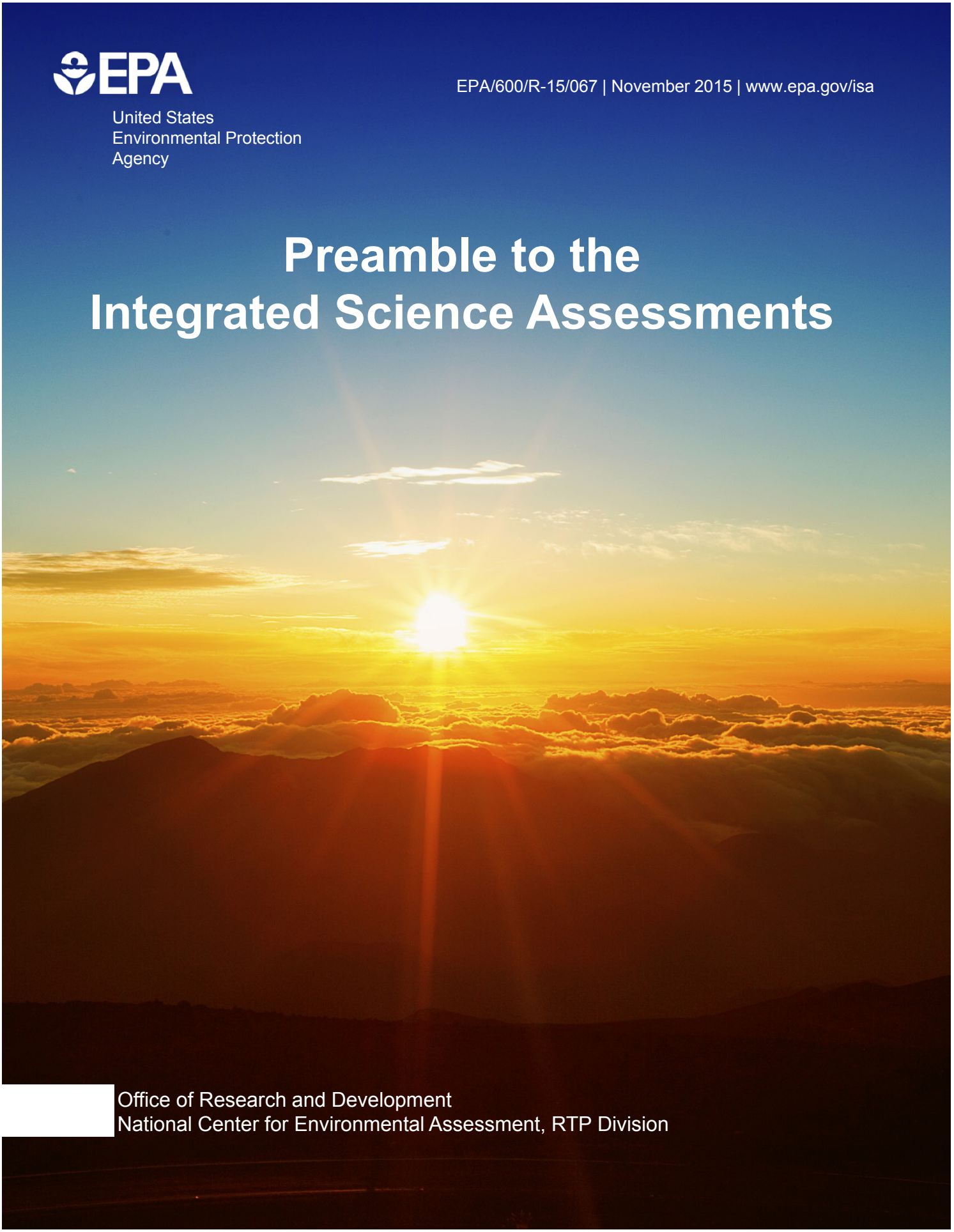
EPA/600/R-15/067 | November 2015 | www.epa.gov/isa

United States
Environmental Protection
Agency

Preamble to the Integrated Science Assessments



Office of Research and Development
National Center for Environmental Assessment, RTP Division

EPA/600/R-15/067
November 2015
www.epa.gov/isa

Preamble to the Integrated Science Assessments

November 2015

National Center for Environmental Assessment—RTP Division
Office of Research and Development
U.S. Environmental Protection Agency
Research Triangle Park, NC

DISCLAIMER

This Preamble serves as a companion document to the Integrated Science Assessments prepared by the U.S. Environmental Protection Agency's National Center for Environmental Assessment in support of reviews of the national ambient air quality standards (NAAQS). It is available on the NCEA website at: www.epa.gov/isa. This document and the approach described herein may be refined in consideration of advice and comments received from the Clean Air Scientific Advisory Committee and the public during subsequent NAAQS reviews. This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

CONTENTS

ACRONYMS AND ABBREVIATIONS

Acronym/Abbreviation	Meaning
CAA	Clean Air Act
CASAC	Clean Air Scientific Advisory Committee
CDC	Centers for Disease Control and Prevention
CO	carbon monoxide
HERO	Health and Environmental Research Online
IARC	International Agency for Research on Cancer
IOM	Institute of Medicine
IRP	Integrated Review Plan
ISA	Integrated Science Assessment
NAAQS	National Ambient Air Quality Standards
NAS	National Academy of Sciences
NO_X	oxides of nitrogen; the sum of nitric oxide and nitrogen dioxide
O_3	ozone
Pb	lead
PM	particulate matter
SES	socioeconomic status
U.S. EPA	U.S. Environmental Protection Agency

PREAMBLE

1. Process of Integrated Science Assessment Development

This Preamble outlines the general process used by the U.S. Environmental Protection Agency (U.S. EPA) for developing an Integrated Science Assessment (ISA) including the framework for evaluating weight of evidence and drawing scientific conclusions and causal judgments. It is available on the U.S. EPA website at: www.epa.gov/isa. The ISA provides a concise review, synthesis, and evaluation of the most policy-relevant science to serve as a scientific foundation for the review of the National Ambient Air Quality Standards (NAAQS).[1] The NAAQS are established based on consideration of the air quality criteria (represented by the ISA) for the pollutants identified by the Administrator using Section 108 of the Clean Air Act (CAA). The pollutants currently identified are carbon monoxide (CO), lead (Pb), oxides of nitrogen, ozone and related photochemical oxidants, particulate matter (PM), and sulfur oxides (CAA, 1990a, b). Figure I depicts the general NAAQS review process. Information for individual NAAQS reviews is available online.[2]

[1] The general process for NAAQS reviews is described at http://www.epa.gov/ttn/naaqs/review.html.
[2] Information for individual NAAQS reviews is available at www.epa.gov/ttn/naaqs.

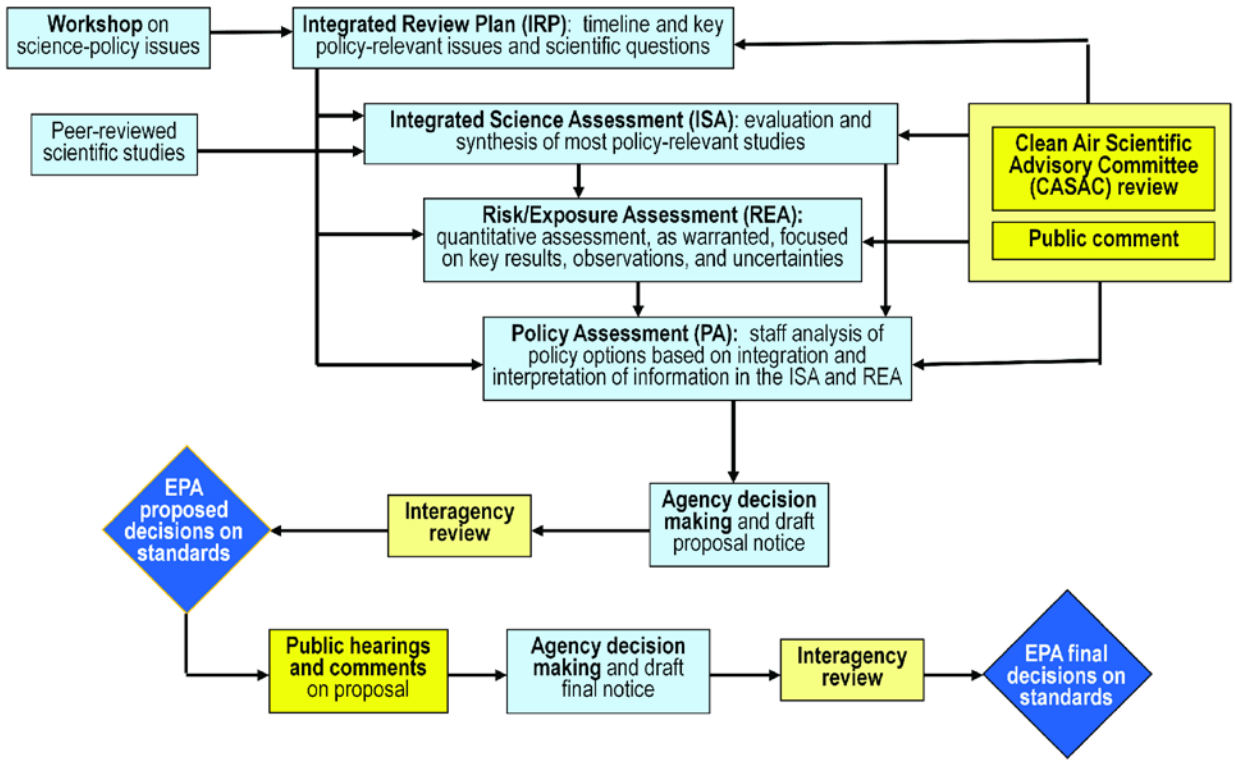


Figure I **Schematic of the key steps in review of the National Ambient Air Quality Standards.**

The development of the ISA is preceded by the release of an Integrated Review Plan (IRP) that discusses the planned scope of the NAAQS review; the planned approaches for developing the key assessment documents [e.g., ISA, Risk and Exposure Assessment (if warranted), Policy Assessment]; and the schedule for release and review of the documents and subsequent rulemaking notices. The key policy-relevant questions included in the IRP serve to clarify and focus the NAAQS review on the critical scientific and policy issues, including addressing uncertainties discussed during the previous review and newly emerging literature. The IRP is informed by a U.S. EPA-hosted public science and policy issue workshop that kicks off review of the NAAQS for each criteria pollutant by seeking input on the current state of the science and engaging stakeholders and experts in discussion of the policy-relevant questions that will frame the review.

