United States Environmental Protection Agency

HAZARD IDENTIFICATION

Steps 1-3: Identification, Evaluation, and Synthesis and Integration of Evidence

Samantha J. Jones, PhD
Associate Director, IRIS Program

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington DC

Development of IRIS Health Assessments: Hazard Identification

Steps Outlined by NAS (2011)

1. Identify Evidence → 2. Evaluate Evidence → 3. Synthesize and Integrate Evidence → 4. Select Studies to Derive Toxicity Values → 5. Derive Toxicity Values

Hazard Identification

- Overview of Hazard Identification in IRIS Assessments
- Identifying Evidence
- Evaluating Evidence
- Synthesizing and Integrating Evidence
- Summary

Overview of Hazard Identification in IRIS Assessments

- Hazard identification = integration of evidence from human, animal, and mechanistic studies in order to draw conclusions about the hazards associated with exposure to a chemical.

- The process for identifying hazards includes:
  - Identifying evidence by systematically searching for and selecting pertinent studies.
  - Evaluating evidence by applying criteria to assess and document the strengths and limitations of individual studies.
  - Synthesizing and integrating across human, animal, and mechanistic evidence streams to characterize the evidence for each effect.
  - Characterizing the overall weight of the evidence using common language to answer the fundamental question: “Does exposure to chemical X cause hazard Y?”

Identifying Evidence: Searching for and Selecting Pertinent Studies

- Identify full reports of primary (original data) studies pertaining to the key questions (e.g., Does the chemical cause liver toxicity?).
- Cast a wide net and screen results for relevance.
- Specify search strategies and cut-off dates of literature searches.
- Results of the literature search posted on the IRIS web site; public provides information on additional studies and ongoing research.

Identifying Evidence: Literature Search Strategy

- Multiple literature search strategies using overlapping core databases for published and unpublished studies (e.g., PubMed, Web of Science, Toxline, TSCATS).
- Search terms including chemical name, CASRN, major metabolites (if relevant), and keywords.
- Initial search using chemical terms; depending on size of database may need to refine with a secondary search, perhaps focusing on a particular topic.
- Augmentation using forward and backward searching.
  - Forward searching: database search of references citing a key publication.
  - Backward searching: database or non-database search of references cited by a key publication.

Identifying Evidence: Screening For Relevance

- Systematic review of each of the citations to determine relevance.
- Screening: Does this report contain primary data pertaining to health effects of chemical X?
  - Neither study quality nor study results considered in this step
  - Initial Review – Title and Abstract
  - Second Review – Full Text
- Possible Outcomes:
  - Not a source of primary data on chemical toxicity: exclude
  - Not a source of primary data on chemical toxicity: exclude but keep as additional resource (e.g., reviews, exposure levels, measurement methods)
  - Possible further review (e.g., language other than English)
  - Move to next step, evaluating evidence (i.e., quality of study)

