the Program is doing to improve IRIS

[5] National Research Council, 2014. Review of EPA's Integrated Risk Information System (IRIS) Process.

assessments consistent with those recommendations. They also provide information on additional activities and initiatives the Program is undertaking to continue to improve.

Chapter 3 – General Recommendations & Guidance from the 2011 NRC Report on Formaldehyde

In 2011, the NRC made the following general recommendations:

- To enhance the clarity of the document, the draft IRIS assessment needs rigorous editing to reduce the volume of text substantially and address redundancies and inconsistencies. Long descriptions of particular studies should be replaced with informative evidence tables. When study details are appropriate, they could be provided in appendices.
- Chapter 1 needs to be expanded to describe more fully the methods of the assessment, including a description of search strategies used to identify studies with the exclusion and inclusion criteria articulated and a better description of the outcomes of the searches and clear descriptions of the weight-of-evidence approaches used for the various noncancer outcomes. The committee emphasizes that it is 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 the RfCs and unit risk estimates.
- Elaborate an overall, documented, and quality-controlled process for IRIS assessments.
- Ensure standardization of review and evaluation approaches among contributors and teams of contributors; for example, include standard approaches for reviews of various types of studies to ensure uniformity.
- Assess disciplinary structure of teams needed to conduct the assessments.

** Bulleted recommendations presented in boxes in Chapters 3 and 4 are direct quotes from the 2011 NRC report.*

In their 2011 report on formaldehyde, the NRC recommended that the IRIS Program enhance the clarity of IRIS assessments, describe more fully the methods of the assessment, assess the disciplinary structure of assessment teams, and elaborate a process for IRIS assessments to ensure standardization of approaches. In response to these recommendations, the IRIS Program has developed several new initiatives and enhanced existing processes.

In July 2013, EPA announced a series of enhancements to its IRIS Program. EPA is implementing the enhancements for ongoing assessments as practicable, with the goal of improving the scientific integrity of assessments, increasing the productivity of the IRIS Program, and increasing transparency so issues are identified and discussed earlier in the assessment development process. These enhancements incorporate additional opportunities for stakeholder and public engagement at various stages of the IRIS process.

More information on the IRIS Program's recent enhancements can be found at http://www.epa.gov/IRIS/process.htm and http://www.epa.gov/IRIS/pdfs/irisprocessfactsheet2013.pdf.

These new initiatives and enhancements help to ensure transparency throughout the IRIS process and assessment development, and that major science decisions are rigorously vetted. Ultimately, these changes will help EPA meet the goal of using the best available science to produce high quality scientific IRIS assessments in a timely and transparent manner. These new initiatives and enhancements are in line with the NRC's recommendations related to improving the development of IRIS assessments and advancing risk assessment in general, including the importance of up front planning and scoping in the risk assessment process[6].

New Document Structure – Implemented

In their report, the NRC recommended that the IRIS Program enhance the clarity of IRIS assessments by reducing the volume of text and addressing redundancies and inconsistencies. The IRIS Program has fully embraced and implemented this recommendation by revising the assessment template to substantially reduce redundancy and the amount of text. This streamlining also reduced the potential occurrence of inconsistencies that arise from revision and transcription errors. The new template provides sections for the literature search and associated strategy, study selection and evaluation, and methods used to develop the assessment.

The new document structure includes an *Executive Summary* in the beginning of each assessment, which provides a concise summary of the major conclusions of the assessment. Additionally, a newly developed *Preamble* describes the methods used to develop the assessment. Each assessment includes information about the literature search and screening strategy used to identify the available scientific evidence, as well as the criteria and rationale for selecting critical studies to be evaluated in the assessment. The main body of the IRIS assessment has been reorganized into two sections, *Hazard Identification* and *Dose-Response Analysis*, to more clearly delineate identification of potential hazards prior to the development of toxicity values and to further reduce the volume of text and redundancies/inconsistencies. Information on chemical and physical properties, toxicokinetics, individual studies, and assessments by other national and International health agencies has been moved to appendices (which are provided as supplemental information) to further improve the flow of the document.

In the *Hazard Identification* chapter of the new document template, the IRIS Program has developed subsections based on organ/system-specific hazards to systematically synthesize and integrate the available evidence for a given chemical (i.e., epidemiologic, toxicological, and mechanistic data). The assessment now uses evidence tables to succinctly summarize the critical studies to be considered in developing the assessment. These tables present the key study design information and findings that support how toxicological hazards are identified. In addition, exposure-response arrays, which graphically depict responses at different exposure levels for studies, are being used as visual tools to inform the hazard characterization. This chapter provides for a strengthened and more integrated and transparent discussion of the available evidence supporting hazard identification.

The *Dose-Response Analysis* section of the new document structure explains the rationale used to select and advance studies for consideration in calculating toxicity values based on conclusions regarding the potential hazards associated with chemical exposure. Key data supporting the dose-response analysis are reported and the methodology and derivation of toxicity values are described. In addition, details of the dose-response analysis—including the data, models, methods, and software—are provided in appendices as supplemental information and described in sufficient detail to allow for independent replication and verification. The *Dose-Response Analysis* section also includes tables and figures showing candidate toxicity values for comparison across studies and

[6] National Research Council, 2009. Science and Decisions: Advancing Risk Assessment

endpoints. Finally, this section of the new document structure includes clear documentation of the conclusions and selection of the overall toxicity values.

> ***The 2013 draft IRIS Toxicological Review of Benzo[a]pyrene provides an example of the new IRIS assessment template and document structure. It can be found at:*** ***http://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=66193 .***

IRIS Assessment Preamble – Implemented

In their report, the NRC recommended that the IRIS Program expand Chapter 1 of IRIS assessments to "describe more fully the methods of the assessment, including a description of search strategies used to identify studies with the exclusion and inclusion criteria clearly articulated and a better description of the outcomes of the searches and clear descriptions of the weight-of-evidence approaches used for the various noncancer outcomes."

In accordance with this recommendation, the IRIS Program has replaced the previous Chapter 1 of IRIS assessments with a section titled *Preamble to IRIS Toxicological Reviews*, which describes the application of existing EPA guidance and the methods and criteria used in developing the assessments. The term "Preamble" is used to emphasize that these methods and criteria are being applied consistently across IRIS assessments. The new Preamble discusses the following topics:

- Scope of the IRIS Program;
- Process for developing and peer-reviewing IRIS assessments;
- Identifying and selecting pertinent studies;
- Evaluating the quality of individual studies;
- Evaluating the overall evidence of each effect;
- Selecting studies for derivation of toxicity values; and
- Deriving toxicity values.

For each of these topics, the Preamble summarizes and cites EPA guidance on methods used in the assessment. The Preamble was first included in the draft IRIS assessments of ammonia and trimethylbenzenes when they were released for public comment in June 2012, and it has been included in all new IRIS assessments since that time.

> ***An example of the Preamble is available in the 2013 draft IRIS Toxicological Review of Benzo[a]pyrene and can be found starting on page xiv of the document available at:*** ***http://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=66193 .***

IRIS Peer Review – Implemented

Rigorous, independent peer review is a cornerstone of the IRIS Program. Every IRIS assessment is reviewed by a group of recognized experts in scientific disciplines relevant for the particular assessment. The peer review process used for IRIS assessments follows EPA guidance on peer review[7]. At this point, IRIS assessments are expected to be reviewed through EPA's Science Advisory Board (SAB) peer reviews (additional details below); however, some peer reviews may

[7] U.S. EPA (2006) Science Policy Council Peer Review Handbook - 3rd Edition, EPA document number EPA/100/B-06/002.(http://www.epa.gov/peerreview/) and the EPA National Center for Environmental Assessment Policy and Procedures for Conducting IRIS Peer Reviews (2009, http://www.epa.gov/iris/pdfs/Policy_IRIS_Peer_Reviews.pdf).

occur through a contractor-organized process. All peer reviews, regardless of the reviewing body, involve a public comment period and public meeting (usually face-to-face). Following peer review, all revised IRIS assessments include an appendix describing how peer review and public comments were addressed. In 2013, EPA announced improvements to its conflict of interest review process for contractor-managed peer reviews. These improvements ensure that the public has the opportunity to review and comment on a peer review panel's composition when influential scientific documents are being reviewed. Additional details about these improvements are available at: http://www.epa.gov/osa/pdfs/epa-process-for-contractor.pdf

Dedicated Chemical Assessment Advisory Committee - Implemented

EPA's SAB has established a new standing committee, the Chemical Assessment Advisory Committee (CAAC), to review IRIS assessments. In the past, the SAB formed a new committee for each chemical assessment that the SAB reviewed. The new CAAC will provide the same high-level, transparent review as previous SAB reviews, but it will provide more continuous and overlapping membership for consistent advice. EPA expects that the majority of IRIS assessments will be reviewed by the CAAC.