This Preamble is a general discussion of the basic steps and criteria used in developing an ISA. Details and considerations specific to an individual ISA are included in the IRP, Preface, and introductory material for that assessment. The general process for ISA development is illustrated in Figure II. An initial step (not shown) is publication of a call

for information in the *Federal Register* that invites the public to provide information relevant to the assessment, such as new or recent publications on health or welfare effects of the pollutant or data from the fields of atmospheric and exposure science.

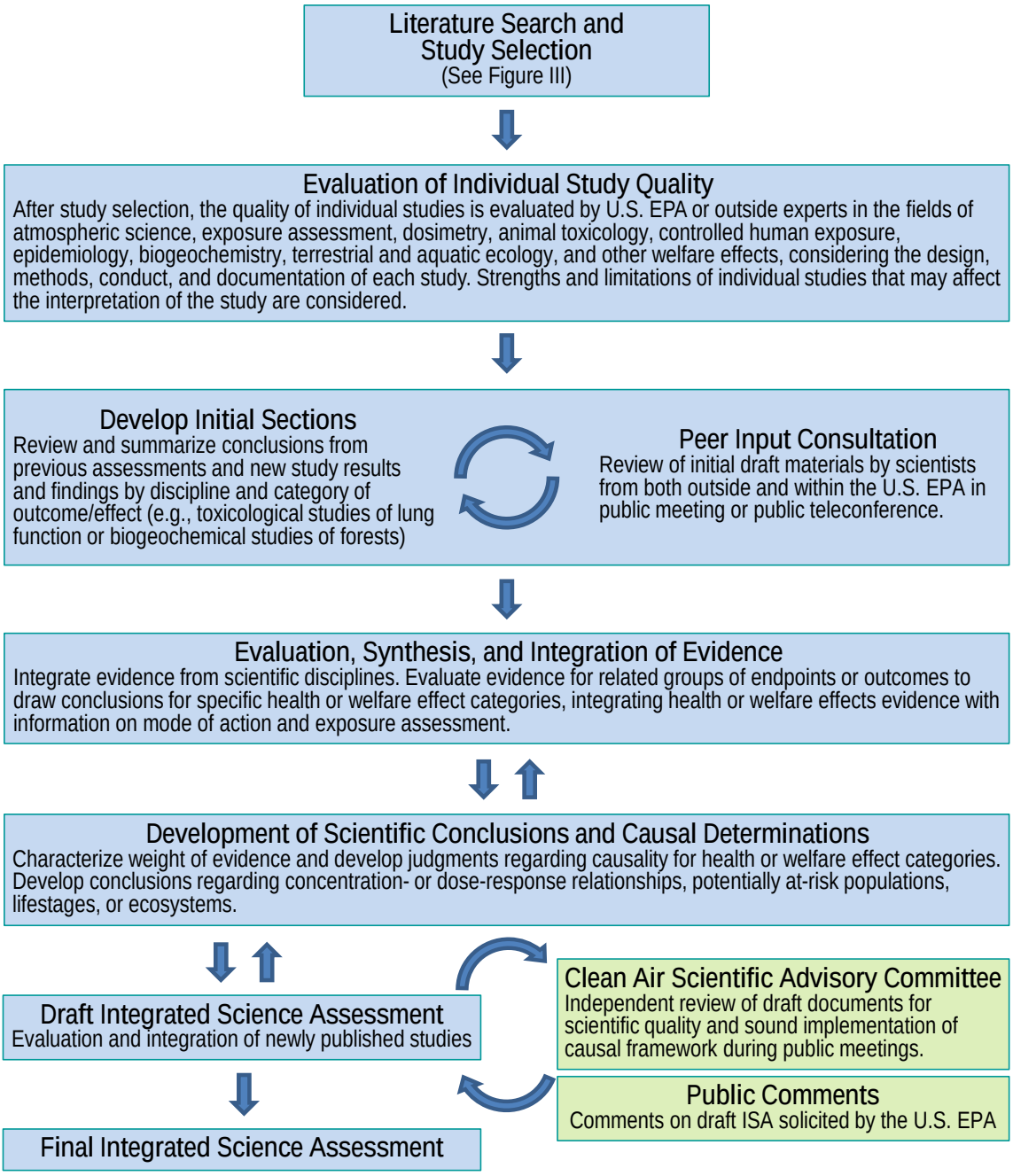


Figure II **Characterization of the general process for Integrated Science Assessment development.**

The fundamental process for developing an ISA includes:

- Literature searches;
- Study selection;
- Evaluation of individual study quality;
- Evaluation, synthesis, and integration of the evidence; and
- Development of scientific conclusions and causal determinations.

In developing an ISA, the U.S. EPA reviews and summarizes the evidence from studies on atmospheric sciences, dosimetry, human exposure, animal toxicology, mode of action, controlled human exposure, epidemiology, biogeochemistry, and/or terrestrial and aquatic ecology and other welfare[1] effects. In the process of developing the first draft ISA, the U.S. EPA may convene a peer input meeting in which the scientific content of preliminary draft materials is reviewed by subject-matter experts to ensure that the ISA is up-to-date and is focused on the most policy-relevant findings. This review also assists the U.S. EPA with integration of evidence within and across disciplines.

The U.S. EPA integrates the evidence from across scientific disciplines or study types and characterizes the weight of evidence for relationships between the pollutant and various outcomes. Integrating evidence on human health or welfare effects involves collaboration between scientists from various disciplines. For example, an evaluation of human health effects evidence would generally include integrating the results from epidemiologic, controlled human exposure, and toxicological studies; consideration of exposure assessment, dosimetry and mode of action; and the application of the causal framework (described below) to draw conclusions.

Integration of results on human health or welfare effects that are logically or mechanistically connected (e.g., respiratory symptoms, asthma exacerbations) informs judgments of causality on the broader health or welfare effect category (e.g., effects on the respiratory system). Using the causal framework described in this Preamble, U.S. EPA scientists consider aspects, such as strength, consistency, coherence, and biological plausibility of the evidence, and develop causality determinations on the nature of the relationships. Causality determinations often entail an iterative process of review and evaluation of the evidence. One or more drafts of the ISA are released for review by the Clean Air Scientific Advisory Committee (CASAC) and the public, and comments

[1] Under CAA Section 302(h) [42 U.S.C. 7602(h)], language referring to "effects on welfare" includes, but is not limited to, "effects on soils, water, crops, vegetation, man-made materials, animals, wildlife, weather, visibility, and climate, damage to and deterioration of property, and hazards to transportation, as well as effects on economic values and on personal comfort and well-being."

received on the characterization of the science as well as the implementation of the causal framework are carefully considered in revising and completing the ISA.

2. Literature Search

In addition to the call for information in the *Federal Register* referenced above, the U.S. EPA maintains an ongoing literature search process to identify relevant scientific studies published since the last ISA for a given criteria pollutant. Search strategies are designed a priori for pollutants and scientific disciplines and iteratively modified to optimize identification of pertinent publications. In addition, papers are identified for inclusion in several additional ways: specialized searches on specific topics, identification of new publications by relational searches conducted using citations from previous assessments, review of tables of contents for journals in which relevant papers may be published, identification of relevant literature by expert scientists, review of citations in previous assessments, and recommendations by the public and CASAC during the call for information and external review processes. References identified through the multipronged search strategy are then "screened" by title and abstract. Those references that are potentially relevant after reading the title are "considered" for inclusion in the ISA and are added to the Health and Environmental Research Online (HERO) database developed by the U.S. EPA.[1] Additional review steps (described in Section 3 below) precede a decision to "include" a study in the ISA. The references cited in the ISA contain a hyperlink to the HERO database. This literature search and study selection process including the "screened", "considered", and "included" references, is depicted in Figure III.

[1] The list of considered references and bibliographic information is accessible to the public through HERO (http://hero.epa.gov/).

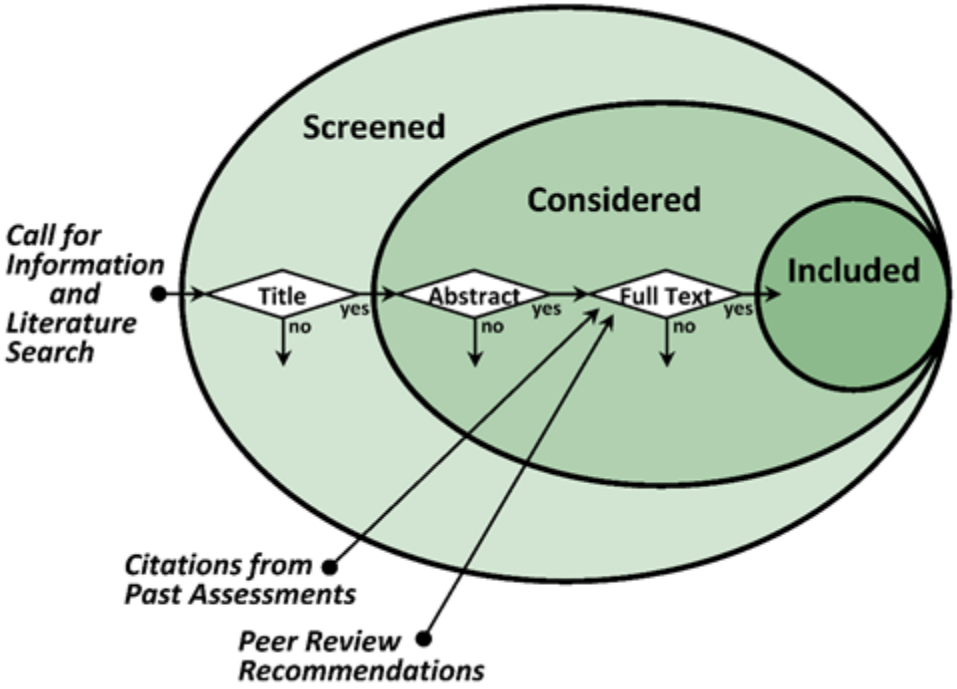


Figure III **Illustration of processes for literature search and study selection process used for development of Integrated Science Assessments.**

Studies and reports that have undergone scientific peer review and have been published (or accepted for publication) are considered for inclusion in the ISA. This includes only studies that have been ethically conducted (e.g., with approval by an Institutional Review Board or Institutional Animal Care and Use Committee). All relevant epidemiologic, controlled human exposure, toxicological, and ecological and other welfare effects studies published since the last review are considered, including those related to exposure-response relationships, mode(s) of action, and populations, lifestages, or ecosystems at increased risk of air pollution-related effects. Studies and data analyses on atmospheric chemistry, air quality and emissions, environmental fate and transport, dosimetry, toxicokinetics, and exposure are also considered for inclusion in the ISA. References considered for inclusion in a specific ISA can be found using the HERO website (http://hero.epa.gov).