Identifying Evidence: Selecting Pertinent Epidemiology Studies

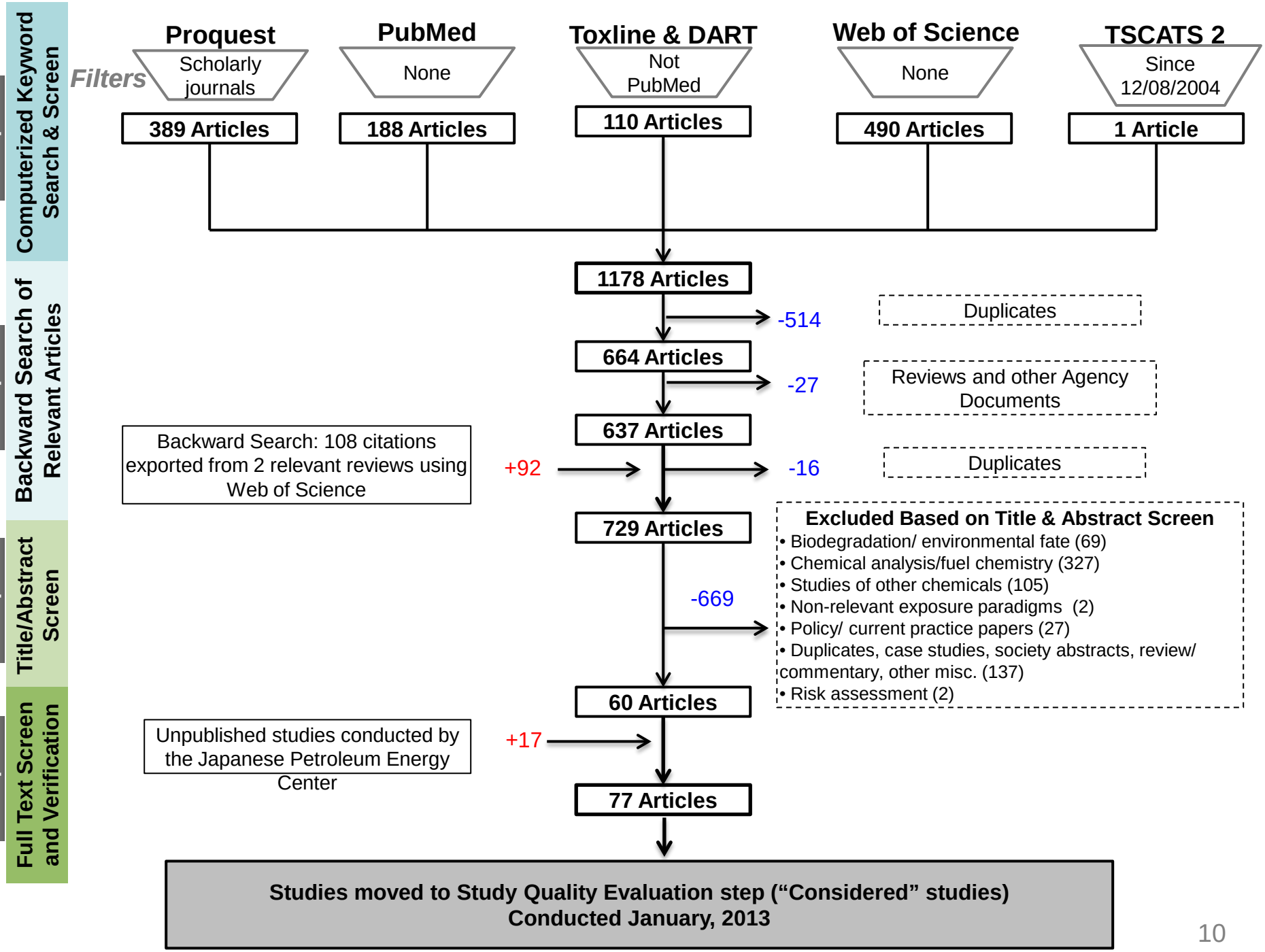
- Cohort studies, case-control studies, and some population-based surveys (e.g., NHANES) provide the strongest epidemiologic evidence.
- Ecological studies (geographic correlation studies) can provide strong evidence if there are large contrasts in exposure between areas as well as minimal exposure variation and population migration.
- Case reports can provide information about a rare effect or about the relevance of analogous results in animals, but their utility is limited.
- The assessment briefly reviews ecological studies and case reports but describes details only if they suggest effects not identified by other studies.

Identifying Evidence: Selecting Pertinent Experimental Studies

- Studies of oral, inhalation, or dermal exposure are considered most pertinent to human environmental exposure.
- Injection or implantation studies are often considered less pertinent but may provide valuable toxicokinetic or mechanistic information.
- Studies of effects from chronic exposure are most pertinent to lifetime human exposure.
- Shorter-duration studies involving animals or humans may provide toxicokinetic or mechanistic information.
- For developmental toxicity and reproductive toxicity, irreversible effects may result from a brief exposure during a critical period of development.