The CAAC is comprised of 26 highly qualified scientists with a broad range of expertise relevant to human health assessment. A group of CAAC members serves on panels reviewing individual IRIS assessments. Panels are augmented with chemical-specific experts or panelists with other areas of expertise needed to review the assessment. The CAAC review process is similar to how other reviews are currently conducted by the SAB and includes the following: the public is invited to nominate peer reviewers for specific assessments; the proposed panels or pools of panelists are posted for public comment; the proposed panelists are screened by an Agency official for conflicts of interest; the final panel is announced prior to the peer review phase.

As part of the process for conducting the peer review, the SAB organizes a public teleconference to take place a few weeks prior to the panel's face-to-face peer review meeting. This teleconference is also available by webinar. The purpose of the teleconference is to discuss the peer review charge questions and to learn about the development of the IRIS assessment under review.

> ***More information on the SAB CAAC can be found at:***
> ***http://yosemite.epa.gov/sab/sabpeople.nsf/WebCommitteesSubcommittees/Chemical%20Assessment%20Advisory%20Committee***

Planning and Scoping– Implemented

The IRIS Program recognizes that it is important to understand the big picture in order to develop an assessment that is most informative and efficient for decision-makers. Having a clear understanding of the overarching environmental problems being addressed in the context of a chemical can help inform what an IRIS assessment will ultimately include. The importance of upfront planning and scoping in the risk assessment process was supported by the NRC in their 2009 report, *Science and Decisions: Advancing Risk Assessment,* where they recommended that EPA provide "greater attention on design in the formative stages of risk assessment." While the NRC was referring to the overall risk assessment paradigm, the spirit of the recommendation supports a scoping step before developing a hazard identification and dose-response assessment (i.e., IRIS assessment).

The IRIS Program is now developing Planning and Scoping summaries for new chemicals and chemicals in early stages of the IRIS process, and conducting internal meetings to identify EPA needs for the assessment. The scoping process involves collecting background information on the chemical, its predominant uses, and the pathways through which humans can be exposed. This early consultation helps ensure that the assessment meets the needs and critical timelines of Agency decision-makers.

The IRIS Program has recently conducted planning and scoping for several chemicals. This information has been provided in the associated preliminary materials released to the public prior to assessment development.

> ***A chemical-specific example of planning and scoping can be found in Chapter 1, "Planning and Scoping Summary" of the Preliminary Materials for Hexabromocyclododecane (HBCD) available at: http://www.epa.gov/ncea/iris/publicmeeting/iris_bimonthly-apr2014/HBCD-preliminary_draft_materials.pdf .***

Problem Formulation – Implemented

The IRIS Program also conducts problem formulation for chemicals prior to assessment development. During this phase, EPA identifies scientific issues that will be important in conducting hazard identification and dose-response assessment. Problem formulation, as conducted within the IRIS Program, draws upon information from other assessments by state, federal, and international health agencies to help identify health endpoints that should be considered in the development of the IRIS assessment. The IRIS Program also uses problem formulation to identify potential scientific issues to be addressed in the assessment, such as human relevance of effects observed in animal studies, questions related to mode of action, or populations with potentially greater susceptibility. The IRIS Program releases this information to the public and discusses it at a public meeting.

> ***Chemical-specific examples of problem formulation can be found Chapter 3 of the "Scoping and Problem Formulation" materials for ethylbenzene and naphthalene available at: http://www.epa.gov/iris/publicmeeting/iris_bimonthly-sep2014/mtg_docs.htm.***

Preliminary Materials for the IRIS Assessment – Implemented

In the early stages of developing the draft assessment, EPA will develop the literature search and associated search and screening strategy and highlight methodological characteristics of studies that will be considered in the evaluation and synthesis of the critical scientific evidence. This evidence is presented in evidence tables and exposure-response arrays. Additionally, as appropriate, anticipated key scientific questions (e.g., emerging areas of research, use of mechanistic information) for the chemical will be included.

The literature search presents the full scope of the scientific literature identified for a chemical, and the associated search and screening strategy describes the processes for identifying the scientific literature, screening studies for consideration, and identifying sources of health effects data, supporting studies, and secondary sources of health effects information. The approach for evaluating methodological features of studies describes the questions and considerations that will

be taken into account when the IRIS Program evaluates the critical studies and synthesizes the evidence for each health effect. Evidence tables succinctly summarize the critical scientific literature and exposure-response arrays graphically depict the health effect responses at different levels of chemical exposure for each study in the evidence tables. The IRIS Program releases this information to the public, receives comments, and holds a public meeting to present these materials and discuss science issues identified by EPA and the public.

> ***A chemical-specific example of Preliminary Materials can be found for Hexabromocyclododecane (HBCD) available at:***
> ***http://www.epa.gov/ncea/iris/publicmeeting/iris_bimonthly-apr2014/HBCD-preliminary_draft_materials.pdf .***

Improved Public Comment and Peer Review – Implemented

During the review stages of the IRIS process, EPA releases the draft assessment and draft peer review charge for public comment and convenes a public meeting to discuss the draft documents and comments. This public meeting replaces the previous IRIS listening session and emphasizes dialogue with stakeholders. The IRIS Program considers the public comments and, in some cases, will revise the draft assessment and peer review charge to respond to the scientific issues raised in the public comments. Additionally, IRIS will summarize the public comments and provide responses to include in the draft assessment. During peer review, EPA will ask the peer review panel to review and comment on whether the IRIS Program adequately addressed the public comments.

Improved Stakeholder Engagement in the IRIS Process and Assessment Development – Implemented

The IRIS Program is committed to proactively engaging with stakeholders and has recently initiated ways to improve stakeholder engagement to help ensure transparency and the use of the best available science in IRIS assessments. Engaging with stakeholders can help facilitate the development of assessments and promote public discussion of key scientific issues. Therefore, scientific engagement with stakeholders and the public is an important part of supporting the best decisions possible.

The IRIS Program considers a stakeholder to be any individual or group that participates in, has an impact on, or could be affected by products produced by the IRIS Program. Public and stakeholder engagement has always been an important part of the IRIS assessment development process. The May 2009 IRIS process provided multiple opportunities for engagement including: (1) public and stakeholder nomination of chemicals for assessment; (2) a public listening session for each draft assessment; (3) public review and comment of draft assessments; (4) a public peer review process; and (5) two opportunities for review and comment on draft assessments by other EPA scientists, other Federal agencies, and the Executive Office of the President.

In November 2012, the IRIS Program convened a public meeting to engage with stakeholders and discuss the IRIS Program in general. The meeting was intended to begin a series of dialogues between the IRIS Program and a broad and diverse group of stakeholders. The goals of the meeting were to: engage stakeholders in the IRIS process; listen to views and needs of IRIS users in an open and respectful environment; facilitate improvements to the IRIS process; and initiate an ongoing dialogue between the IRIS Program and stakeholders. A summary of this meeting is available on the IRIS website at: http://www.epa.gov/iris/publicmeeting/stakeholders-kickoff/index.htm.

Based on this interaction with stakeholders, in combination with recommendations provided by the NRC, in July 2013, EPA announced enhancements to the IRIS Program. The 2013 enhanced IRIS assessment development process includes the following additional opportunities for engagement:

- Before beginning to develop a draft assessment, the IRIS Program will conduct an internal planning and scoping meeting to identify EPA needs for the assessment. The IRIS Program will then release planning and scoping information and convene a public meeting focused on scientific issues and studies that may inform EPA's plan for developing the assessment (i.e., problem formulation).
- In the early stages prior to assessment development, EPA will release preliminary materials and receive public comments. The IRIS Program will hold a public meeting to present these materials and discuss science issues.
- During the review stages of the IRIS process, EPA will release the draft assessment and draft peer review charge for public comment and receive comments. The IRIS Program will hold a public meeting to present the draft assessment and discuss science issues prior to external peer review.

Since the enhancements were announced, the IRIS Program has introduced a series of bimonthly public meetings to allow the public the opportunity to provide input and participate in discussions about preliminary materials and draft IRIS assessments for specific chemicals (as noted in the latter two bullets above). The first meeting, held December 12-13, 2013, provided stakeholders with the opportunity to give input and participate in an open discussion regarding preliminary materials that were prepared for IRIS chemicals prior to the development of the draft assessment. The discussion at this meeting centered on the following chemicals: ethyl tert-butyl ether (ETBE); tert-butyl alcohol (tert-butanol); and hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX). This meeting also was an opportunity for the public to provide input and discussion on draft assessments and draft charges to the peer review panel prior to external peer review for the following chemicals: ethylene oxide (EtO) and benzo[a]pyrene (BaP). Additional information about this meeting is available at: http://www.epa.gov/iris/publicmeeting/iris_bimonthly-dec2013/index.htm.