Each ISA builds upon the conclusions of previous assessments for the pollutant under review. The U.S. EPA focuses on peer-reviewed literature published following the completion of the previous review and on any new interpretations of previous literature, integrating the results of recent scientific studies with previous findings. Important earlier studies may be discussed in detail to reinforce key concepts and conclusions or for

reinterpretation in light of newer data. Earlier studies also are the primary focus in some areas of the document where research efforts have subsided, or if these earlier studies remain the definitive works available in the literature.

3. Study Selection

References considered for inclusion in the ISA undergo abstract and full-text review to determine whether they will be included in the ISA. The selection process is based on the extent to which the study is informative, pertinent, and policy relevant. Informative, pertinent, and policy-relevant studies include those that describe or provide a basis for the relationship between the criteria pollutant and effects, including studies that offer innovation in method or design and studies that reduce uncertainty on critical issues. Emphasis is placed on studies that examine effects associated with pollutant concentrations and exposure conditions relevant to current human population and ecosystem exposures, and particularly those pertaining to concentrations currently found in ambient air. Other studies are included if they contain unique data, such as a previously unreported effect or mode of action for an observed effect, or examine multiple concentrations to elucidate exposure-response relationships.

4. Evaluation of Individual Study Quality

After selecting studies for inclusion, the individual study quality is evaluated by considering the design, methods, conduct, and documentation of each study, but not the study results. This uniform approach aims to consider the strengths, limitations, and possible roles of chance, confounding, and other biases that may affect the interpretation of individual studies and the strength of inference from the results of the study. Particular study quality aspects, relevance, or limitations of some of these features do not necessarily define a less informative study or exclude a study from consideration in an ISA. As stated initially, the intent of the ISA is to provide a concise review, synthesis, and evaluation of the most policy-relevant science to serve as a scientific foundation for the review of the NAAQS, not extensive summaries of all human health and welfare effects studies for a pollutant. Of most importance for inclusion of a study is whether it provides useful qualitative or quantitative information on exposure-response relationships for effects associated with pollutant exposures at doses or concentrations relevant to ambient conditions that can inform decisions on whether to retain or revise the NAAQS.

In general, in assessing the scientific quality of studies on health and welfare effects, the following considerations are taken into account.

- Were study design, study groups, methods, data, and results clearly presented in relation to the study objectives to allow for study evaluation? Were limitations and any underlying assumptions of the design and other aspects of the study stated?
- Were the ecosystems, study site(s), study populations, subjects, or organism models adequately selected, and are they sufficiently well defined to allow for meaningful comparisons between study or exposure groups?
- Are the air quality, exposure, or dose metrics of adequate quality and are they sufficiently representative of or pertinent to ambient conditions?
- Are the health, ecological, or other welfare effect measurements meaningful, valid, and reliable?
- Were likely covariates or modifying factors adequately controlled or taken into account in the study design and statistical analysis?
- Do the analytical methods provide adequate sensitivity and precision to support conclusions?
- Were the statistical analyses appropriate, properly performed, and properly interpreted?

Additional study quality considerations specific to particular disciplines are discussed below.

a. Atmospheric Science and Exposure Assessment

Atmospheric science and exposure assessment studies that are considered for inclusion in the ISA focus on measurement of, behavior of, and exposure to ambient air pollution using quality-assured field, experimental, and/or modeling techniques. The most informative measurement-based studies will include detailed descriptive statistics for measurements taken at varying spatial and temporal scales. These studies will also include a clear and comprehensive description of measurement techniques and quality-control procedures used. Quality-control metrics (e.g., method detection limits) and quantitative relationships between and within pollutant measurements (e.g., regression slopes, intercepts, and fit statistics) should be provided when appropriate. Measurements that include contrasting conditions for various time periods (e.g., weekday/weekend, season), populations, regions, and categories (e.g., urban/rural) are particularly useful. The most informative modeling-based studies will incorporate appropriate chemistry, transport, dispersion, and/or exposure modeling techniques with a clear and comprehensive description of model evaluation procedures, metrics, and technique strengths and limitations. The ISA also may include analyses of data pertinent to characterizing air quality or exposure, such as emissions sources and ambient air pollutant concentrations. Sources of monitoring and modeling data should be clearly referenced and described to foster transparency and reproducibility of any analysis. In

general, atmospheric science studies and data analyses focusing on locations pertinent to the U.S. will have maximum value in informing review of the NAAQS.

Exposure measurement error, which refers to inaccuracies in the characterization of the exposures of study participants, can be an important contributor to uncertainty in air pollution epidemiologic study results. Exposure measurement error can influence observed epidemiologic associations between ambient pollutant concentrations and health outcomes by biasing effect estimates toward or away from the null and widening confidence intervals around those estimates (Zeger et al., 2000). Factors that could influence exposure estimates include, but are not limited to: choice of exposure metric, spatial variability of the pollutant concentration, nonambient sources of exposure, topography of the natural and built environment, meteorology, instrument errors, time-activity patterns, and differential infiltration of air pollutants into indoor environments. The influence of these factors on effect estimates also depends on epidemiologic study design. For example, when longitudinal studies depend on spatial contrasts in exposure estimates, it is important that the exposure estimates correspond in space to the population of interest. Likewise for time-series studies, the temporal variability of the exposure estimate must correspond temporally to the true exposures of the study population.

b. Epidemiology

In evaluating individual study quality for inference about health effects in epidemiologic studies, the U.S. EPA considers, in addition to the general quality considerations discussed previously, whether a given study: (1) presents information on associations with short- or long-term pollutant exposures at or near conditions relevant to ambient exposures; (2) addresses potential confounding, particularly by other pollutants; (3) assesses potential effect modifiers; (4) evaluates health endpoints and populations, groups, or lifestages not previously extensively researched; and (5) evaluates important methodological issues related to interpretation of the health evidence (e.g., lag or time period between exposure and effects, model specifications, thresholds).

In evaluating epidemiologic evidence, one important consideration is potential confounding. Confounding is "…a confusion of effects. Specifically, the apparent effect of the exposure of interest is distorted because the effect of an extraneous factor is mistaken for or mixed with the actual exposure effect (which may be null)" (Rothman and Greenland, 1998). A confounder is associated with both the exposure and the effect; for example, confounding can occur between correlated pollutants that are associated with the same effect. One approach to remove spurious associations due to possible confounders is to control for characteristics that may differ between exposed and

unexposed persons; this is frequently termed "adjustment." Scientific judgment is needed to evaluate likely sources and extent of confounding, together with consideration of how well the existing constellation of study designs, results, and analyses address the potential for erroneous inferences.

Several statistical methods are available to detect and control for potential confounders; however, none of these methods is completely satisfactory. Multivariable regression models constitute one tool for estimating the association between exposure and outcome after adjusting for characteristics of participants that might confound the results. Because much of the uncertainty in inferring causality may be due to potential confounding by copollutants, evaluation of copollutant confounding in individual studies is of particular importance. The use of copollutant regression models has been the prevailing approach for controlling for potential confounding by copollutants in air pollution health effects studies. Trying to determine whether an individual pollutant is independently associated with the health outcome of interest from copollutant regression models is made difficult by the possibility that one or more air pollutants is acting as a surrogate for an unmeasured or poorly measured pollutant or for a particular mixture of pollutants. In addition, pollutants may independently exert effects on the same system; for example, several pollutants may be associated with a respiratory effect through either the same or different modes of action. Despite these limitations, the use of copollutant models is still the prevailing approach employed in most air pollution epidemiologic studies and provides some insight into the potential for confounding or interaction among pollutants.

Confidence that unmeasured confounders are not producing the findings is increased when multiple studies are conducted in various settings using different subjects or exposures, each of which might eliminate another source of confounding from consideration. For example, multicity studies can provide insight on potential confounding through the use of a consistent method to analyze data from across locations with different concentrations of copollutants and other covariates. Intervention studies, because of their quasi-experimental nature, can be particularly useful in characterizing causation.

Another important consideration in the evaluation of epidemiologic studies is effect-measure modification, which occurs when the effect differs between subgroups or strata; for example, effect estimates that vary by age group or a potential risk factor. As stated by Rothman and Greenland (1998):

> "Effect-measure modification differs from confounding in several ways. The main difference is that, whereas confounding is a bias that the investigator hopes to prevent or remove from the effect estimate, effect-measure modification is a property of the effect under study …In

epidemiologic analysis one tries to eliminate confounding but one tries to detect and estimate effect-measure modification."

When a risk factor is a confounder, it is the true cause of the association observed between the exposure and the outcome; when a risk factor is an effect modifier, it changes the magnitude of the association between the exposure and the outcome in stratified analyses. For example, the presence of a pre-existing disease or indicator of low socioeconomic status (SES) (e.g., educational attainment, household income) may act as effect modifiers if they are associated with increased risk of effects related to air pollution exposure. It is often possible to stratify the relationship between health outcome and exposure by one or more of these potential effect modifiers. For variables that modify the association, effect estimates in each stratum will be different from one another and different from the overall estimate, indicating a different exposure-response relationship may exist in populations represented by these variables.

c. Controlled Human Exposure and Animal Toxicology

Controlled human exposure and animal toxicological studies experimentally evaluate the health effects of administered exposures in human volunteers and animal models under highly controlled laboratory conditions. Controlled human exposure studies are also referred to as human clinical studies. In controlled human exposure and animal toxicological experiments, investigators expose subjects or animals to known concentrations of air pollutants under carefully regulated environmental conditions and activity levels. In addition to the general quality considerations discussed previously, evaluation of controlled human exposure and animal toxicological studies includes assessing the design and methodology of each study with focus on (1) characterization of the intake dose, dosing regimen, and exposure route; (2) characterization of the pollutant(s); (3) sample size and statistical power to detect differences; and (4) control of other variables that could influence the occurrence of effects. The evaluation of study design generally includes consideration of factors that minimize bias in results, such as randomization, blinding, and allocation concealment of study subjects, investigators, and research staff, and unexplained loss of animals or withdrawal/exclusion of subjects. Additionally, studies must include appropriate control groups to allow for accurate interpretation of results relative to exposure. Emphasis is placed on studies that address concentration-dependent responses or time-course of responses and studies that investigate potentially at-risk lifestages or populations (e.g., older adults, groups with pre-existing disease).