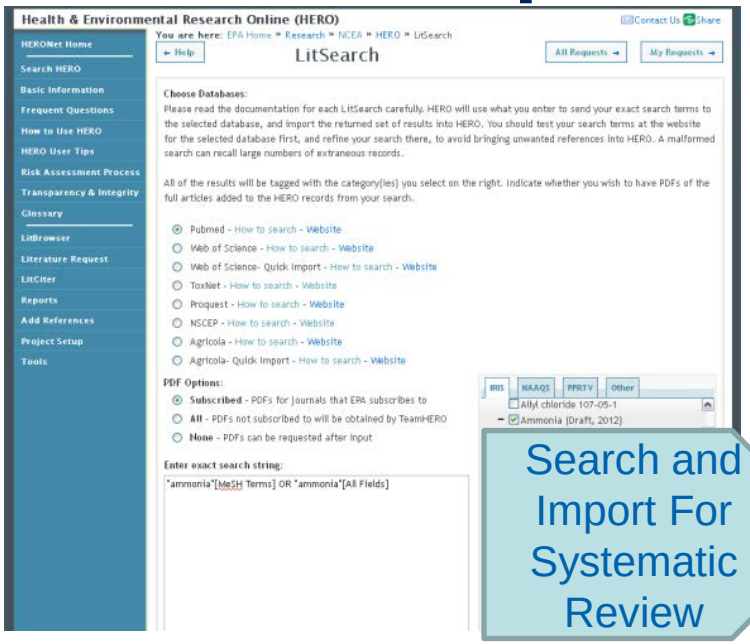
Identifying Evidence: Documentation

- Documentation is an essential component of the systematic review process.
  - Narrative describing the databases, dates of search, search terms and fields, and context of search.
  - Flow diagram for primary studies that are moved forward to study quality evaluation.
- Integrate into Health and Environmental Research Online (HERO).
  - Database system to facilitate complete, sustainable and effective assessment development.
  - House bibliographic citations and study data from scientific literature.
  - Hyperlinks to references in the text; references accessible to public through website (http://hero.epa.gov).


Step 1
Computerized Keyword Search & Screen
Step 2
Backward Search of Relevant Articles
Step 3
Title/Abstract Screen
Step 4
Full Text Screen and Verification
Filters
Proquest
Scholarly journals
PubMed
None
Toxline & DART
Not PubMed
Web of Science
None
TSCATS 2
Since 12/08/2004
389 Articles
188 Articles
110 Articles
490 Articles
1 Article
1178 Articles
-514
Duplicates
664 Articles
-27
Reviews and other Agency Documents
637 Articles
Backward Search: 108 citations exported from 2 relevant reviews using Web of Science
+92
-16
Duplicates
729 Articles
-669
Excluded Based on Title & Abstract Screen
• Biodegradation/ environmental fate (69)
• Chemical analysis/fuel chemistry (327)
• Studies of other chemicals (105)
• Non-relevant exposure paradigms (2)
• Policy/ current practice papers (27)
• Duplicates, case studies, society abstracts, review/ commentary, other misc. (137)
• Risk assessment (2)
60 Articles
Unpublished studies conducted by the Japanese Petroleum Energy Center
+17
77 Articles
Studies moved to Study Quality Evaluation step ("Considered" studies)
Conducted January, 2013

