Evidence Synthesis and Integration in the IRIS Program

Xabier Arzuaga (based on Handbook materials developed by the IRIS Systematic Review Workgroup, particularly Barbara Glenn and Andrew Kraft)
National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency
Xabier Arzuaga | Arzuaga.Xabier@epa.gov

Introduction

Systematic reviews conducted as part of developing IRIS assessments (Figure 1) consist of structured processes for identifying the relevant evidence, evaluating individual studies, summarizing the relevant evidence (i.e., evidence synthesis), and arriving at summary conclusions regarding the overall body of evidence (i.e., evidence integration). These approaches were developed through discussions within EPA, and were informed by multiple reviews by the National Research Council (2011; 2014; 2018). In addition, IRIS assessments include quantitative toxicity values based on the evidence identified as most informative during the systematic reviews. The standard operating procedures, including frameworks and considerations for developing the different parts of the systematic reviews, are outlined in an internal document (IRIS Handbook; Figure 2).

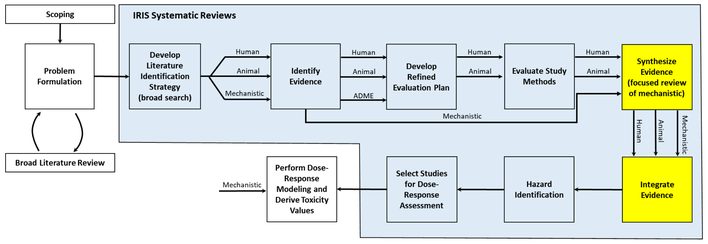

Figure 1. Systematic reviews in the IRIS Program: Figure adapted from the 2014 National Research Council review of the IRIS Program (adapted to show current workflows). Evidence synthesis and integration steps are highlighted.

Figure 2. IRIS Handbook: SOPs on approaches and considerations for applying principles of systematic review to IRIS assessments, including general frameworks, and examples. Evidence synthesis and integration steps are highlighted.

Overview of the Process

For each potential human health hazard, the evidence synthesis builds from the outcome-specific evaluations of individual studies, and discusses additional considerations across the sets of pertinent studies to summarize the available evidence in a manner that informs an evaluation of the body of evidence during evidence integration. Evidence integration is a two-step process based on structured, example-based frameworks for applying an adapted set of considerations described by Sir Bradford Hill (1965), first to each line of evidence, and then across all evidence. The general process is outlined in Figure 3.

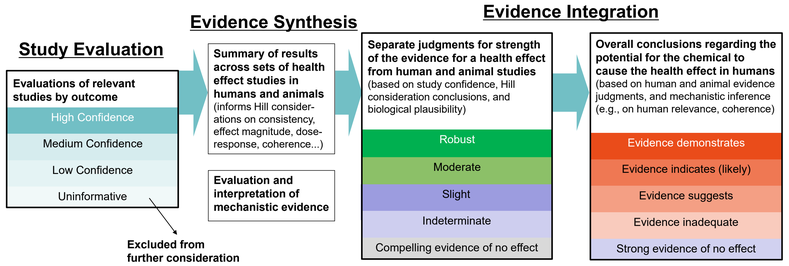

Figure 3. Outline of IRIS Evidence Synthesis and Integration. Human and animal evidence syntheses build from individual study evaluations and directly inform evidence integration across all lines of evidence.

Disclaimer: The views expressed are those of the authors and do not represent the views or policies of the U.S. EPA.

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

Evidence Synthesis

Summarize the information within each line of evidence (human, animal mechanistic), and analyze and present study results relevant to a given health effect to facilitate integration judgments.

- Narratives, not study summaries, focused on analyses that directly inform Hill considerations
- Human and animal health effect evidence is analyzed and synthesized separately. Mechanistic evidence is synthesized to inform the human and animal evidence conclusions (not shown).
- A primary goal of the evidence synthesis is to evaluate potential sources of heterogeneity across the study results (Figure 4), which informs evaluations of each Hill criterion.