The December 2013 bimonthly public meeting was the first opportunity for stakeholders to comment on preliminary materials and draft assessments following the introduction of the enhancements in July 2013. These meetings allow stakeholders and IRIS scientists to engage in robust and productive scientific discussion and exchange information and perspectives on science issues associated with the chemicals undergoing assessment. The objective of this discussion is to ensure that the subsequent development of the draft IRIS assessments will reflect the most critical scientific issues and various perspectives on those issues. In 2014, EPA held four bimonthly meetings that discussed preliminary materials on seven chemicals at varying stages of assessment development. Six meetings have been scheduled for 2015. Additional information on the bimonthly public science meetings is available at: http://www.epa.gov/iris/publicmeeting/index.htm.

In October 2014, EPA announced an agreement with the NRC to provide independent expert advice to the IRIS Program on the scientific and technical aspects of IRIS chemical assessments through participation in the IRIS bimonthly public meetings. This initiative is consistent with recommendations made in the NRC May 2014 report, *Review of EPA's Integrated Risk Information System (IRIS) Process* on expanding the breadth of perspectives made available to the Agency. Addition of experts to IRIS bimonthly public meetings ensures an independent and diverse range of

expert scientific and technical perspectives are represented at the meetings and reflects EPA's commitment to scientific rigor and integrity.

Public Peer Consultation Science Workshops – Implemented

The IRIS Program has begun to hold public peer consultation science workshops to enhance the input from the scientific community as assessments are designed. The overarching goal of these workshops is to better interpret and evaluate the latest scientific evidence. Information regarding specific peer consultation workshops is announced to the public in advance of the meetings. The goal of each workshop varies. For example, the workshops may focus on the state-of-the-science for a particular chemical or provide a forum for discussion with experts about certain cross-cutting scientific issues that may impact the development of a scientifically complex assessment. In the past year, the IRIS Program held peer consultation workshops focusing on:

IRIS Workshop on the NRC Recommendations (October 15-16, 2014): The purpose of this workshop was to discuss recommendations from the National Academies' National Research Council (NRC) May 2014 report on ways to further improve the scientific quality of IRIS assessments. The discussions at the workshop focused on NRC recommendations related to:

1. Systematic integration of the primary lines of evidence (human, animal, mechanistic) considered in IRIS assessments, including the use of guided expert judgment and structured processes for evidence integration;
2. Adapting systematic review methodologies for IRIS, including discussion of factors that can lead to study bias;
3. Advancing dose-response analysis through combining multiple studies, and the types of values that should be derived in quantitative analysis of toxicity; and
4. Advancing dose-responses analysis through better characterization of uncertainty and variability, including discussion of how uncertainty and variability estimates affect users of IRIS assessments as well as practical approaches for characterizing uncertainty and variability

The workshop was open to the public and broadcast by webinar/teleconference. Additional information is available at: http://www.epa.gov/iris/irisworkshops/NRC_workshop/index.htm.

State-of-the-Science Workshop to Discuss Issues Relevant for Assessing the Health Hazards of Formaldehyde Inhalation (April 30-May 1, 2014): The purpose of this workshop was to discuss issues relevant for assessing the health hazards of formaldehyde inhalation. The workshop included discussion of several scientific issues related to the development of the draft IRIS assessment of formaldehyde (inhalation exposure). It focused on the following three themes:

1. Epidemiological research examining the potential association between formaldehyde exposure and lymphohematopoietic cancers (leukemia and lymphomas);
2. Mechanistic evidence relevant to formaldehyde inhalation exposure and these types of cancers; and
3. Evidence pertaining to the influence of formaldehyde that is produced endogenously (by the body during normal biological processes) on the toxicity of inhaled formaldehyde, and implications for the health assessment.

This workshop was open to the public and broadcast by webinar/teleconference.Additional information is available at: http://www.epa.gov/iris/irisworkshops/fa/index.htm.

State-of-the-Science Workshop on Chemically-induced Mouse Lung Tumors (January 7-8, 2014): The purpose of this workshop was to discuss the available data and interpretation of results from studies of mouse alveolar/bronchiolar adenomas and carcinomas (lung tumors) following exposure to chemical agents and the relevance of such tumors in mice to human cancer risk. Several lines of research have investigated whether or not these types of lung tumors are formed by a mode of action (MOA) which is specific to mice and are relevant to tumor formation or other toxicity in humans. This is an important issue for the IRIS assessments for naphthalene, styrene, and ethylbenzene. This workshop was open to the public and broadcast by webinar/teleconference. Additional information is available at: http://www.epa.gov/iris/irisworkshops/mltw/index.htm.

Scientific Workshop on Hexavalent Chromium (September 19 & 25, 2013): The purpose of this workshop was to discuss scientific issues relevant to evaluating the potential human health effects of ingesting hexavalent chromium. An important component of determining the cancer causing potential of ingested hexavalent chromium is understanding the rates at which this metal is effectively detoxified in the gastrointestinal tract. To address this, EPA convened a state-of-the-science workshop where an expert panel discussed this issue. This workshop was open to the public and convened entirely by webinar/teleconference. Additional information is available at: http://www.epa.gov/iris/irisworkshops/cr6/index.htm

Workshop on Applying Systematic Review to Assessments of Health Effects of Chemical Exposures (August 26, 2013): EPA is developing and implementing approaches to enhance its scientific assessments, particularly in the area of increasing transparency and clarity related to evaluating evidence and drawing conclusions. To inform these efforts, EPA held a public workshop to provide scientific input on issues related to the use of a systematic review process for evaluating potential health hazards of chemical exposures. The goal for this workshop was to receive scientific input regarding approaches for different steps within a systematic review, such as evaluating individual studies, synthesizing evidence within a particular discipline, and integrating evidence across different disciplines to draw scientific conclusions and causality determinations. This workshop was open to the public and broadcast by webinar/teleconference. Additional information is available at: http://www.epa.gov/iris/irisworkshops/systematicreview/index.htm

The IRIS Program is also developing additional scientific workshops to take place in 2015. These workshops will be announced on the IRIS website.

> ***More information on the IRIS Program's public meetings and workshops can be found at http://www.epa.gov/iris/publicmeeting/index.htm.***

Discipline-Specific Workgroups and Interdisciplinary Science Teams – Implemented

The IRIS Program has created discipline-specific workgroups which coordinate across assessments to ensure consistency, solve cross-cutting issues, and advance scientific understanding that contributes to decision-making in IRIS assessments. The discipline-specific workgroups cover topics related to: reproductive/developmental toxicity, neurotoxicity, respiratory/inhalation toxicity, systemic and general toxicity, immunotoxicity, cancer, epidemiology, toxicity pathways/genetic toxicity, statistics and dose-response analysis, and physiologically-based pharmacokinetic modeling.

The expertise needed for each chemical undergoing assessment by the IRIS Program is chemical-specific. The various areas of expertise that are needed are identified in the early stages of planning

and document development. For each assessment, discipline-specific workgroups and relevant scientific personnel are assigned to lead or assist in the development of the assessment.

IRIS Needs Assessment/Multi-year Plan – **In Progress**

As part of its 2013 review of EPA's chemical assessment process, GAO recommended that the Agency conduct an assessment to determine the demand for IRIS assessments. In response to this recommendation, EPA initiated an IRIS multi-year planning effort which prioritizes the chemicals on the 2012 IRIS Agenda. This Agency-wide initiative was implemented to ensure that the IRIS Program's chemical assessment agenda is best suited to meet the Agency's regulatory needs. It is anticipated that the results of this multi-year planning effort will be shared with the public in the third quarter FY2015.

Stopping Rules – **Implemented**

The IRIS Program has developed a set of "stopping rules" for new data and scientific issues to help ensure that IRIS assessments are not delayed by new research findings or ongoing debate of scientific issues after certain process points have passed.

- For new data: All published and ongoing studies relevant to the assessment should be identified by the end of the public meeting to discuss the preliminary materials (i.e., the literature search, evidence tables, and exposure-response arrays). Exceptions may be made when new information appears that could change the major findings of an assessment or when the peer review panel explicitly recommends that EPA incorporate recent research before completing an assessment.
- For scientific issues: Alternative interpretations of the science and perspectives on how to bridge scientific uncertainty should be raised early in the assessment development process. Opportunities to propose scientific approaches and interpretations are available: during the public meeting to discuss the preliminary materials; during the public comment period and public meeting to discuss the draft IRIS assessment and draft peer review charge; or during the public external peer review process. Scientific issues that are raised, but not resolved, will be highlighted for the peer review panel for their input.