Controlled human exposure or animal toxicological studies that approximate expected human exposures in terms of concentration, duration, and route of exposure are of particular interest. Relevant pollutant exposures are considered to be those generally within two orders of magnitude of recent ambient concentrations. This range in relevant exposures is to account for differences in dosimetry, toxicokinetics, and biological sensitivity of various species, strains, or potentially at-risk populations. Studies using higher concentration exposures or doses will be considered to the extent that they provide information relevant to understanding mode of action or mechanisms, inter-species variation, or at-risk human populations. In vitro studies may provide mechanistic insight for effects examined in vivo or in epidemiologic studies.

d. Ecological and Other Welfare Effects

Ecological effects considered in the ISAs typically include several of the topics given as examples by the Clean Air Act definition in Section 302(h) related to effects on welfare including soils, water, vegetation, animals, and wildlife. Additional topic areas that may be evaluated by an ISA include visibility, weather, and climate, as well as materials damage, economic values, and impacts to personal comfort and well-being. In evaluating studies that consider ecological and other welfare effects, in addition to assessing the general quality considerations discussed previously, emphasis is placed on studies that evaluate effects at or near ambient concentrations. Studies conducted in any country that contribute significantly to the general understanding of air pollutant effects may be evaluated for relevance to U.S. air quality considerations and inclusion in the ISA.

For ecological effects, studies at higher concentrations are used to evaluate ecological effects only when they are part of a range of concentrations that also included more ambient-relevant concentrations, or when they inform understanding of modes of action and illustrate the wide range of sensitivity to air pollutants across taxa or across biomes and ecoregions. In evaluating quantitative exposure-response relationships, emphasis is placed on findings from studies conducted in the U.S. and Canada as having ecological and climatic conditions most relevant for review of the NAAQS. The type of experimental approach used in the study (e.g., controlled laboratory exposure, growth chamber, open-top chamber, mesocosm, gradient, field study, etc.) is also evaluated when considering the applicability of the results to the review of criteria air pollutant effects.

In evaluating studies on climate and visibility, emphasis is placed on studies that use well-established measurement and modeling techniques, especially those that report uncertainty or compare results from an ensemble of techniques. Novel methods may also be informative in addressing knowledge gaps not well characterized by existing techniques. Relevant climate studies include those evaluating direct and indirect climate

impacts of criteria air pollutants at a global scale, while for visibility, studies conducted in the U.S. and Canada provide information more applicable for review of the NAAQS. In both cases, studies that evaluate effects by source sector or region, such as regional climate modeling studies, are particularly informative. Studies that report impacts of multiple PM components for visibility, and, for climate, multiple criteria pollutants are useful in evaluating interactions and the relative contributions of atmospheric constituents. For example, in evaluating the climate forcing effects of ozone (O_3), it is useful to understand the atmospheric chemistry involving CO and NO_X (the sum of nitric oxide and nitrogen dioxide) that affects atmospheric concentrations of O_3. Visibility preference and valuation studies that explicitly separate preferences for visibility from concerns about health risks of air pollution are particularly relevant in considering a welfare-based secondary NAAQS for pollutants that affect visibility.

5. Evaluation, Synthesis, and Integration of Evidence across Disciplines and Development of Scientific Conclusions and Causal Determinations

The U.S. EPA has developed an approach for integrating the scientific evidence gained from the array of studies discussed above in order to draw conclusions regarding the causal nature of ambient air pollutant-related health or welfare effects. Evidence from all disciplines is integrated to evaluate consistency and inconsistency in the pattern of effects as well as strengths and limitations of the evidence across disciplines. Part of this approach includes a framework for making determinations with regard to the existence of a causal relationship between the pollutant in ambient air and health or welfare effects (described in Section 5.b). This framework establishes uniform language concerning causality and brings specificity to the conclusions.

a. Evaluation, Synthesis, and Integration of Evidence across Disciplines

The ISA focuses on evaluation of the findings from the body of evidence across disciplines, drawing upon the results of all studies judged of adequate quality and relevance per the criteria described previously. Evidence from across scientific disciplines for related and similar health or welfare effects is evaluated, synthesized, and integrated to develop conclusions and causality determinations. This process includes evaluating strengths and weaknesses in the overall collection of studies across disciplines. Confidence in the collective body of evidence is based on evaluation of study design and quality. The roles of different types of evidence in drawing the conclusions varies by

pollutant or assessment, as does the availability of different types of evidence for causality determination. Consideration of human health effects are informed by controlled human exposure, epidemiologic, and toxicological studies. Evidence on ecological and other welfare effects may be drawn from a variety of experimental approaches (e.g., greenhouse, laboratory, field) and numerous disciplines (e.g., community ecology, biogeochemistry, paleontological/historical reconstructions). Other evidence, including mechanistic, toxicokinetics, and exposure assessment, may be highlighted if it is relevant to the evaluation of health and welfare effects and is of sufficient importance to affect the overall evaluation. Causal inference can be strengthened by integrating evidence across disciplines. A weak inference from one line of evidence can be addressed by other lines of evidence, and coherence of these lines of evidence can add support to a cause-effect interpretation of the association. Interpretation of the body of epidemiologic associations as evidence of causal relationships involves assessing the full evidence base with regard to elimination of alternative explanations for the association.

Evaluation and integration of evidence must also include consideration of uncertainty, which is inherent in scientific findings. "Uncertainty" can be defined as a deficit of knowledge to describe the existing state or future outcome with accuracy and precision (e.g., the lack of knowledge about the correct value for a specific measure or estimate). Uncertainty analysis may be qualitative or quantitative in nature. In many cases, the analysis is qualitative and can include professional judgment or inferences based on analogy with similar situations. Quantitative uncertainty analysis may include use of simple measures (e.g., ranges) and analytical techniques. Quantitative uncertainty analysis might progress to more complex measures and techniques, if needed for decision support. Various approaches to evaluating uncertainty include classical statistical methods, sensitivity analysis, or probabilistic uncertainty analysis, in order of increasing complexity and data requirements. However, data may not be available for all aspects of an assessment, and those data that are available may be of questionable or unknown quality. Ultimately, the assessment is based on a number of assumptions with varying degrees of uncertainty. While the ISA may include quantitative analysis approaches such as meta-regression in some situations, generally qualitative evaluation of uncertainties is used to assess the evidence from across studies.

Publication bias is another source of uncertainty that can impact the magnitude of health or welfare risk estimates. It is well understood that studies reporting non-null findings are more likely to be published than reports of null findings. Publication bias can result in overestimation of effect estimate sizes (Ioannidis, 2008). For example, effect estimates from single-city epidemiologic studies have been found to be generally larger than those from multicity studies. This is an indication of publication bias because null or negative

single-city results may be reported in multicity analyses but might not be published independently (Bell et al., 2005).

Health-specific considerations

Potential strengths and limitations of the body of studies can vary across disciplines and are evaluated during data synthesis and integration. Direct evidence of a relationship between pollutant exposures and human health effects may come from controlled human exposure studies. These studies can also provide important information on the biological plausibility of associations observed in epidemiologic studies and inform determinations of factors that may increase or decrease the risk of health effects in certain populations. In some instances, controlled human exposure studies can be used to characterize concentration-response relationships at pollutant concentrations relevant to ambient conditions. Controlled human exposures are typically conducted using a randomized crossover design, with subjects exposed both to the pollutant and a clean air control. In this way, subjects serve as their own experimental controls, effectively limiting the variance associated with potential inter-individual confounders. Limitations that must be considered in evaluating controlled human study findings include the generally small sample size and short exposure time used in experimental studies, and that severe health outcomes are not assessed. By experimental design, controlled human exposure studies are structured to evaluate physiological or biomolecular outcomes in response to exposure to a specific air pollutant and/or combination of pollutants. In addition, the study design generally precludes inclusion of subjects with serious health conditions or heightened risks of exposure, and therefore, the results often cannot be generalized to an entire population, which includes populations or lifestages at potentially increased risk of air pollutant-induced effects. Although some controlled human exposure studies have included health-compromised individuals, such as those with mild or moderate respiratory or cardiovascular disease, these individuals may also be relatively healthy and may not represent the most sensitive individuals in the population. Thus, observed effects in these studies may underestimate the response in certain populations. In addition, the study design is limited to exposures and endpoints that are not expected to result in severe health outcomes.

Epidemiologic studies provide important information on the associations between health effects and exposure of human populations to ambient air pollution. In epidemiologic or observational studies of humans, the investigator tends not to control exposures or intervene with the study population. Broadly, observational studies can describe associations between exposures and effects. These studies fall into several categories and include, for example, cross-sectional, prospective cohort, time-series, and panel studies. Each type of study has various strengths and limitations. Cross-sectional ecologic studies

use health outcome, exposure, and covariate data available at the community level (e.g., annual mortality rates and pollutant concentrations), but do not have individual-level data. Cross-sectional studies may have limited power to evaluate an extensive set of confounding factors because these studies examine between-subject or between-location comparisons. Prospective cohort studies include some data collected at the individual level, typically health outcome data, and in some cases, individual-level data on exposure and covariates are collected. Time-series and case-crossover studies are often used to evaluate the relationship between day-to-day changes in air pollution exposures and a specific health outcome at the population-level (i.e., mortality, hospital admissions, or emergency department visits). Panel studies include repeated measurements of health outcomes, such as respiratory symptoms or heart rate variability, at the individual level. "Natural experiments" offer the opportunity to investigate changes in health related to a change in exposure, such as closure of a pollution source.

When evaluating the collective body of epidemiologic studies, consideration of many study design factors and limitations must be taken into account to properly inform their interpretation. One key consideration is the evaluation of the potential independent contribution of the pollutant to a health outcome when it is a component of a complex air pollutant mixture. Reported effect estimates in epidemiologic studies may reflect (1) independent effects on health outcomes, (2) effects of the pollutant acting as an indicator of a copollutant or a complex ambient air pollution mixture, and (3) effects resulting from interactions between that pollutant and copollutants.

The third main type of health effects evidence, animal toxicological studies, provides information on the pollutant's biological action under controlled and monitored exposure circumstances. Taking into account biological differences among species, these studies contribute to our understanding of potential health effects, exposure-response relationships, and modes of action. Further, animal models can inform determinations of factors that may increase or decrease the risk of health effects in certain populations. These studies evaluate the effects of exposures to a variety of pollutants in a highly controlled laboratory setting and allow exploration of toxicological pathways or mechanisms by which a pollutant may cause effects. Understanding the biological mechanisms underlying various health outcomes can be crucial in establishing or negating causality. In the absence of human studies data, extensive, well-conducted animal toxicological studies can support determinations of causality, if the evidence base indicates that similar responses are expected in humans under ambient exposure conditions.