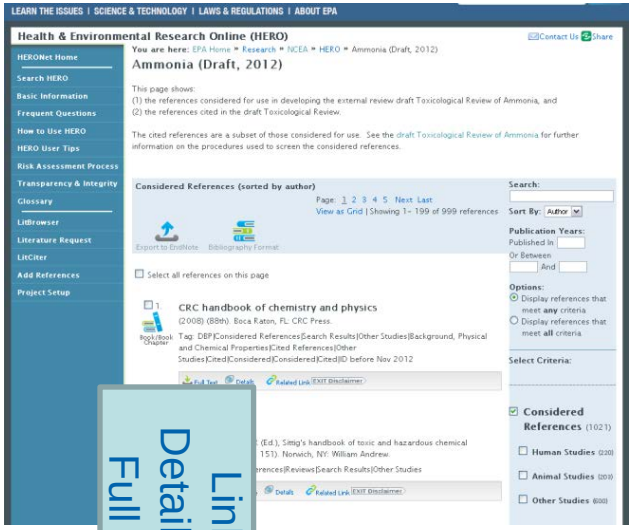
Components of HERO

LitSearch



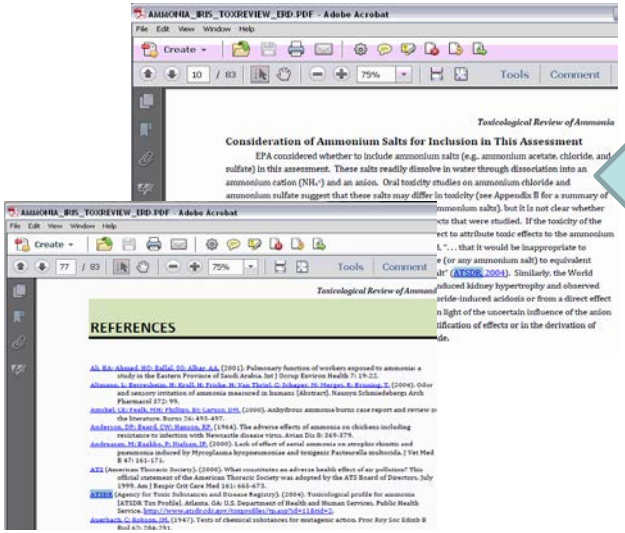
Search and Import For Systematic Review

Project Page



Link to Details and Full Text

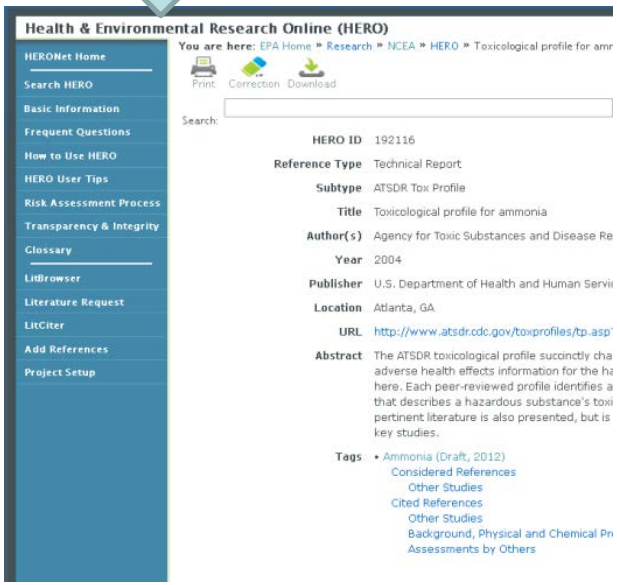
Draft and Final Assessments



LitCiter
"cite while you write"

Link to Details and Full Text

Reference Details



Evaluating Evidence for Hazard Identification

- Evaluate the quality of each individual study.
- Characterize the strengths and limitations that would affect the interpretation of or confidence in the results.
- Focus on study design, methodological aspects, and documentation.
- Apply a series of focused questions, tailored to the studies and health effect under review.
- Analysis independent of consideration of the direction or magnitude of the study's results.

Evaluating Evidence: Quality of Observational Epidemiology Studies

Target questions to **specific features** that may influence the interpretation of results:

- Study population; target population; setting
- Participation rate; loss to follow-up (e.g., attrition rate)
- Comparability (e.g., exposed and non-exposed; cases and controls)
- Exposure measures (e.g., procedure, range)
- Outcome measures (e.g., mortality, incidence)
- Data presentation and statistical analysis
- Consideration of likely confounding factors

Evaluating Evidence: Quality of Experimental Studies

Similar to process for epidemiology studies, questions are targeted to **specific features** that may influence the interpretation of results:

- Test subjects (e.g., species, sex, strain, age, number examined, and controls).
  - Research involving human subjects is considered only if conducted according to ethical principles.
- Exposure
  - Confidence in administration (e.g., intended dose/ concentration), test article.
  - Exposure groups, frequency, duration, and timing.
- Toxicity endpoints (e.g., method and timing of evaluation)
- Data presentation and statistics
- Reporting