Additional information about the stopping rules is available at:
http://www.epa.gov/iris/pdfs/IRIS_stoppingrules.pdf.

Chapter 4
– Specific Recommendations & Guidance from the 2011 NRC Report on Formaldehyde.

In 2011, the NRC made twenty-five specific recommendations in five broad categories:

- evidence identification,
- evidence evaluation,
- weight-of-evidence evaluation,
- selection of studies for derivation of toxicity values, and
- calculation of toxicity values.

The IRIS Program has been working to improve the approaches for identifying and selecting pertinent studies; evaluating and displaying studies; strengthening and improving integration of evidence for hazard identification; and increasing transparency in dose-response analysis. The IRIS Program is adopting the principles of systematic review in IRIS assessments with regard to providing an overview of methods and points to consider in the process of developing and documenting decisions. The focus of IRIS assessments is typically on the evidence of health effects (any kind of health effects) of a particular chemical. This is, by definition, a broad topic. The systematic review process that has been developed and applied within the clinical medicine arena (evidence-based medicine) is generally applied to narrower, more focused questions. Nonetheless, the experiences within the clinical medicine field provide a strong foundation to draw upon. The IRIS Program held a workshop in August 2013 on this topic in order to have a public discussion of systematic review approaches that may be applicable to IRIS assessments. Additional information about the workshop is available at:
http://www.epa.gov/iris/irisworkshops/systematicreview/index.htm

An IRIS assessment is made up of multiple systematic reviews. It is an iterative process that ultimately identifies relevant scientific information needed to address key assessment-specific questions. The initial steps of the systematic review process formulate specific strategies to identify and select studies relating to each key question, evaluate study methods based on clearly defined criteria, and transparently document the process and its outcomes. Synthesizing and integrating data also falls under the purview of systematic review.

One of the strengths of systematic review is its ability to identify relevant studies, published and unpublished, pertaining to the question of interest (e.g., what are the health effects of a chemical?). Additionally, by transparently presenting all decision points and the rationale for each decision, bias in study selection and evaluation is eliminated.

Evidence Identification: Literature Collection and Collation Phase

NRC Recommendations:

- Select outcomes on the basis of available evidence and understanding of mode of action.
- Establish standard protocols for evidence identification.
- Develop a template for description of the search approach.
- Use a database, such as the Health and Environmental Research Online (HERO) database, to capture study information and relevant quantitative data.

Identifying and Selecting Pertinent Studies – Implemented

The IRIS Program has incorporated a systematic and transparent approach for identifying evidence in the new IRIS assessment document structure, with a separate section that provides a detailed description of the literature search and associated search and screening strategy. The literature search presents the full scope of the scientific literature identified for a chemical. The search and screening strategy describes the processes for identifying the scientific literature, screening studies for consideration, and identifying the sources of health effects data, along with supporting studies and secondary sources of health effects information. The strategy serves as a standard protocol for evidence identification for IRIS assessments.

The IRIS Program developed a template for this strategy. It is included as a separate section in the new document structure. While the approach is consistent across IRIS assessments, the detailed search strategies and results of the literature search and screening are specific to each chemical assessment. Additionally, the strategy utilizes a graphical display documenting how initial search findings are narrowed to the final studies that are selected for further evaluation. This evidence identification process is first documented and released publicly as part of the preliminary materials developed prior to assessment development. Later, as the assessment is developed, the literature search and screening strategy will be expanded upon, as appropriate, for each endpoint. This section provides a link to the Health and Environmental Research Online (HERO) database (www.epa.gov/hero), an external database that contains the references that are identified as a result of the literature search.

The Preamble to IRIS Toxicological Reviews includes additional information in Section 3 on "Identifying and selecting pertinent studies."

A chemical-specific example of the implementation of this recommendation is available in Chapter 2, "Draft Literature Search and Screening Strategy" of the Preliminary Materials for Hexabromocyclododecane (HBCD) available at:
http://www.epa.gov/iris/publicmeeting/iris_bimonthly-apr2014/HBCD-preliminary_draft_materials.pdf

Evidence Evaluation: Hazard Identification

NRC Recommendations:

- All critical studies need to be thoroughly evaluated with standardized approaches that are clearly formulated and based on the type of research, for example, observational, epidemiologic, or animal bioassays. The findings of the reviews might be presented in tables to ensure transparency.
- Standardize the presentation of reviewed studies in tabular or graphic form to capture the key dimensions of study characteristics, weight of evidence, and utility as a basis for deriving reference values and unit risks.
- Standardized evidence tables for all health outcomes need to be developed. If there were appropriate tables, long text descriptions of studies could be moved to an appendix or deleted.
- Develop templates for evidence tables, forest plots, or other displays.
- Establish protocols for review of major types of studies, such as epidemiologic and bioassay.

Standardized Approaches to Evaluating Evidence – In Progress

The IRIS Program is improving the approach to evaluating evidence and standardizing the documentation of this evaluation. This step in the systematic review process involves a uniform evaluation of a variety of methodological features (e.g., study design, exposure measurement details, data analysis, and presentation) of studies that will be considered in the overall evaluation and synthesis of evidence for each health effect. Critical studies identified after the literature search and screen are evaluated for aspects of the design, conduct, or reporting that could affect the interpretation of results and the overall contribution to the synthesis of evidence for determining hazard potential. Much of the key information for conducting this evaluation can generally be found in the study's methods section and in how the study results are reported. Importantly, this evaluation does not consider study results or, more specifically, the direction or magnitude of any reported effects. For example, standard issues for evaluation of experimental animal data identified by the NRC and adopted in this approach include consideration of the species and sex of animals studied, dosing information (dose spacing, dose duration, and route of exposure), endpoints considered, and the relevance of the endpoints to the human endpoints of concern.

The purpose of this step is generally not to eliminate studies, but rather to evaluate studies with respect to potential methodological considerations (e.g., the purity of a chemical used in a study, study protocols that may result in systematic underestimation of the frequency of an effect) that could affect the interpretation of and relative confidence in the results. It is worth emphasizing that the systematic evaluation of studies is conducted on multiple levels. The evaluation, to a certain extent, can be conducted at an early stage of assessment development (i.e., after identifying the scientific literature and developing evidence tables); however, the complete evaluation would be conducted during the data evaluation and synthesis of hazard characterization. This first-level review is presented in the preliminary materials that the IRIS Program is preparing for chemicals prior to assessment development. Ultimately, this systematic evaluation may inform decisions about which studies to use for hazard identification and it informs decisions about which studies to move forward for dose-response modeling for derivation of toxicity values.

The Preamble to IRIS Toxicological Reviews includes additional information in Section 4 ("Evaluating the quality of individual studies").

A chemical-specific example of the implementation of this recommendation is available in Chapter 3, "Selection of Studies for Hazard Identification" of the Preliminary Materials for Hexabromocyclododecane (HBCD) available at: http://www.epa.gov/iris/publicmeeting/iris_bimonthly-apr2014/HBCD-preliminary_draft_materials. pdf.

Standardized Presentation of Reviewed Studies (i.e., Evidence Tables and Exposure-Response Arrays) – Implemented

The IRIS Program has developed templates for evidence tables to present key study data from critical studies in a standardized tabular format. The evidence tables succinctly summarize the study design and findings (both positive and negative results), organized by specific outcome or endpoint of toxicity, and also facilitate the evaluation described above. In general, the evidence tables include all studies that inform the overall synthesis of evidence for hazard potential. The studies that are considered to be most informative will depend on the extent and nature of the database for a given chemical, but may encompass a range of study designs and include epidemiology, toxicology, and other toxicity data, when appropriate. Additionally, exposure-response arrays graphically depict the health effect responses at different levels of chemical exposure for each study in the evidence tables.

A chemical-specific example of the implementation of this recommendation is available in Appendix A, "Preliminary Evidence Tables and Exposure-Response Arrays" of the Preliminary Materials for Hexabromocyclododecane (HBCD) available at: http://www.epa.gov/iris/publicmeeting/iris_bimonthly-apr2014/HBCD-preliminary_draft_materials.pdf .