Interpretations of animal toxicological studies are affected by limitations associated with extrapolation between animal and human responses. The differences between humans

and other species have to be taken into consideration, including metabolism, hormonal regulation, breathing pattern, and differences in lung structure and anatomy. Also, in spite of a high degree of homology and the existence of a high percentage of orthologous genes across humans and rodents (particularly mice), extrapolation of molecular alterations at the gene or protein level is complicated by species-specific differences in transcriptional regulation and/or signaling. Given these differences, uncertainties are associated with quantitative extrapolations of observed pollutant-induced pathophysiological alterations between laboratory animals and humans, as those alterations are under the control of widely varying biochemical, endocrine, and neuronal factors.

Ecological- and welfare-specific considerations

For ecological effects assessment, both laboratory and field studies (including field experiments and observational studies) can provide useful data for causal determination. Because conditions can be controlled in laboratory studies, responses may be less variable and smaller effects may be easier to detect. However, the control conditions may limit the range of responses (e.g., animals may not be able to seek alternative food sources) or incompletely reflect pollutant bioavailability, so the responses under controlled conditions may not reflect responses that would occur in the natural environment. In addition, larger scale processes are difficult to reproduce in the laboratory.

Field observational studies measure biological changes in uncontrolled situations with high natural variability (in organismal genetics or in abiotic seasonal, climatic, or soil-related factors) and describe an association between a disturbance and an ecological effect. Field data can provide important information to assess multiple stressors or circumstances where site-specific factors significantly influence exposure. Field data are also often useful for analyzing pollutant effects at larger geographic scales and higher levels of biological organization. However, because conditions are not controlled, variability of the response is expected to be higher and may mask effects. Field surveys are most useful for linking stressors with effects when stressor and effect levels are measured concurrently. The presence of confounding factors can make it difficult to attribute observed effects to specific stressors.

Ecological impacts of pollutants are also evaluated in studies “intermediate” between the lower variability typically associated with laboratory exposures and high natural variability usually found in field studies. Some studies use environmental media collected from the field to examine the biological responses under controlled laboratory conditions. Other studies are experiments performed in the natural environment that control for some, but not all, of the environmental or genetic variability (e.g., mesocosm studies).

This type of study in manipulated natural environments can be considered a hybrid between a field experiment and laboratory study because some sources of response variation are removed through use of control conditions, while others are included to mimic natural variation. Such studies make it possible to observe community and/or ecosystem dynamics and provide strong evidence for causality when combined with findings of studies that have been made under more controlled conditions.

b. Considerations in Developing Scientific Conclusions and Causal Determinations

In its evaluation and integration of the scientific evidence on health or welfare effects of criteria pollutants, the U.S. EPA determines the weight of evidence in support of causation and characterizes the strength of any resulting causal classification. The U.S. EPA also evaluates the quantitative evidence and draws scientific conclusions, to the extent possible, regarding the concentration-response relationships and the loads to ecosystems, exposures, doses or concentrations, exposure duration, and pattern of exposures at which effects are observed.

Approaches to assessing the separate and combined lines of human health evidence (e.g., epidemiologic, controlled human exposure, and animal toxicological studies) have been formulated by a number of regulatory and science agencies, including the National Academy of Sciences (NAS) Institute of Medicine (IOM, 2008), the International Agency for Research on Cancer (IARC, 2006), the U.S. EPA (2005), and the Centers for Disease Control and Prevention (CDC, 2004)]. Causal inference criteria have also been described for ecological effects evidence (U.S. EPA, 1998; Fox, 1991). These formalized approaches offer guidance for assessing causality. The frameworks of each are similar in nature, although adapted to different purposes, and have proven effective in providing a uniform structure and language for causal determinations.

The 1964 Surgeon General's report on tobacco smoking defined "cause" as a "significant, effectual relationship between an agent and an associated disorder or disease in the host" (HEW, 1964). More generally, a cause is defined as an agent that brings about an effect or a result. An association is the statistical relationship among variables, but alone, it is insufficient proof of a causal relationship between an exposure and a health outcome. Unlike an association, a causal claim supports the creation of counterfactual claims; that is, a claim about what the world would have been like under different or changed circumstances (IOM, 2008).

Many of the health and welfare outcomes reported in studies have complex etiologies. Diseases such as asthma, coronary heart disease, or cancer are typically initiated by

multiple agents. Outcomes depend on a variety of factors, such as age, genetic background, nutritional status, immune competence, and social factors (IOM, 2008; Gee and Payne-Sturges, 2004). Effects on ecosystems are also often multifactorial with a complex web of causation. Further, exposure to a combination of agents could cause synergistic or antagonistic effects. Thus, the observed risk may represent the net effect of many actions and counteractions.

To aid judgment, various "aspects"[1] of causality have been discussed by many philosophers and scientists. The 1964 Surgeon General's report on tobacco smoking discussed criteria for the evaluation of epidemiologic studies, focusing on consistency, strength, specificity, temporal relationship, and coherence (HEW, 1964). Sir Austin Bradford Hill (Hill, 1965) articulated aspects of causality in epidemiology and public health that have been widely used (IOM, 2008; IARC, 2006; U.S. EPA, 2005; CDC, 2004). These aspects (Hill, 1965) have been modified (Table I) for use in causal determinations specific to health and welfare effects for pollutant exposures (U.S. EPA, 2009).[2] Although these aspects provide a framework for assessing the evidence, they do not lend themselves to being considered in terms of simple formulas or fixed rules of evidence leading to conclusions about causality (Hill, 1965). For example, one cannot simply count the number of studies reporting statistically significant results or statistically nonsignificant results and reach credible conclusions about the relative weight of evidence and the likelihood of causality. Rather, these aspects provide a framework for systematic appraisal of the body of evidence, informed by peer and public comment and advice, which includes weighing alternative views on controversial issues. In addition, it is important to note that the aspects in Table I cannot be used as a strict checklist, but rather to determine the weight of evidence for inferring causality. In particular, not meeting one or more of the principles does not automatically preclude a determination of causality [see discussion in (CDC, 2004)].

[1] The "aspects" described by Sir Austin Bradford Hill (Hill, 1965) have become, in the subsequent literature, more commonly described as "criteria." The original term "aspects" is used here to avoid confusion with "criteria" as it is used, with different meaning, in the Clean Air Act.

[2] The Hill aspects were developed for interpretation of epidemiologic results. They have been modified here for use with a broader array of data (i.e., epidemiologic, controlled human exposure, ecological, and animal toxicological studies as well as in vitro data) and to be more consistent with the EPA Guidelines for Carcinogen Risk Assessment.

Table I Aspects to aid in judging causality.

Aspect	Description
Consistency	An inference of causality is strengthened when a pattern of elevated risks is observed across several independent studies. The reproducibility of findings constitutes one of the strongest arguments for causality. Statistical significance is not the sole criterion by which the presence or absence of an effect is determined. If there are discordant results among investigations, possible reasons such as differences in exposure, confounding factors, and the power of the study are considered.
Coherence	An inference of causality from one line of evidence (e.g., epidemiologic, controlled human exposure, animal, or ecological studies) may be strengthened by other lines of evidence that support a cause-and-effect interpretation of the association. There may be coherence in demonstrating effects from evidence across various fields and/or across multiple study designs or related health endpoints within one scientific line of evidence. For example, evidence on welfare effects may be drawn from a variety of experimental approaches (e.g., greenhouse, laboratory, and field) and subdisciplines of ecology (e.g., community ecology, biogeochemistry, and paleontological/historical reconstructions).
Biological plausibility	An inference of causality is strengthened by results from experimental studies or other sources demonstrating biologically plausible mechanisms. A proposed mechanism, which is based on experimental evidence and which links exposure to an agent to a given effect, is an important source of support for causality.
Biological gradient (exposure-response relationship)	A well-characterized exposure-response relationship (e.g., increasing effects associated with greater exposure) strongly suggests cause and effect, especially when such relationships are also observed for duration of exposure (e.g., increasing effects observed following longer exposure times).
Strength of the observed association	The finding of large, precise risks increases confidence that the association is not likely due to chance, bias, or other factors. However, it is noted that a small magnitude in an effect estimate may or may not represent a substantial effect in a population.
Experimental evidence	Strong evidence for causality can be provided through “natural experiments” when a change in exposure is found to result in a change in occurrence or frequency of health or welfare effects.
Temporality of the observed association	Evidence of a temporal sequence between the introduction of an agent and appearance of the effect constitutes another argument in favor of causality.
Specificity of the observed association	Evidence linking a specific outcome to an exposure can provide a strong argument for causation. However, it must be recognized that rarely, if ever, does exposure to a pollutant invariably predict the occurrence of an outcome, and that a given outcome may have multiple causes.
Analogy	Structure activity relationships and information on the agent’s structural analogs can provide insight into whether an association is causal. Similarly, information on mode of action for a chemical, as one of many structural analogs, can inform decisions regarding likely causality.

Consistency of findings across studies is informed by the repeated observation of effects or associations across multiple independent studies. Further strength is provided by reproducibility of findings in different populations under different circumstances.

However, discordant results among independent investigations may be explained by differences in study methods, random errors, exposure, confounding factors, or study power, and thus may not be used to rule out a causal connection.

In evaluating the consistency of findings across studies, the U.S. EPA emphasizes the importance of examining the pattern of results across various studies and does not focus solely on statistical significance or the magnitude of the direction of the association as criteria of study reliability. Statistical significance is influenced by a variety of factors including, but not limited to, the size of the study, exposure and outcome measurement error, and statistical model specifications. Statistical significance may be informative; however, it is just one of the means of evaluating confidence in the observed relationship and assessing the probability of chance as an explanation. Other indicators of reliability such as the consistency and coherence of a body of studies as well as other confirming data may be used to justify reliance on the results of a body of epidemiologic studies, even if results in individual studies lack statistical significance. Traditionally, statistical significance is used to a larger extent to evaluate the findings of controlled human exposure and animal toxicology studies. Understanding that statistical inferences may result in both false positives and false negatives, consideration is given to both trends in data and reproducibility of results. Thus, in drawing judgments regarding causality, the U.S. EPA emphasizes statistically significant findings from experimental studies, but does not limit its focus or consideration to statistically significant results in epidemiologic studies.

In evaluating the strength of the observed association, the U.S. EPA considers both the magnitude and statistical precision (i.e., width of confidence interval) of the association in epidemiologic studies. In a large study that accounts for several potential confounding factors, a strong association can serve to increase confidence that a finding is not due to a weak unmeasured confounder, chance, or other biases. However, in a study that accounts for several potential confounding factors and other sources of bias, a weak association does not rule out a causal connection. The health effects evaluated in the ISAs tend to have multiple risk factors that likely vary in strength of effect, and the magnitude of effect of air pollution exposure will depend on the prevalence of other risk factors in the study population. Further, a small effect size can be important from a public health impact perspective. The air pollution-related change in a health effect observed in a study can represent a shift in the distribution of responses in the study population and potentially an increase in the proportion of individuals with clinically important effects.