Evaluating Evidence: Documentation

- Document the relevant details from the evaluation in a table; recording answers to the targeted questions.
- Sort studies into tiers according to the level of information they provide.
  - A study in the top quality tier would typically use an appropriate study design, have high-quality measures of exposure and outcome, and use adequate methods to analyze and present results. These studies would be given the greatest consideration within the context of hazard identification.
- Judgments made regarding tiering of a study for use in hazard identification are documented.

Evaluating Evidence: Documentation of Specific Features

Reference, Setting and Design	Participants, Selection, Follow-up	Comparability	Exposure Measure and Range	Outcome Measure	Consideration of Likely Confounding	Analysis and Presentation of Results (Estimate and Variability)	Sample Size; Power	Evaluation of Major Limitations
Lee et al., 1995 US (New York) chemical X production plant (cohort)	All men, age at baseline not reported; duration ≥ 12 months (mean 2.2 years), worked at plant 1960 – 1972 - plant operations began in 1945. Follow-up through 1990, 2% loss to follow-up, mean follow-up time 32 years	External (state mortality rates) referent; age and time-period matched (5 year groupings). Healthy worker effect seen for CVD (SMR 0.7) and all cancers (SMR 0.9). Internal referent: "no" exposure group	Exposure based on job records and personal/air monitoring; cumulative exposure calculated based on summations across all jobs (duration times average exposure)	Mortality (death certificates, ICD-8 and 9, underlying and contributing causes of death)	External comparison: use of age and time-period matched mortality rates.	SMR and 95% CI	Brain cancer: 4 obs cases	Low statistical power; not an inception cohort (had to "survive" to 1960 to be included)
Johnson et al., 1996 US (24 states) (case-control)	All deaths 1984-1986. Controls (died of causes other than cancer; frequency matched by age, sex, state and race)	Matching procedures for cases and controls	Death certificate occupation data; job exposure matrix developed to assess 11 chemical exposures	Mortality (death certificates, ICD-9), underlying cause of death)	Sex-specific odds ratios adjusted for marital status, race, socioeconomic status (3-levels), age at death	OR and 95% CI	10,540 cases, 42,160 controls	Non-differential exposure misclassification likely, particularly for women (lower quality occupation data for women)