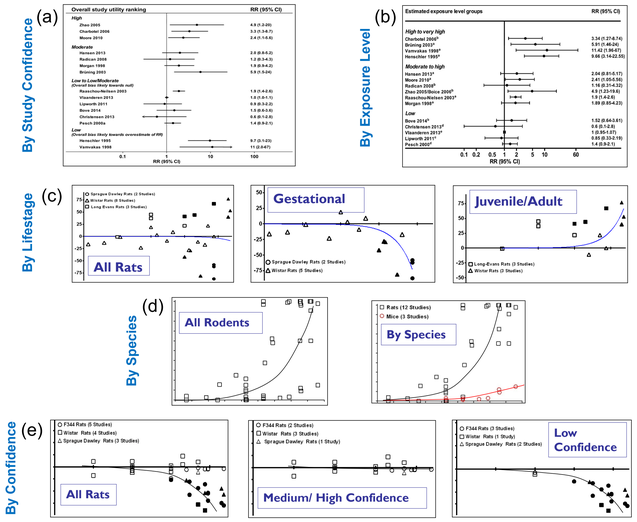

Figure 4. Evaluating Study Heterogeneity During Evidence Synthesis: (a) RoC Monograph on Trichloroethylene (2015); (b) EPA Toxicological Review of Trichloroethylene (2011); (c-e) "Edited" data from examples in draft IRIS assessments on hormones (c), pathology (d), and behavior (e).

Transitioning from Synthesis to Integration

The results of the analyses conducted during evidence synthesis inform an evaluation of each Hill consideration (Table 1) for the human and animal evidence relevant to a given health effect.

	Human Evidence Stream	Animal Evidence Stream
Individual Studies	• *High* or *medium* confidence studies provide stronger evidence within evaluations of each Hill consideration • Interpreting results considers biological as well as statistical significance, and findings across studies	
Consistency	• Different studies or populations increase strength	• Different studies, species, or labs increase strength • Analyze across study confidence, sensitivity, exposure levels/duration, lifestage, species or other factors • Unexplained inconsistency decreases evidence strength
Dose-response	• Simple or complex (nonlinear) relationships within or across studies provide stronger evidence • Dose-dependence that is expected, but missing, can weaken evidence (after considering the findings in the context of other available studies and biological understanding)	
Magnitude, Precision	• Large or severe effects can increase strength; further consider imprecise findings (e.g., across studies) • Small changes don't necessarily reduce evidence strength (consider variability, historical data, and bias)	
Coherence	• Biologically related findings within an organ system, within or across studies, or across populations (e.g., sex) increases evidence strength (considering the temporal- and dose-dependence of the relationship) • An observed lack of expected changes (e.g., based on biological linkage) reduces evidence strength	
Mechanistic Evidence on Biological Plausibility	• Informed by mechanistic evidence on the biological development of the health effect or toxicokinetic/ dynamic knowledge of the chemical or related chemicals • Mechanistic evidence in humans or animals of precursors or biomarkers of health effects, or of changes in established biological pathways or a theoretical mode-of-action, can strengthen evidence • Lack of mechanistic understanding does not weaken evidence outright, but it can if well-conducted experiments exist and demonstrate that effects are unlikely	

Table 1. Factors that increase or decrease the strength of the human and animal evidence for a health effect. Expert judgments are organized using adapted Hill considerations (not shown are temporality- addressed during epidemiology study evaluation, and natural experiments- very rare that is important to highlight).

Evidence Integration

Develop summary judgments of the evidence relevant to a human health effect within the evidence integration narrative

- A two-step process (Figure 5) involving transparent and structured approaches for drawing summary conclusions (examples in Figure 6) across all lines of evidence.
- Evidence profile tables (Figure 7) document the primary decisions and rationales.