Weight-of-Evidence Evaluation: Integration of Evidence for Hazard Identification

NRC Recommendations:

- Strengthened, more integrative, and more transparent discussions of weight of evidence are needed. The discussions would benefit from more rigorous and systematic coverage of the various determinants of weight of evidence, such as consistency.
- Review use of existing weight-of-evidence guidelines.
- Standardize approach to using weight-of-evidence guidelines.
- Conduct agency workshops on approaches to implementing weight-of-evidence guidelines.
- Develop uniform language to describe strength of evidence on noncancer effects.
- Expand and harmonize the approach for characterizing uncertainty and variability.
- To the extent possible, unify consideration of outcomes around common modes of action rather than considering multiple outcomes separately.

Integration of Evidence for Hazard Identification – In Progress

The IRIS Program has strengthened and increased transparency in the weight-of-evidence[8] approach for identifying hazards in IRIS assessments. Hazard identification involves the integration of evidence from human, animal, and mechanistic studies in order to draw conclusions about the hazards associated with exposure to a chemical. In general, IRIS assessments integrate evidence in the context of Hill (1965), which outlines aspects — such as consistency, strength, coherence, specificity, does-response, temporality, and biological plausibility — for consideration of causality in epidemiologic investigations that were later modified by others and extended to experimental studies (U.S. EPA, 2005a).

All results, both positive and negative, of potentially relevant studies that have been evaluated for quality are considered (U.S. EPA, 2002) to answer the fundamental question: "Does exposure to chemical X cause hazard Y?" This requires a critical weighing of the available evidence (U.S. EPA, 2005a; 1994), but is not to be interpreted as a simple tallying of the number of positive and negative studies (U.S. EPA, 2002). Hazards are identified by an informed, expert evaluation and integration of the human, animal, and mechanistic evidence streams.

The Preamble to IRIS Toxicological Reviews includes additional information in Section 5 ("Evaluating the overall evidence of each effect").

A chemical-specific example of the implementation of this recommendation is available in Chapter 1, "Hazard Identification" of the 2013 draft IRIS Toxicological Review of Benzo[a]pyrene and can be found at:
http://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=66193 .

[8] The terminology describing weight-of-evidence approaches is evolving. In the 2014 NRC report, the committee found the term "evidence integration to be more useful and more descriptive of the process that occurs after the completion of systematic reviews." For the purposes of this Report to Congress, the term weight-of-evidence is used to be consistent with the 2011 NRC report terminology.

Currently, the IRIS Program is using existing guidelines that address these issues to inform assessments. In addition, the IRIS Program is taking a more systematic approach in analyzing the available human, animal, and mechanistic data being used in IRIS assessments. In conducting this analysis and developing the synthesis, the IRIS Program evaluates the data for the:

- strength of the relationship between the exposure and response and the presence of a dose-response relationship;
- specificity of the response to chemical exposure and whether the exposure precedes the effect;
- consistency of the association between the chemical exposure and response; and
- biological plausibility of the response or effect and its relevance to humans.

The IRIS Program uses this weight of evidence approach to identify the potential hazards associated with chemical exposure.

In May 2014, the NRC released their report reviewing the IRIS assessment development process. As part of this review, the NRC reviewed current methods for evidence-based reviews and made several recommendations with respect to integrating scientific evidence for chemical hazard and dose-response assessments. In their report, the NRC states that EPA should continue to improve its evidence-integration process incrementally and enhance the transparency of its process. They note that EPA should either maintain its current guided-expert-judgment process but make its application more transparent or adopt a structured (or GRADE-like) process for evaluating evidence and rating recommendations along the lines of the approach that NTP has taken. If EPA does move to a structured evidence-integration process, it should combine resources with NTP to leverage the intellectual resources and scientific experience in both organizations. The committee does not offer a preference, but suggests that EPA consider which approach best fits its plans for the IRIS process. The NRC recommendations will inform the IRIS Program's efforts in this area going forward.

The IRIS Program recognizes the benefit of adopting a formal weight-of-evidence framework, which may include elements of both structured and guided expert judgement processes, to define standardized classification of causality. The IRIS Program convened a workshop in October 2014 to discuss approaches to evidence integration. Several workshop participants emphasized the need for both expert judgement and structure. As part of this workshop, the various approaches that are currently in use were acknowledged and compared for their strengths and limitations. The workshop included scientists with expertise in the classification of chemicals for various health effects. The workshop was open to the public. The Agency is in the process of evaluating the information received during the workshop (as well as 2014 NRC report) and anticipates making decisions about weight-of-evidence evaluations as we move forward with assessment development in 2015.

Additional information, including the workshop agenda and presentations, is available at: http://www.epa.gov/iris/irisworkshops/NRC_workshop/index.htm.

Selection of Studies for Derivation of Toxicity Values

NRC Recommendations:

- The rationales for the selection of the studies that are advanced for consideration in calculating the RfCs and unit risks need to be expanded. All candidate RfCs should be evaluated together with the aid of graphic displays that incorporate selected information on attributes relevant to the database.
- Establish clear guidelines for study selection.
- Balance strengths and weaknesses.
- Weigh human vs. experimental evidence.
- Determine whether combining estimates among studies is warranted.

Selection of Studies for Dose-Response Analysis – Implemented

The IRIS Program has improved the process for selecting studies for derivation of toxicity values as well as increasing the transparency about this process by providing an improved discussion and rationale. After identifying hazards (e.g., developmental, reproductive, and immunological), the IRIS Program evaluates studies within each effect category in order to identify a subset of studies to be considered for the derivation of toxicity values.

The first step is determining whether the quantitative exposure and response data are available to derive a point of departure (POD). The POD can be a no-observed-adverse-effect-level [NOAEL], lowest-observed-adverse-effect-level [LOAEL], or the benchmark dose/concentration lower confidence limit [BMDL/BMCL]). Additional attributes (aspects of the study, data characteristics, and relevant considerations) pertinent to deriving toxicity values are used as criteria to evaluate the subset of studies for dose-response analysis (described in more detail in EPA guidance documents). Thus, the most relevant, informative studies are selected to move forward. The new document structure provides a transparent discussion of the studies identified for dose-response analysis.

> ***The Preamble to IRIS Toxicological Reviews includes additional information in Section 6 ("Selecting studies for dose-response analysis").***
>
> ***A chemical-specific example of the implementation of this recommendation is available in Chapter 2, "Dose-Response Analysis" of the 2013 draft IRIS Toxicological Review of Benzo[a]pyrene and can be found at:***
> ***http://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=66193 .***

Considerations for Combining Data for Dose-Response Modeling – In Progress

The IRIS Program is now routinely considering whether combining data among studies is warranted for the derivation of toxicity values. For most IRIS assessments, the POD had been derived based on data from a single study dataset. This is because in most cases, datasets are often expected to be heterogeneous for biological or study design reasons.

However, there are cases where conducting dose-response modeling after combining data from multiple studies can be considered, resulting in a single POD based on multiple datasets. For instance, this may be useful to increase precision in the POD or to quantify the impact of specific sources of heterogeneity. The IRIS Program has developed considerations for combining data for

dose-response modeling to be taken into account when performing dose-response analysis for an IRIS assessment.

In addition, multiple PODs or toxicity values can be combined (considering, for example, the highest quality studies, the most sensitive outcomes, or a clustering of values) to derive a single, overall toxicity value (or "meta-value"). For example, the IRIS assessment for trichloroethylene (TCE) identified multiple candidate reference doses (RfDs) that fell within a narrow dose range, and selected an overall RfD that reflected the midpoint among the similar candidate RfDs. This RfD is supported by multiple effects/studies (i.e., less sensitive to limitations of individual studies) (for more information, see: http://www.epa.gov/iris/subst/0199.htm).

Calculation of Reference Values and Unit Risks

NRC Recommendations:

- Describe and justify assumptions and models used. This step includes review of dosimetry models and the implications of the models for uncertainty factors; determination of appropriate points of departure (such as benchmark dose, no-observed-adverse-effect level, and lowest observed-adverse-effect level), and assessment of the analyses that underlies the points of departure.
- Provide explanation of the risk-estimation modeling processes (for example, a statistical or biologic model fit to the data) that are used to develop a unit risk estimate.
- Provide adequate documentation for conclusions and estimation of reference values and unit risks. As noted by the committee throughout the present report, sufficient support for conclusions in the formaldehyde draft IRIS assessment is often lacking. Given that the development of specific IRIS assessments and their conclusions are of interest to many stakeholders, it is important that they provide sufficient references and supporting documentation for their conclusions. Detailed appendixes, which might be made available only electronically, should be provided, when appropriate.
- Assess the sensitivity of derived estimates to model assumptions and end points selected. This step should include appropriate tabular and graphic displays to illustrate the range of the estimates and the effect of uncertainty factors on the estimates.