In making judgments regarding causality, the biological plausibility of effects resulting from air pollutant exposure is considered. Experimental results from in vivo studies involving animal models and humans, as well as from in vitro studies when appropriate,

may be used to establish biological plausibility and to interpret other lines of evidence (e.g., health effects from epidemiologic studies). Biological plausibility is often provided from understanding the mode of action by which exposure to a pollutant leads to health effects. This understanding may encompass several different levels of biological organization including, but not limited to, molecular and cellular events in the pathways leading to disease. While a complete understanding of the mode of action is not considered necessary for making causal determinations within the ISA, biological plausibility plays a key role.

c. Framework for Causal Determinations

In the ISA, the U.S. EPA assesses the body of relevant literature, building upon evidence available during previous NAAQS reviews, to draw conclusions on the causal relationships between relevant pollutant exposures and health or environmental effects. ISAs use a five-level hierarchy that classifies the weight of evidence for causation.[1] This weight-of-evidence evaluation is based on the integration of findings from various lines of evidence from across health and environmental effect disciplines that are integrated into a qualitative statement about the overall weight of the evidence and causality. The five descriptors for causal determination are described in Table II.

Table II Weight of evidence for causal determination.

	Health Effects	Ecological and Other Welfare Effects
Causal relationship	Evidence is sufficient to conclude that there is a causal relationship with relevant pollutant exposures (e.g., doses or exposures generally within one to two orders of magnitude of recent concentrations). That is, the pollutant has been shown to result in health effects in studies in which chance, confounding, and other biases could be ruled out with reasonable confidence. For example: (1) controlled human exposure studies that demonstrate consistent effects, or (2) observational studies that cannot be explained by plausible alternatives or that are supported by other lines of evidence (e.g., animal studies or mode of action information). Generally, the determination is based on multiple high-quality studies conducted by multiple research groups.	Evidence is sufficient to conclude that there is a causal relationship with relevant pollutant exposures. That is, the pollutant has been shown to result in effects in studies in which chance, confounding, and other biases could be ruled out with reasonable confidence. Controlled exposure studies (laboratory or small- to medium-scale field studies) provide the strongest evidence for causality, but the scope of inference may be limited. Generally, the determination is based on multiple studies conducted by multiple research groups, and evidence that is considered sufficient to infer a causal relationship is usually obtained from the joint consideration of many lines of evidence that reinforce each other.

[1] The CDC and IOM frameworks use a four-level hierarchy for the strength of the evidence. A five-level hierarchy is used here to be consistent with the five-level hierarchy used in the U.S. *EPA Guidelines for Carcinogen Risk Assessment* and to provide a more nuanced set of categories.

Table II (Continued): Weight of evidence for causal determination.

	Health Effects	Ecological and Other Welfare Effects
Likely to be a causal relationship	Evidence is sufficient to conclude that a causal relationship is likely to exist with relevant pollutant exposures. That is, the pollutant has been shown to result in health effects in studies where results are not explained by chance, confounding, and other biases, but uncertainties remain in the evidence overall. For example: (1) observational studies show an association, but copollutant exposures are difficult to address and/or other lines of evidence (controlled human exposure, animal, or mode of action information) are limited or inconsistent, or (2) animal toxicological evidence from multiple studies from different laboratories demonstrate effects, but limited or no human data are available. Generally, the determination is based on multiple high-quality studies.	Evidence is sufficient to conclude that there is a likely causal association with relevant pollutant exposures. That is, an association has been observed between the pollutant and the outcome in studies in which chance, confounding, and other biases are minimized but uncertainties remain. For example, field studies show a relationship, but suspected interacting factors cannot be controlled, and other lines of evidence are limited or inconsistent. Generally, the determination is based on multiple studies by multiple research groups.
Suggestive of, but not sufficient to infer, a causal relationship	Evidence is suggestive of a causal relationship with relevant pollutant exposures but is limited, and chance, confounding, and other biases cannot be ruled out. For example: (1) when the body of evidence is relatively small, at least one high-quality epidemiologic study shows an association with a given health outcome and/or at least one high-quality toxicological study shows effects relevant to humans in animal species, or (2) when the body of evidence is relatively large, evidence from studies of varying quality is generally supportive but not entirely consistent, and there may be coherence across lines of evidence (e.g., animal studies or mode of action information) to support the determination.	Evidence is suggestive of a causal relationship with relevant pollutant exposures, but chance, confounding, and other biases cannot be ruled out. For example, at least one high-quality study shows an effect, but the results of other studies are inconsistent.
Inadequate to infer a causal relationship	Evidence is inadequate to determine that a causal relationship exists with relevant pollutant exposures. The available studies are of insufficient quantity, quality, consistency, or statistical power to permit a conclusion regarding the presence or absence of an effect.	Evidence is inadequate to determine that a causal relationship exists with relevant pollutant exposures. The available studies are of insufficient quality, consistency, or statistical power to permit a conclusion regarding the presence or absence of an effect.
Not likely to be a causal relationship	Evidence indicates there is no causal relationship with relevant pollutant exposures. Several adequate studies, covering the full range of levels of exposure that human beings are known to encounter and considering at-risk populations and lifestages, are mutually consistent in not showing an effect at any level of exposure.	Evidence indicates there is no causal relationship with relevant pollutant exposures. Several adequate studies examining relationships with relevant exposures are consistent in failing to show an effect at any level of exposure.

This standardized language was drawn from sources across the federal government and wider scientific community, especially the U.S. EPA *Guidelines for Carcinogen Risk Assessment* (U.S. EPA, 2005), U.S. Surgeon General's report, *The Health Consequences of Smoking* (CDC, 2004), and NAS IOM document, *Improving the Presumptive Disability Decision-Making Process for Veterans* (IOM, 2008), a comprehensive report on evaluating causality.

This framework:

- describes the kinds of scientific evidence used in making determinations on causal relationships between exposure and health or welfare effects,
- summarizes the key aspects of the evaluation of evidence necessary to reach a conclusion about the existence of a causal relationship,
- identifies issues and approaches related to uncertainty, and
- classifies and characterizes the weight of evidence in support of a general causal determination.

Determination of causality involves evaluating and integrating evidence for different types of health or welfare effects associated with short- and long-term exposure periods. In drawing conclusions regarding causality, evidence is evaluated for major outcome categories or groups of related endpoints (e.g., respiratory effects, vegetation growth), integrating evidence from across disciplines, and evaluating the coherence of evidence across a spectrum of related endpoints. In discussing the causal determination, the U.S. EPA characterizes the evidence on which the judgment is based, including strength of evidence for individual endpoints within the outcome category or group of related endpoints.

In drawing judgments regarding causality for the criteria air pollutants, the ISA focuses on evidence of effects in the range of relevant pollutant exposures or doses and not on determination of causality at any dose. Emphasis is placed on evidence of effects at doses (e.g., blood Pb concentration) or exposures (e.g., air concentrations) that are relevant to, or somewhat above, those currently experienced by the population. The extent to which studies of higher concentrations are considered varies by pollutant and major outcome category, but generally includes those with doses or exposures in the range of one to two orders of magnitude above current or ambient conditions to account for intra-species variability and toxicokinetic or toxicodynamic differences between experimental animals and humans. Studies that use higher doses or exposures may also be considered to the extent that they provide useful information to inform understanding of mode of action, inter-species differences, or factors that may increase risk of effects for a population and if biological mechanisms have not been demonstrated to differ based on exposure concentration. Thus, a causality determination is based on weight-of-evidence evaluation for health or welfare effects, focusing on the evidence from exposures or doses generally ranging from recent ambient concentrations to one or two orders of magnitude above recent ambient concentrations.

In addition, the U.S. EPA evaluates evidence relevant to understanding the quantitative relationships between pollutant exposures and health or welfare effects. This includes evaluating the form of concentration-response or dose-response relationships and, to the

extent possible, drawing conclusions on the concentrations at which effects are observed. The ISA also draws scientific conclusions regarding important exposure conditions for effects and populations and lifestages that may be at greater risk for effects, as described in the following two sections on public health and public welfare impacts.

6. Public Health Impact

Once a determination is made regarding the causality of relationship between the pollutant and outcome category, the public health impact of exposure to the pollutant is evaluated. Important questions regarding the public health impact include:

- What populations and lifestages appear to be differentially affected (i.e., at greater or less risk of experiencing effects)?
- What exposure conditions (dose or exposure, duration, and pattern) are important?
- What is the severity of the effect (e.g., clinical relevance)?
- What is the concentration-response, exposure-response, or dose-response relationship in the human population?
- What is the interrelationship between incidence and severity of effect?

To address these questions, the entirety of quantitative evidence is evaluated to characterize pollutant concentrations and exposure durations at which effects were observed for exposed populations, including populations and lifestages potentially at increased risk. To accomplish this, evidence is considered from multiple and diverse types of studies, and a study or set of studies that best approximates the concentration-response relationships between health outcomes and the pollutant may be identified. Controlled human exposure studies provide the most direct and quantifiable exposure-response data on the human health effects of pollutant exposures, although they tend to examine potential at-risk populations and lifestages to a limited extent and tend to have small sample sizes for between-group comparisons. To the extent available, the ISA evaluates results from epidemiologic studies that characterize the shape of the relationship between a pollutant and a health outcome. Animal data may also inform evaluation of concentration-response relationships, particularly relative to modes of action and characteristics of at-risk populations.

a. Approach to Identifying, Evaluating, and Characterizing At-Risk Factors

A critical part of assessing the public health impact of an air pollutant is the identification, evaluation, and characterization of populations potentially at greater risk of

an air pollutant-related health effect. Under the Clean Air Act, the primary NAAQS are intended to protect public health with an adequate margin of safety. In doing so, protection is provided for both the population as a whole and those groups potentially at increased risk for health effects from exposure to a criteria air pollutant. To inform decisions on the NAAQS, the ISA evaluates the currently available information regarding those factors (e.g., lifestage, pre-existing disease) that contribute to portions of the population being at greater risk for an air pollutant-related health effect.