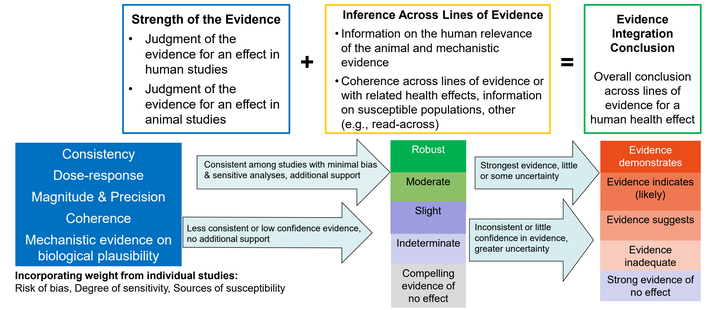

Figure 5. Evidence Integration Decision Process and Explanations

	Strength of the Evidence Judgments, Made Separately for Human and Animal Evidence
Robust	A set of consistent *high* or *medium* confidence, independent experiments reasonably ruling out alternative explanations; any conflicting set of studies is weaker. Additional criteria must also be met: **Human evidence**: Observed across populations, with clear dose-response evidence **Animal evidence**: Observed across labs or species, with multiple lines of additional support (e.g., pronounced severity or frequency; clear dose-response; coherence; a well-supported MOA).
	Overall Evidence Integration
Evidence Demonstrates	A very high level of certainty that exposure causes the health effect in humans: • The strongest evidence judgment (***robust***) for the human evidence stream • A moderately strong human evidence judgment (***moderate***) and the strongest animal evidence judgment (*robust*) alongside strong mechanistic evidence that MOAs and key precursors in animals are anticipated to occur in humans

Figure 6. Examples of Criteria for Evidence Integration Judgments (i.e., strongest judgments)

Step 1 – Evidence Integration of Human or Animal Evidence | **Step 2 – Overall Integration**

Studies and interpretation	Factors that increase strength	Factors that decrease strength	Summary of findings	Within stream evidence judgments	Inference across evidence streams	Overall conclusion
[Health Effect or Outcome Grouping]						
Evidence from Human Studies (Route) • References • Study confidence • Study design description	• Consistency • Dose-response gradient • Coherence of observed effects • Effect size • Mechanistic evidence providing plausibility • Medium or high confidence studies	• Unexplained inconsistency • Imprecision • Low confidence studies • Evidence demonstrating implausibility	• Results across studies • Human mechanistic evidence informing biological plausibility	Describe strength of the evidence from human studies, and primary basis: ⊕⊕⊕ Robust ⊕⊕◯ Moderate ⊕◯◯ Slight ◯◯◯ Indeterminate – – – Compelling evidence of no effect	• Human relevance of findings in animals • Cross-stream coherence • Other inferences: ○ Information on susceptibility ○ MOA analysis inferences ○ Relevant information from other sources (e.g., read across)	Describe conclusion(s) for the integration of all available evidence: ⊕⊕⊕ Evidence demonstrates ⊕⊕◯ Evidence indicates ⊕◯◯ Evidence suggests ◯◯◯ Evidence inadequate – – – Strong evidence supports no effect Summarize the models and range of dose levels upon which the conclusions were primarily reliant
Evidence for an Effect in Animals (Route) • References • Study confidence • Study design description	• Consistency and/or Replication • Dose-response gradient • Coherence of observed effects • Effect size Mechanistic evidence providing plausibility • Medium or high confidence studies	• Unexplained inconsistency • Imprecision • Low confidence studies • Evidence demonstrating implausibility	• Results across studies • Animal mechanistic evidence informing biological plausibility	Describe strength of the evidence for an effect in animals, and primary basis: ⊕⊕⊕ Robust ⊕⊕◯ Moderate ⊕◯◯ Slight ◯◯◯ Indeterminate – – – Compelling evidence of no effect		

Figure 7. Evidence Profile Table (Template): Documents the story of the evidence and supporting rationale for evidence integration decisions (note: may be subdivided, e.g., by study design)

Transitioning from Integration to Dose-Response

Evidence integration directly informs study selection and toxicity value derivation (Figure 8).

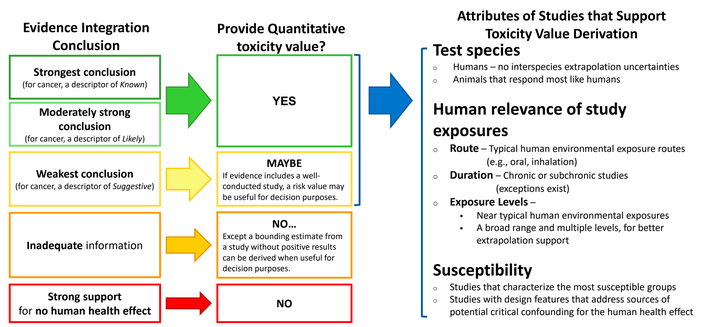

Figure 8. Considerations for Dose-Response: Note: study confidence informs study selection (not shown).