Conducting and Documenting Dose-Response Modeling and Deriving Toxicity Values – Implemented

IRIS assessments, in general, include dose-response analysis to derive toxicity values. In response to NRC recommendations,the IRIS Program has improved the quality control of the overall dose-response modeling process and increased transparency by documenting the approach for conducting dose-response modeling. Part of this documentation is achieved with the addition of considerations for selecting organ/system-specific and overall toxicity values, and a streamlined dose-response modeling output (both part of the new document structure). Additionally, tools and approaches to manage data and ensure quality (e.g., Data Management and Quality Control for Dose-Response Modeling) in dose-response analyses have been developed. The objectives are to minimize errors, maintain a transparent system for data management, automate tasks where possible, and maintain an archive of data and calculations used to develop assessments.

The IRIS Program has improved the documentation of dose-response modeling. Preamble Section 7 provides a description of the process for dose-response analysis. In addition, the text describing the dose-response analysis will include a description of how the toxicity values were derived and will cite EPA guidelines, where appropriate. EPA is working to address the recommendations on characterizingand communicatinguncertainty discussed in the 2014 NRC report in IRIS assessments.

> ***The Preamble to IRIS Toxicological Reviews includes additional information in Section 7 ("Deriving toxicity values").***
>
> ***A chemical-specific example of the implementation of this recommendation is available in Chapter 2, "Dose-Response Analysis" of the 2013 draft IRIS Toxicological Review of Benzo[a]pyrene and in Appendix E, "Dose-Response Modeling for the Derivation of Reference Values for Effects other than Cancer and the Derivation of Cancer Risk***

Estimates" in the Supplemental Information for Benzo[a]pyrene. These documents can be found at: http://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=66193 .

Chapter 5 – Updates on the IRIS Assessments of Formaldehyde and Acrylonitrile

The examples provided in previous chapters are the IRIS Program's best illustrations of how the NRC recommendations are being implemented in assessments. They also highlight the scientific methods the Program is using to assess, synthesize, and draw conclusions regarding likely human health effects associated with exposures to substances. Consistent with the advice of the NRC, EPA has been implementing their recommendations using a phased approach, making the most extensive changes to documents that are in the earlier steps of the assessment development process. For assessments that were in the later stages of development when the NRC report was published, EPA has been implementing the recommendations as feasible without taking the assessments backwards to earlier steps of the process. The IRIS assessments of formaldehyde and acrylonitrile are currently being revised to incorporate the NRC recommendations. Additional information regarding the status of these assessments is provided below.

IRIS Assessment of Formaldehyde: EPA's IRIS Program has been working on an updated assessment of formaldehyde (inhalation route of exposure). In June 2010, EPA released a draft assessment for public comment and independent expert scientific peer review by the NRC. The NRC completed their review and published their review report in April 2011. Since that time, the IRIS Program has been working to fully address the NRC's recommendations, as well as the comments received from the public.

To inform the development of the revised IRIS assessment of formaldehyde, EPA convened a workshop on April 30 and May 1, 2014, to facilitate scientific discussion about three broad issues and the scientific challenges they pose for assessing the health hazards of inhaling formaldehyde:

1. Epidemiological research examining the potential association between formaldehyde exposure and lymphohematopoietic cancers (leukemia and lymphomas);
2. Mechanistic evidence relevant to formaldehyde inhalation exposure and these types of cancers; and
3. The influence of formaldehyde that is produced endogenously (by the body during normal biological processes) when assessing the health hazards (especially excess cancer risk) of inhaled formaldehyde.

This workshop was open to the public and available either in person (in Arlington, VA) or by webinar/teleconference. Additional details about the workshop are available on the IRIS website at: http://www.epa.gov/iris/irisworkshops/fa/index.htm

Information from this workshop will inform the development of the revised draft IRIS assessment of formaldehyde. In developing the revised draft assessment, EPA is following the NRC's recommendations, as stated in the NRC's Report on Formaldehyde and highlighted in Chapters 3 and 4 of this Report to Congress. The revised draft formaldehyde assessment will be released for public comment and rigorous, independent expert peer review by the SAB CAAC.

IRIS Assessment of Acrylonitrile: In June 2011, EPA released a draft IRIS assessment for acrylonitrile for public comment. A public listening session for the draft acrylonitrile assessment was held in August 2011. Since that time, the IRIS Program has been working to develop a revised draft assessment. In developing the revised draft assessment, EPA is making revisions in response to the public comments and is following the NRC's recommendations, as stated in the NRC's Report on Formaldehyde and highlighted in Chapters 3 and 4 of this Report to Congress. The revised draft acrylonitrile assessment will be released for public comment and rigorous, independent expert peer review by the SAB CAAC.

Appendix A – 2014 Consolidated Appropriations Act

The 2014 Consolidated Appropriations Act included the following language related to EPA's IRIS Program:

Integrated Risk Information System (IRIS).- The Committees note that House Report 112-331 directed EPA to contract with the National Academy of Sciences (NAS) to conduct reviews of IRIS assessments with the goal of improving EPA's IRIS assessments. The Committees recognize that the agreed-upon NAS review is ongoing and that the Agency is taking steps to address previous NAS recommendations. To that end, the Agency shall include in each draft and final IRIS assessment released in fiscal year 2014, documentation describing how EPA has implemented or addressed NAS Chapter 7 recommendations. If any recommendations were not incorporated, the Agency should explain its rationale.

Further, EPA should ensure the new draft of the formaldehyde assessment reflects those recommended improvements. Specifically, EPA should adhere to the recommendation in Chapter 7 of the NAS report that "strengthened, more integrative and more transparent discussions of weight of the evidence are needed." Conducting a risk assessment for formaldehyde presents many challenges, due largely to the significant database for this compound. Although several evaluations have been conducted, none has formally integrated toxicological and epidemiological evidence. EPA should ensure the forthcoming revised draft IRIS assessment of formaldehyde is a model of transparency and represents an objective and robust integration of the scientific evidence.

The Committees understand EPA has decided to make further revisions to the acrylonitrile assessment to more fully address scientific issues in the assessment. Therefore, the Agency is directed to review methods previously used to evaluate and interpret the body of available scientific data, including the weight-of-evidence approach. Further, and no later than May 1, 2014, the Agency shall provide to the House and Senate Committees on Appropriations a progress report that describes the Agency's implementation of NAS Chapter 7 recommendations for fiscal years 2012 and 2013.

The progress report shall include a chapter on whether there are more appropriate scientific methods to assess, synthesize, and draw conclusions regarding likely human health effects associated with likely exposures to substances. The Agency also should discuss the current re-evaluation of the formaldehyde and acrylonitrile assessments as well as any other assessments that may be relevant as case studies. This chapter should include a discussion of the methods previously used by the Agency to evaluate and interpret the body of available scientific data and include descriptions of any quantitative methods used to combine evidence to support hypotheses, such as the weight-of-evidence approach.

Appendix B – Findings and Recommendations from the 2014 NRC Report, *Review of EPA's Integrated Risk Information System (IRIS) Process*

Findings	Recommendations
General Process Issues	
Finding: The committee is impressed and encouraged by EPA's progress, recognizing that the implementation of the recommendations in the NRC formaldehyde report is still in process. If current trajectories are maintained and objectives still to be implemented are successfully brought to fruition, the IRIS process will have become much more effective and efficient in achieving its basic goal of developing human-health assessments that can provide the scientific foundation for ensuring that risks posed to public health by chemicals are assessed and managed optimally.	Recommendation: EPA needs to complete the changes in the IRIS process that are in response to the recommendations in the NRC formaldehyde report and specifically complete documents, such as the draft handbook, that provide detailed guidance for developing IRIS assessments. When those changes and the detailed guidance, such as the draft handbook, have been completed, there should be an independent and comprehensive review that evaluates how well EPA has implemented all the new guidance. The present committee is completing its report while those revisions are still in progress.
Finding: Although it is clear that quality control (QC) of the IRIS assessment process is critical for the outcome of the program, the documents provided do not sufficiently discuss the QC processes or provide guidelines that adequately separate the technical methods from the activities of QC management and program oversight. For example, the role of the CASTs in the QC process is not specifically described.	Recommendation: EPA should provide a quality-management plan that includes clear methods for continuing assessments of the quality of the process. The roles of the various internal entities involved in the process, such as the CASTs, should be described. The assessments should be used to improve the overall process and the performance of EPA staff and contractors.
	Recommendation: When extracting data for evidentiary tables, EPA should use at least two reviewers to assess each study independently for risk of bias. The reliability of the independent coding should be calculated; if there is good agreement, multiple reviewers might not be necessary.