Studies often use a variety of terms to classify factors and subsequently populations that may be at increased risk of an air pollutant-related health effect, including "susceptible," "vulnerable," "sensitive," and "at-risk," with recent literature introducing the term "response-modifying factor" (Vinikoor-Imler et al., 2014; Sacks et al., 2011; U.S. EPA, 2010, 2009). The inconsistency in the definitions for each of these terms across the scientific literature has shifted the focus away from answering the key questions: Which populations are at increased risk and what evidence forms the basis of this conclusion (Vinikoor-Imler et al., 2014)? Due to the lack of a consensus on terminology in the scientific community, the term "susceptible populations" was used in previous reviews and ISAs (Sacks et al., 2011; U.S. EPA, 2010, 2009) to encompass these various factors. However, it was recognized that even using the term "susceptible populations" was problematic because it often refers to populations at increased risk specifically due to biological or intrinsic factors such as pre-existing disease or lifestage. As such, starting with the ISA for Ozone and Related Photochemical Oxidants (U.S. EPA, 2013), the terminology "at-risk" was introduced to define populations and lifestages potentially at increased risk of an air pollutant-related health effect. In assessing the overall public health impact of an air pollutant, the ISA focuses on identifying, evaluating, and characterizing "at-risk" factors to address the main question of what populations and lifestages are at increased risk of an air pollutant-related health effect. Each "at-risk" factor is evaluated with a focus on identifying whether the factor contributes to a population at increased risk of an air pollutant-related health effect. Some factors may lead to a reduction in risk, and these are acknowledged during the evaluation process. However, for the purposes of identifying those populations or lifestages at increased risk to inform decisions on the NAAQS, the focus of this ISA is on characterizing those factors that may increase risk.

A population or lifestage may be at increased risk for various reasons, which can generally be divided into four broad categories. The first category of factors often is referred to as intrinsic. Intrinsic factors can increase risk for an effect through a biological mechanism and include genetic or developmental factors, race, sex, lifestage, or the presence of pre-existing diseases. For example, people in this category would have a steeper concentration-risk relationship and a greater or more severe effect at a given

pollutant concentration compared to those not in the category. The second category often is referred to as extrinsic or nonbiological. These factors include SES (e.g., educational attainment, income, access to healthcare), activity pattern, and exercise level. The third category includes factors that can increase risk by increasing internal dose at a given exposure concentration. Individuals in this category could have a greater dose of delivered pollutant because of breathing patterns and could include children who are typically more active outdoors. In addition, some groups could have greater exposure (concentration × time) regardless of the delivered dose, such as outdoor workers. The final category encompasses factors that may increase risk for experiencing a greater exposure based on exposure to a higher concentration. For example, populations that live near roadways could be exposed to higher pollutant concentrations. Some factors described above are multifaceted and may influence the risk of an air pollutant-related health effect through a combination of ways (e.g., SES). Additionally, it is recognized that some portions of the population or lifestages may be at increased risk of an air pollutant-related health effect because they experience insults from a combination of factors. The emphasis is to identify and understand the factors that potentially increase the risk of air pollutant-related health effects, regardless of whether the increased risk is due to intrinsic factors, extrinsic factors, increased dose/exposure, or a combination due to the often interconnectedness of factors.

To identify at-risk factors that potentially lead to some portions of the population being at increased risk of air pollution-related health effects, the evidence is systematically evaluated across relevant scientific disciplines (i.e., exposure sciences, dosimetry, toxicology, epidemiology). The evaluation process consists of evaluating studies that conduct stratified analyses (i.e., epidemiologic, controlled human exposure) to compare populations or lifestages exposed to similar air pollutant concentrations within the same study design. Experimental studies also provide an important line of evidence in evaluating factors that can lead to increased risk of an air pollutant-related health effect. Specifically, toxicological studies conducted using animal models of disease and controlled human exposure studies that examine individuals with underlying disease or genetic polymorphisms can provide coherence with the health effects observed in epidemiologic studies as well as an understanding of biological plausibility. The potential increased risk of an air pollutant-related health effect may also be determined from studies that examine factors that result in differential air pollutant exposures. The characterization of each at-risk factor consists of evaluating the evidence across scientific disciplines and assessing the overall confidence that a specific factor may result in a population or lifestage being at increased risk of an air pollutant-related health effect. The categories considered for evaluating the potential increased risk of an air pollutant-related health effect are “adequate evidence,” “suggestive evidence,” “inadequate evidence,” and “evidence of no effect.” They are described in more detail in Table III.

Table III Characterization of evidence for potential at-risk factors.

Classification	Health Effects
Adequate evidence	There is substantial, consistent evidence within a discipline to conclude that a factor results in a population or lifestage being at increased risk of air pollutant-related health effect(s) relative to some reference population or lifestage. Where applicable, this evidence includes coherence across disciplines. Evidence includes multiple high-quality studies.
Suggestive evidence	The collective evidence suggests that a factor results in a population or lifestage being at increased risk of air pollutant-related health effect(s) relative to some reference population or lifestage, but the evidence is limited due to some inconsistency within a discipline or, where applicable, a lack of coherence across disciplines.
Inadequate evidence	The collective evidence is inadequate to determine whether a factor results in a population or lifestage being at increased risk of air pollutant-related health effect(s) relative to some reference population or lifestage. The available studies are of insufficient quantity, quality, consistency, and/or statistical power to permit a conclusion to be drawn.
Evidence of no effect	There is substantial, consistent evidence within a discipline to conclude that a factor does not result in a population or lifestage being at increased risk of air pollutant-related health effect(s) relative to some reference population or lifestage. Where applicable, the evidence includes coherence across disciplines. Evidence includes multiple high-quality studies.

b. Evaluating Adversity of Human Health Effects

In evaluating health evidence, a number of factors can be considered in delineating between adverse and nonadverse health effects resulting from exposure to air pollution. Some health outcomes, such as hospitalization for respiratory or cardiovascular diseases, are clearly adverse. It is more difficult to determine the extent of change that constitutes adversity in more subtle health measures. These more subtle health effects include a wide variety of responses, such as alterations in markers of inflammation or oxidative stress, changes in pulmonary function or heart rate variability, or alterations in neurocognitive function measures. The challenge is to determine the magnitude of change in these measures when there is no clear point at which a change becomes adverse. The extent to which a change in health measure constitutes an adverse health effect may vary between populations and lifestages. Some changes that may not be considered adverse in healthy individuals would be potentially adverse in more at-risk individuals.

Professional scientific societies may evaluate the magnitude of change in an outcome or event that is considered adverse. For example, in an official statement titled *What Constitutes an Adverse Health Effect of Air Pollution?* (ATS, 2000), the American Thoracic Society described transient decrements in lung function as adverse when accompanied by clinical symptoms. Additionally, an air pollution-induced shift in the

population distribution of a given risk factor for a health outcome was viewed as adverse, even though it may not increase the risk of any one individual to an unacceptable level. For example, a population with asthma could have a distribution of lung function such that no identifiable individual has a level associated with significant impairment. Exposure to air pollution could shift the distribution such that no identifiable individual experiences clinically relevant effects. This shift toward decreased lung function, however, could be considered adverse because individuals within the population would have diminished reserve function and therefore would be at increased risk to further environmental insult. The committee also observed that elevations of biomarkers, such as cell number and types, cytokines, and reactive oxygen species, may signal risk for ongoing injury and clinical effects or may simply indicate transient responses that can provide insights into mechanisms of injury, thus illustrating the lack of clear boundaries that separate adverse from nonadverse effects.

The more subtle health outcomes may be connected mechanistically to health events that are clearly adverse. For example, air pollution may affect markers of transient myocardial ischemia such as ST-segment abnormalities or onset of exertional angina. These effects may not be apparent to the individual, yet may still increase the risk of a number of cardiac events, including myocardial infarction and sudden death. Thus, small changes in physiological measures may not appear to be clearly adverse when considered alone, but may be a part of a coherent and biologically plausible chain of related health outcomes that range up to responses that are very clearly adverse, such as hospitalization or mortality.

c. Concentration-Response Relationships

An important consideration in characterizing the public health impacts associated with exposure to a pollutant is whether the concentration-response relationship is linear across the range of concentrations or if nonlinear relationships exist along any part of this range. The shape of the concentration-response curve at and below the level of the current standards is of particular interest. Various sources of variability and uncertainty, such as low data density in the lower concentration range, possible influence of exposure measurement error, and variability among individuals with respect to air pollution health effects, tend to smooth and “linearize” the concentration-response function and thus can obscure the existence of a threshold or nonlinear relationship. Because individual thresholds vary from person-to-person due to individual differences such as genetic differences or pre-existing disease conditions (and even can vary from one time to another for a given person), it can be difficult to demonstrate that a threshold exists in a population study. These sources of variability and uncertainty may explain why the

available human data at ambient concentrations for some environmental pollutants (e.g., PM, O_3, Pb, environmental tobacco smoke, radiation) do not exhibit population-level thresholds for cancer or noncancer health effects, even though likely mechanisms include nonlinear processes for some key events.

7. Public Welfare Impact

Once a determination is made regarding the causality of relationships between the pollutant and outcome category, important questions regarding the public welfare impact include:

- What endpoints or services appear to be differentially affected (i.e., at greater or less risk of experiencing effects)? What elements of the ecosystem (e.g., types, regions, taxonomic groups, populations, functions, etc.) appear to be affected, or are more sensitive to effects? Are there differences between locations or materials in welfare effects responses, such as impaired visibility or materials damage?
- What is concluded from the evidence with regard to other types of welfare effects?
- Under what exposure conditions (amount deposited or concentration, duration, and pattern) are effects seen?
- What is the shape of the concentration-response, exposure-response, or dose-response relationship?

To address these questions, the entirety of quantitative evidence is evaluated to characterize pollutant concentrations and exposure durations at which effects were observed. To accomplish this, evidence is considered from multiple and diverse types of studies, and a study or set of studies that best approximates the concentration-response relationships between welfare outcomes and the pollutant may be identified. Controlled experimental studies provide the most direct and quantifiable exposure-response data on the effects of pollutant exposures. To the extent available, the ISA also evaluates results from less controlled field studies that characterize the shape of the relationship between a pollutant and an outcome. Other types of data may also inform evaluation of concentration-response relationships, particularly relative to modes of action and characteristics of at-risk ecosystems.

a. Evaluating Adversity of Ecological and Other Welfare Effects

The final step in assessing the public welfare impact of an air pollutant is the evaluation of the level considered to be adverse. A secondary standard, as defined in Section 109(b)(2) of the CAA must "specify a level of air quality the attainment and

maintenance of which, in the judgment of the Administrator, based on such criteria, is requisite to protect the public welfare from any known or anticipated adverse effects associated with the presence of such air pollutant in the ambient air." In setting standards that are "requisite" to protect public health and welfare, as provided in Section 109(b), the U.S. EPA's task is to establish standards that are neither more nor less stringent than necessary for these purposes.