Findings	Recommendations
Finding: The current scoping process for obtaining input from within the agency is clear, but opportunities for stakeholder input from outside EPA early in the process are less clear.	Recommendation: EPA should continue its efforts to develop clear and transparent processes that allow external stakeholder input early in the IRIS process. It should develop communication and outreach tools that are tailored to meet the needs of the various stakeholder groups. For example, EPA might enhance its engagement with the scientific community through interactions at professional-society meetings, advertised workshops, and seminars. In contrast, greater use of social media might help to improve communications with environmental advocacy groups and the public.
Finding: EPA has taken steps to expand opportunities for stakeholder input and discussion that are likely to improve assessment quality. However, not all stakeholders with an interest in the IRIS process have the resources to provide timely comments.	Recommendation: Similar to other EPA technical-assistance programs, EPA should consider ways to provide technical assistance to under-resourced stakeholders to help them to develop and provide input to the IRIS program.
Finding: Promoting efficiency in the IRIS program is paramount given the constraint of inevitably shrinking resources. Thus, the committee agrees with EPA that stopping rules are needed given that the process for some IRIS assessments has become too long as revisions are repeatedly made to the assessments to accommodate new evidence and review comments.	Recommendation: The stopping rules should be explicit and transparent, should describe when and why the window for evidence inclusion should be expanded, and should be sufficiently flexible to accommodate truly pivotal studies. Such rules could be included in the preamble.
	Recommendation: Regarding promotion of efficiencies, EPA should continue to expand its efforts to develop computer systems that facilitate storage and annotation of information relevant to the IRIS mission and to develop automated literature and screening procedures, sometimes referred to as text-mining.
Finding: The draft handbook and other materials are useful but lack explicit guidance as to the methods and nature of the use of expert judgment throughout the full scope of the assessment-development process, from literature searching and screening through integrating evidence to analyzing the dose-response relationship and deriving final toxicity values.	Recommendation: More details need to be provided on the recognition and applications of expert judgment throughout the assessment-development process, especially in the later stages of the process. The points at which expert judgment is applied should be identified, those applying the judgment should be listed, and consideration should be given to harmonizing the use of expert judgment at various points in the process.

Findings	Recommendations
Problem Formulation and Protocol Development	
Finding: The materials provided to the committee by EPA describe the need for carefully constructed literature searches but do not provide sufficient distinction between an initial survey of the literature to identify putative adverse outcomes of interest and the comprehensive literature search that is conducted as part of a systematic review of an identified putative outcome.	Recommendation: EPA should establish a transparent process for initially identifying all putative adverse outcomes through a broad search of the literature. The agency should then develop a process that uses guided expert judgment to identify the specific adverse outcomes to be investigated, each of which would then be subjected to systematic review of human, animal, and in vitro or mechanistic data.
	Recommendation: For all literature searches, EPA should consult with an information specialist who is trained in conducting systematic reviews.
Finding: A protocol is an essential element of a systematic review. It makes the methods and the process of the review transparent, can provide the opportunity for peer review of the methods, and stands as a record of the review.	Recommendation: EPA should include protocols for all systematic reviews conducted for a specific IRIS assessment as appendixes to the assessment.
Evidence Identification	
Finding: EPA has been responsive to recommendations in the NRC formaldehyde report regarding evidence identification and is well on the way to adopting a more rigorous approach to evidence identification that would meet standards for systematic reviews. This finding is based on a comparison of the draft EPA materials provided to the committee with IOM standards.	Recommendation: The trajectory of change needs to be maintained.

Findings	Recommendations
Finding: Current descriptions of search strategies appear inconsistently comprehensive, particularly regarding (a) the roles of trained information specialists; (b) the requirements for contractors; (c) the descriptions of search strategies for each database and source searched; (d) critical details concerning the search, such as the specific dates of each search and the specific publication dates included; and (e) the periodic need to consider modifying the databases and languages to be searched in updated and new reviews. The committee acknowledges that recent assessments other than the ones that it reviewed might already address some of the indicated concerns.	Recommendation: The current process can be enhanced with more explicit documentation of methods. Protocols for IRIS assessments should include a section on evidence identification that is written in collaboration with information specialists trained in systematic reviews and that includes a search strategy for each systematic-review question being addressed in the assessment. Specifically, the protocols should provide a line-by-line description of the search strategy, the date of the search, and publication dates searched and, as noted in Chapter 3, explicitly state the inclusion and exclusion criteria for studies.
	Recommendation: Evidence identification should involve a predetermined search of key sources, follow a search strategy based on empirical research, and be reported in a standardized way that allows replication by others. The search strategies and sources should be modified as needed on the basis of new evidence on best practices. Contractors who perform the evidence identification for the systematic review should adhere to the same standards and provide evidence of experience and expertise in the field.
Finding: One problem for systematic reviews in toxicology is identifying and retrieving toxicologic information outside the peer-reviewed public literature.	Recommendation: EPA should consider developing specific resources, such as registries, that could be used to identify and retrieve information about toxicology studies reported outside the literature accessible by electronic searching. In the medical field, clinical-trial registries and US legislation that has required studies to register in ClinicalTrials.gov have been an important step in ensuring that the total number of studies that are undertaken is known.

Findings	Recommendations
Finding: Replicability and quality control are critical in scientific undertakings, including data management. Although that general principle is evident in IRIS assessments that were reviewed, tasks appear to be assigned to a single information specialist or review author. There was no evidence of the information specialist's or reviewer's training or of review of work by others who have similar expertise. As discussed in Chapter 2, an evaluation of validity and reliability through inter-rater comparisons is important and helps to determine whether multiple reviewers are needed. This aspect is missing from the IOM standards.	Recommendation: EPA is encouraged to use at least two reviewers who work independently to screen and select studies, pending an evaluation of validity and reliability that might indicate that multiple reviewers are not warranted. It is important that the reviewers use standardized procedures and forms.
Finding: Another important aspect of quality control in systematic reviews is ensuring that information is not double-counted. Explicit recognition of and mechanisms for dealing with multiple publications that include overlapping data from the same study are important components of data management that are not yet evident in the draft handbook.	Recommendation: EPA should engage information specialists trained in systematic reviews in the process of evidence identification, for example, by having an information specialist peer review the proposed evidence-identification strategy in the protocol for the systematic review.
Finding: The committee did not find enough empirical evidence pertaining to the systematic-review process in toxicological studies to permit it to comment specifically on reporting biases and other methodologic issues, except by analogy to other, related fields of scientific inquiry. It is not clear, for example, whether a reporting bias is associated with the language of publication for toxicological studies and the other types of research publications that support IRIS assessments or whether any such bias (if it exists) might be restricted to specific countries or periods.	Recommendation: EPA should encourage and support research on reporting biases and other methodologic topics relevant to the systematic-review process in toxicology.
Finding: The draft preamble and handbook provide a good start for developing a systematic, quality-controlled process for identifying evidence for IRIS assessments.	Recommendation: EPA should continue to document and standardize its evidence-identification process by adopting (or adapting, where appropriate) the relevant IOM standards described in Table 4-1. It is anticipated that its efforts will further strengthen the overall consistency, reliability, and transparency of the evidence-identification process.

Findings	Recommendations
Evidence Evaluation	
Finding: The checklist developed by EPA that is presented in the preamble and detailed in the draft handbook addresses many of the concerns raised by the NRC formaldehyde report. EPA also has developed broad guidance for the assessment of the quality of observational studies of exposed human populations and, to a smaller extent, animal toxicology studies. It has not developed criteria for the evaluation of mechanistic toxicology studies. Still lacking is a clear picture of the assessment tools that EPA will develop to assess risk of bias and of how existing assessment tools will be adapted.	Recommendation: To advance the development of tools for assessing risk of bias in different types of studies (human, animal, and mechanistic) used in IRIS assessments, EPA should explicitly identify factors, in addition to those discussed in this chapter, that can lead to bias in animal studies—such as control for litter effects, dosing, and methods for exposure assessment—so that these factors are consistently evaluated for experimental studies. Likewise, EPA should consider a tool for assessing risk of bias in in vitro studies.
Finding: The development of standards for evaluating individual studies for risk of bias is most advanced in human clinical research. Even in that setting, the evidence base to support the standards is modest and expert guidance varies. Furthermore, many of the individual criteria included in risk-of-bias assessment tools, particularly for animal studies and epidemiologic studies, have not been empirically tested to determine how the various sources of bias influence the results of individual studies. The validity and reliability of the tools also have not been tested. Finding: Thus, the committee acknowledges that incorporating risk-of-bias assessments into the IRIS process might take additional time; the ability to do so will vary with the complexity and extent of data on each chemical and with the resources available to EPA. However, the use of standard risk-of-bias criteria by trained coders has been shown to be efficient.	Recommendation: When considering any method for evaluating individual studies, EPA should select a method that is transparent, reproducible, and scientifically defensible. Whenever possible, there should be empirical evidence that the methodologic characteristics that are being assessed in the IRIS protocol have systematic effects on the direction or magnitude of the outcome. The methodologic characteristics that are known to be associated with a risk of bias should be included in the assessment tool. Additional quality-assessment items relevant to a particular systematic-review question also could be included in the EPA assessment tool.

Findings	Recommendations
	Recommendation: EPA should carry out, support, or encourage research on the development and evaluation of empirically based instruments for assessing bias in human, animal, and mechanistic studies relevant to chemical-hazard identification. Specifically, there is a need to test existing animal-research assessment tools on other animal models of chemical exposures to ensure their relevance and generalizability to chemical-hazard identification. Furthermore, EPA might consider pooling data collected for IRIS assessment to determine whether, among various contexts, candidate risk-of-bias items are associated with overestimates or underestimates of effect.
	Recommendation: Although additional methodologic work might be needed to establish empirically supported criteria for animal or mechanistic studies, an IRIS assessment needs to include a transparent evaluation of the risk of bias of studies used by EPA as a primary source of data for the hazard assessment. EPA should specify the empirically based criteria it will use to assess risk of bias for each type of study design in each type of data stream. Recommendation: To maintain transparency, EPA should publish its risk-of-bias assessments as part of its IRIS assessments. It could add tables that describe the assessment of each risk-of-bias criterion for each study and provide a summary of the extent of the risk of bias in the descriptions of each study in the evidence tables.
Finding: The nomenclature of the various factors that are considered in evaluating risk of bias is variable and not well standardized among the scientific fields relevant to IRIS assessments. Such terminology has not been standardized for IRIS assessments.	Recommendation: EPA should develop terminology for potential sources of bias with definitions that can be applied during systematic reviews.
Finding: Although reviews of human clinical studies have shown that study funding sources and financial ties of investigators are associated with research outcomes that are favorable for the sponsors, less is known about the extent of funding bias in animal research.	Recommendation: Funding sources should be considered in the risk-of-bias assessment conducted for systematic reviews that are part of an IRIS assessment.

Findings	Recommendations
Finding: An important weakness of all existing tools for assessing methodologic characteristics of published research is that assessment requires full reporting of the research methods. EPA might be hampered by differences in traditions of reporting risk of bias among fields in the scientific literature.	Recommendation: EPA should contact investigators to obtain missing information that is needed for the evaluation of risk of bias and other quality characteristics of included studies. The committee expects that, as happened in the clinical literature in which additional reporting standards for journals were implemented (Turner et al. 2012), the reporting of toxicologic research will eventually improve as risk-of- bias assessments are incorporated into the IRIS program. However, a coordinated approach by government agencies, researchers, publishers, and professional societies will be needed to improve the completeness and accuracy of the reporting of toxicology studies in the near future.
Finding: EPA has not developed procedures that describe how the evidence evaluation for individual studies will be incorporated, either qualitatively or quantitatively, into an overall assessment.	Recommendation: The risk-of-bias assessment of individual studies should be carried forward and incorporated into the evaluation of evidence among data streams.
Evidence Integration for Hazard Identification	
Finding: Critical considerations in evaluating a method for integrating a diverse body of evidence for hazard identification are whether the method can be made transparent, whether it can be feasibly implemented under the sorts of resource constraints evident in today's funding environment, and whether it is scientifically defensible.	Recommendation: EPA should continue to improve its evidence-integration process incrementally and enhance the transparency of its process. It should either maintain its current guided-expert-judgment process, but make its application more transparent, or adopt a structured (or GRADE-like) process for evaluating evidence and rating recommendations along the lines that NTP has taken. If EPA does move to a structured evidence-integration process, it should combine resources with NTP to leverage the intellectual resources and scientific experience in both organizations. The committee does not offer a preference but suggests that EPA consider which approach best fits its plans for the IRIS process.

Findings	Recommendations
Finding: Quantitative approaches to integrating evidence will be increasingly needed by and useful to EPA.	Recommendation: EPA should expand its ability to perform quantitative modeling of evidence integration; in particular, it should develop the capacity to do Bayesian modeling of chemical hazards. That technique could be helpful in modeling assumptions about the relevance of a variety of animal models to each other and to humans, in incorporating mechanistic knowledge to model the relevance of animal models to humans and the relevance of human data for similar but distinct chemicals, and in providing a general framework within which to update scientific knowledge rationally as new data become available. The committee emphasizes that the capacity for quantitative modeling should be developed in parallel with improvements in existing IRIS evidence-integration procedures and that IRIS assessments should not be delayed while this capacity is being developed.
Finding: EPA has instituted procedures to improve transparency, but additional gains can be achieved in this arena. For example, the draft IRIS preamble provided to the committee states that "to make clear how much the epidemiologic evidence contributes to the overall weight of the evidence, the assessment may select a standard descriptor to characterize the epidemiologic evidence of association between exposure to the agent and occurrence of a health effect" (EPA 2013a, p. B-6). A set of descriptor statements was provided, but they were not used in the recent IRIS draft assessments of methanol and benzo[a]pyrene.	Recommendation: EPA should develop templates for structured narrative justifications of the evidence-integration process and conclusion. The premises and structure of the argument for or against a chemical's posing a hazard should be made as explicit as possible, should be connected explicitly to evidence tables produced in previous stages of the IRIS process, and should consider all lines of evidence (human, animal, and mechanistic) used to reach major conclusions.
Finding: EPA guidelines for evidence integration for cancer and noncancer end points are different; the cancer guidelines are more developed and more specific.	Recommendation: Guidelines for evidence integration for cancer and noncancer end points should be more uniform.

Findings	Recommendations
Derivation of Toxicity Values	
Finding: EPA develops toxicity values for health effects for which there is "credible evidence of hazard" after chemical exposure and of an adverse outcome.	Recommendation: EPA should develop criteria for determining when evidence is sufficient to derive toxicity values. One approach would be to restrict formal dose-response assessments to when a standard descriptor characterizes the level of confidence as medium or high (as in the case of noncancer end points) or as "carcinogenic to humans" or "likely to be carcinogenic to humans" for carcinogenic compounds. Another approach, if EPA adopts probabilistic hazard classification, is to conduct formal dose-response assessments only when the posterior probability that a human hazard exists exceeds a predetermined threshold, such as 50% (more likely than not likely that the hazard exists).
Finding: EPA has made a number of substantive changes in the IRIS program since the publication of the NRC formaldehyde report, including the derivation and graphical presentation of multiple dose-response values and a shift away from choosing a particular study as the "best" study for derivation of dose-response estimates.	Recommendation: EPA should continue its shift toward the use of multiple studies rather than single studies for dose-response assessment but with increased attention to risk of bias, study quality, and relevance in assessing human dose-response relationships. For that purpose, EPA will need to develop a clear set of criteria for judging the relative merits of individual mechanistic, animal, and epidemiologic studies for estimating human dose-response relationships.
Finding: Although subjective judgments (such as identifying which studies should be included and how they should be weighted) remain inherent in formal analyses, calculation of toxicity values needs to be prespecified, transparent, and reproducible once those judgments are made.	Recommendation: EPA should use formal methods for combining multiple studies and the derivation of IRIS toxicity values with an emphasis on a transparent and replicable process.

Findings	Recommendations
Finding: EPA could improve documentation and presentation of dose-response information.	Recommendation: EPA should clearly present two dose-response estimates: a central estimate (such as a maximum likelihood estimate or a posterior mean) and a lower-bound estimate for a POD from which a toxicity value is derived. The lower bound becomes an upper bound for a cancer slope factor but remains a lower bound for a reference value.
Finding: Advanced analytic methods, such as Bayesian methods, for integrating data for dose-response assessments and deriving toxicity estimates are underused by the IRIS program.	Recommendation: As the IRIS program evolves, EPA should develop and expand its use of Bayesian or other formal quantitative methods in data integration for dose-response assessment and derivation of toxicity values.
Finding: IRIS-specific guidelines for consistent, coherent, and transparent assessment and communication of uncertainty remain incompletely developed. The inconsistent treatment of uncertainties remains a source of confusion and causes difficulty in characterizing and communicating uncertainty.	Recommendation: Uncertainty analysis should be conducted systematically and coherently in IRIS assessments. To that end, EPA should develop IRIS-specific guidelines to frame uncertainty analysis and uncertainty communication. Moreover, uncertainty analysis should become an integral component of the IRIS process.