Adversity of ecological effects can be understood in terms ranging in biological level of organization from the cellular level to the individual organism and to the population, community, and ecosystem levels. In the context of ecology, a population is a group of individuals of the same species, and a community is an assemblage of populations of different species that inhabit an area and interact with one another. An ecosystem is the interactive system formed from all living organisms and their abiotic (physical and chemical) environment within a given area (IPCC, 2007). The boundaries of what could be called an ecosystem are somewhat arbitrary, depending on the focus of interest or study. Thus, the extent of an ecosystem may range from very small spatial scales to, ultimately, the entire Earth (IPCC, 2007).

Effects on an individual organism are generally not considered to be adverse to public welfare. However if effects occur to enough individuals within a population, then communities and ecosystems may be disrupted. Changes to populations, communities, and ecosystems can in turn result in an alteration of ecosystem processes. Ecosystem processes are defined as the metabolic functions of ecosystems, including energy flow, elemental cycling, and the production, consumption, and decomposition of organic matter (U.S. EPA, 2002). Growth, reproduction, and mortality are species-level endpoints that may be clearly linked to community and ecosystem effects and are considered to be adverse when negatively affected. Other endpoints, such as changes in behavior and physiological stress, can decrease ecological fitness of an organism but are harder to link unequivocally to effects at the population, community, and ecosystem level. Support for consideration of adversity beyond the species level by making explicit the linkages between stress-related effects at the species and effects at the ecosystem level is found in *A Framework for Assessing and Reporting on Ecological Condition: an SAB report* (U.S. EPA, 2002). Additionally, the National Acid Precipitation Assessment Program (NAPAP, 1991) uses the following working definition of "adverse ecological effects" in the preparation of reports to Congress mandated by the Clean Air Act: "any injury (i.e., loss of chemical or physical quality or viability) to any ecological or ecosystem component, up to and including the regional level, over both long and short terms."

Beyond species-level impacts, consideration of ecosystem services allows for evaluation of how pollutant exposure may adversely impact species or processes of particular

economic or cultural importance to humans. On a broader scale, ecosystem services may provide indicators for ecological impacts. Ecosystem services are the benefits that people obtain from ecosystems (UNEP, 2003). According to the *Millennium Ecosystem Assessment*, ecosystem services include "provisioning services such as food and water; regulating services such as regulation of floods, drought, land degradation, and disease; supporting services such as soil formation and nutrient cycling; and cultural services such as recreational, spiritual, religious, and other nonmaterial benefits" (UNEP, 2003). For example, a more subtle ecological effect of pollution exposure may result in a clearly adverse impact on ecosystem services if it results in a population decline in a species that is recreationally or culturally important.

A consideration in evaluating adversity of climate-related effects is that criteria air pollutants have both direct and indirect effects on radiative forcing. For example, CO has a relatively small direct forcing effect, but it influences the concentrations of other atmospheric species, such as O_3 and methane, which are important contributors to climate forcing. PM has both direct and indirect effects. For example, black carbon and sulfate contribute directly to warming and cooling, respectively, while aerosols are involved in cloud formation which affects climate indirectly. Thus, it is crucial to consider the role of multiple pollutants together in evaluating the climate impact of criteria pollutants. Although climate effects of criteria air pollutants impact terrestrial and aquatic environments in diverse ways over multiple time scales, their effect on temperature is the main metric of adversity, with some consideration of proximate effects such as precipitation and relatively rapid feedbacks impacting the composition of the troposphere. Downstream effects such as land use changes are more difficult to link back to changes in concentrations of individual pollutants regulated under the NAAQS. The relative adversity of U.S. versus global emissions and concentrations is informed by regional climate modeling studies, including consideration of uncertainty and spatial and temporal variability.

The adversity of visibility impacts may be expressed in terms of psychological stress, such as impairment of aesthetic quality or enjoyment of the environment, or in monetary terms, such as willingness to pay to improve air quality. Understanding the relationship between pollutant concentration and perception of visibility, including distinguishing between concerns about health risks due to air pollution and perceived visibility impairment, can be crucial in evaluating the level of protection provided by a welfare-based secondary NAAQS when impacts on visibility are among the welfare effects that are potentially relevant for a pollutant.

Adversity of materials damage is evaluated considering the impact to human and economic well-being. Physical damage and soiling impair aesthetic qualities and function

of materials. Additionally, damage to property and cultural heritage sites due to pollutant deposition may be considered adverse.

b. Quantitative Relationships: Effects on Welfare

Evaluations of causality generally consider the probability of quantitative changes in welfare effects in response to exposure. A challenge to the quantification of exposure-response relationships for ecological effects is the great regional and local spatial variability, as well as temporal variability, in ecosystems. Thus, exposure-response relationships are often determined for a specific ecological system and scale, rather than at the national or even regional scale. Quantitative relationships, therefore, are estimated site by site and may differ greatly between ecosystems.

References

ATS (American Thoracic Society). (2000). What constitutes an adverse health effect of air pollution? Am J Respir Crit Care Med 161: 665-673. http://dx.doi.org/10.1164/ajrccm.161.2.ats4-00

Bell, ML; Dominici, F; Samet, JM. (2005). A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study. Epidemiology 16: 436-445. http://dx.doi.org/10.1097/01.ede.0000165817.40152.85

CAA (Clean Air Act). (1990a). Clean Air Act, as amended by Pub. L. No. 101-549, section 108: Air quality criteria and control techniques, 42 USC 7408. http://www.law.cornell.edu/uscode/text/42/7408

CAA (Clean Air Act). (1990b). Clean Air Act, as amended by Pub. L. No. 101-549, section 109: National primary and secondary ambient air quality standards, 42 USC 7409. http://www.epa.gov/air/caa/title1.html#ia

CDC (Centers for Disease Control and Prevention). (2004). The health consequences of smoking: A report of the Surgeon General. Washington, DC: U.S. Department of Health and Human Services. http://www.surgeongeneral.gov/library/smokingconsequences/

Fox, GA. (1991). Practical causal inference for ecoepidemiologists. J Toxicol Environ Health A 33: 359-373. http://dx.doi.org/10.1080/15287399109531535

Gee, GC; Payne-Sturges, DC. (2004). Environmental health disparities: A framework integrating psychosocial and environmental concepts [Review]. Environ Health Perspect 112: 1645-1653. http://dx.doi.org/10.1289/ehp.7074

HEW (U.S. Department of Health, Education and Welfare). (1964). Smoking and health: Report of the advisory committee to the surgeon general of the public health service. Washington, DC: U.S. Department of Health, Education, and Welfare. http://profiles.nlm.nih.gov/ps/retrieve/ResourceMetadata/NNBBMQ

Hill, AB. (1965). The environment and disease: Association or causation? Proc R Soc Med 58: 295-300.

IARC (International Agency for Research on Cancer). (2006). IARC monographs on the evaluation of carcinogenic risks to humans: Preamble. Lyon, France: World Health Organization. http://monographs.iarc.fr/ENG/Preamble/

Ioannidis, JPA. (2008). Why most discovered true associations are inflated [Review]. Epidemiology 19: 640-648. http://dx.doi.org/10.1097/EDE.0b013e31818131e7

IOM (Institute of Medicine). (2008). Improving the presumptive disability decision-making process for veterans. In JM Samet; CC Bodurow (Eds.), Improving the Presumptive Disability Decision-making Process for Veterans. Washington, DC: National Academies Press. http://www.nap.edu/openbook.php?record_id=11908

IPCC (Intergovernmental Panel on Climate Change). (2007). Climate change 2007: Impacts, adaptation and vulnerability. Cambridge, UK: Cambridge University Press. http://www.ipcc.ch/ipccreports/ar4-wg2.htm

NAPAP (National Acid Precipitation Assessment Program). (1991). The experience and legacy of NAPAP: Report of the Oversight Review Board of the National Acid Precipitation Assessment Program. Washington, DC.

Rothman, KJ; Greenland, S. (1998). Modern epidemiology. In Modern Epidemiology (2nd ed.). Philadelphia, PA: Lippincott, Williams, & Wilkins.

Sacks, JD; Stanek, LW; Luben, TJ; Johns, DO; Buckley, BJ; Brown, JS; Ross, M. (2011). Particulate-matter induced health effects: Who is susceptible? [Review]. Environ Health Perspect 119: 446-454. http://dx.doi.org/10.1289/ehp.1002255

U.S. EPA (U.S. Environmental Protection Agency). (1998). Guidelines for ecological risk assessment [EPA Report]. (EPA/630/R-95/002F). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. http://www.epa.gov/raf/publications/guidelines-ecological-risk-assessment.htm

U.S. EPA (U.S. Environmental Protection Agency). (2002). A framework for assessing and reporting on ecological condition: An SAB report [EPA Report]. (EPA-SAB-EPEC-02-009). Washington, DC. http://yosemite.epa.gov/sab%5CSABPRODUCT.NSF/7700D7673673CE83852570CA0075458A/$File/epec02009.pdf

U.S. EPA (U.S. Environmental Protection Agency). (2005). Guidelines for carcinogen risk assessment. (EPA/630/P-03/001F). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. http://www2.epa.gov/osa/guidelines-carcinogen-risk-assessment

U.S. EPA (U.S. Environmental Protection Agency). (2009). Integrated science assessment for particulate matter [EPA Report]. (EPA/600/R-08/139F). Research Triangle Park, NC: U.S. Environmental Protection Agency, National Center for Environmental Assessment. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=216546

U.S. EPA (U.S. Environmental Protection Agency). (2010). Integrated science assessment for carbon monoxide [EPA Report]. (EPA/600/R-09/019F). Research Triangle Park, NC: U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=218686

U.S. EPA (U.S. Environmental Protection Agency). (2013). Integrated science assessment for ozone and related photochemical oxidants. (EPA/600/R-10/076F). Research Triangle Park, NC: U.S. Environmental Protection Agency, National Center for Environmental Assessment. http://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=247492

UNEP (United Nations Environment Programme). (2003). Ecosystems and human well-being: A framework for assessment. Washington, DC: Island Press.

Vinikoor-Imler, LC; Owens, EO; Nichols, JL; Ross, M; Brown, JS; Sacks, JD. (2014). Evaluating potential response-modifying factors for associations between ozone and health outcomes: a weight-of-evidence approach [Review]. Environ Health Perspect 122: 1166-1176. http://dx.doi.org/10.1289/ehp.1307541

Zeger, SL; Thomas, D; Dominici, F; Samet, JM; Schwartz, J; Dockery, D; Cohen, A. (2000). Exposure measurement error in time-series studies of air pollution: Concepts and consequences. Environ Health Perspect 108: 419-426.