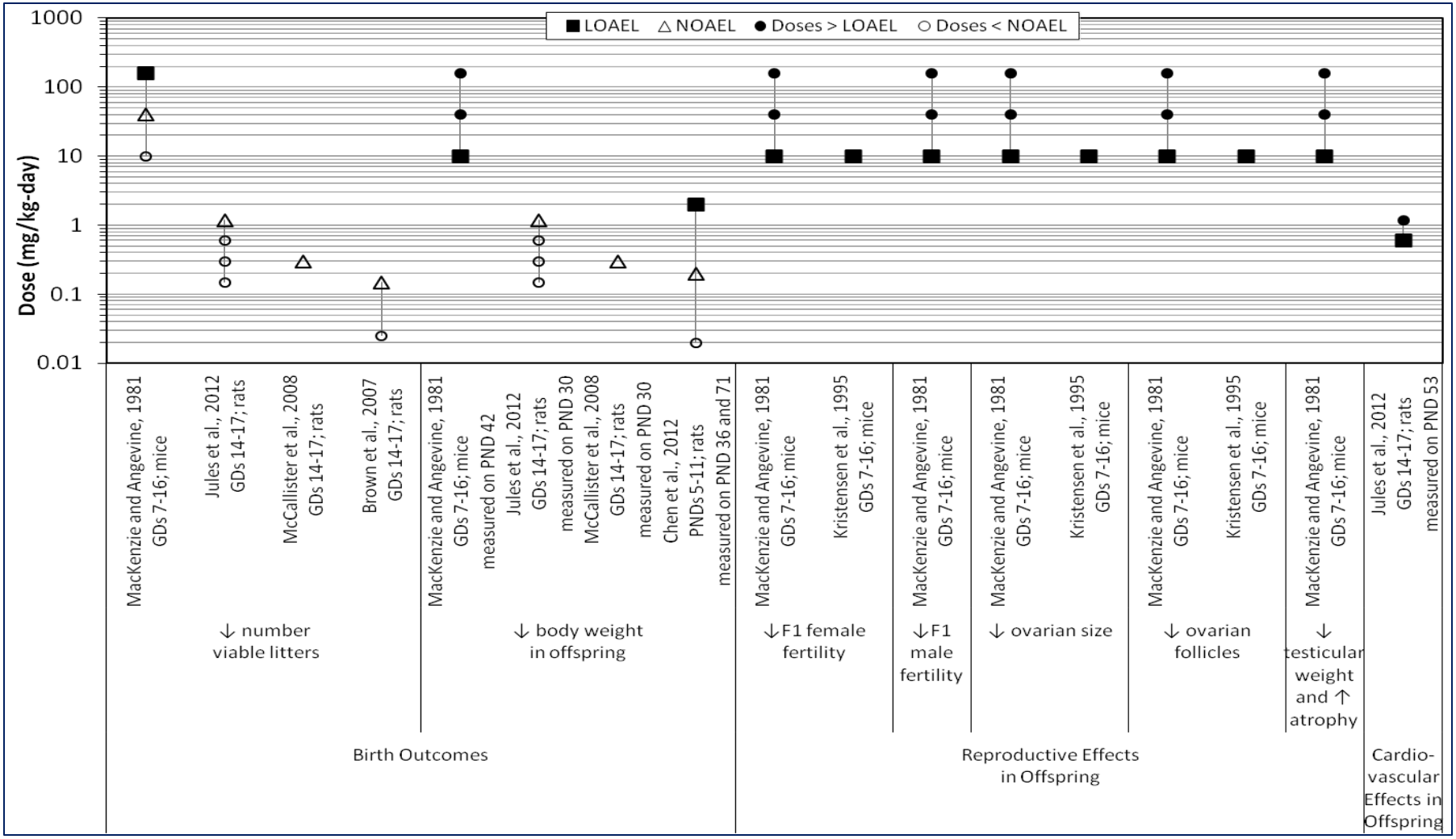
Standardized Reporting of Studies

- Use evidence tables and exposure-response arrays to present information and results related to a specific hazard.
  - If the database is extensive, the tables and arrays may be organized into two or more tiers based on the relevance and quality of the studies.
- Display findings of studies evaluating a relevant exposure scenario (taking into consideration route, timing, and dose).
- Display the spectrum of available study results; not restricted to statistically significant or positive associations.

Standardized Presentation of Studies: Evidence Table

Study Design and Reference	Results
Birth outcomes and postnatal growth	
Mackenzie and Angevine (1981) CD-1 mice, 30 or 60 F0 females/ dose 0, 10, 40, or 160 mg/kg-d by gavage GDs 7–16	↓ number of F0 females with viable litters: 46/60, 21/30, 44/60, and 13/30* ↓ F1 body weight at PND 20: Response relative to control: 0, 4, -7*, and -13* ↓ F1 body weight at PND 42: Response relative to control: 0, -6*, -6*, and -10* (no difference in pup weight at PND 4)
Jules et al. (2012) Long-Evans rats, 6–17 F0 females/dose 0, 0.15, 0.3, 0.6, or 1.2 mg/kg-d by gavage GDs 14–17	No overt signs of toxicity in dams or offspring, differences in pup body weight, or number of pups/litter
McCallister et al. (2008) Long-Evans Hooded rats, 5–6/group 0 or 0.3 mg/kg-d by gavage GDs 14–17	No difference in number of pups/litter No overt maternal or pup toxicity No difference in liver:body weight Increased brain:body weight ratio at PNDs 15 and 30 (data not shown)
Reproductive effects in offspring	
Mackenzie and Angevine (1981) CD-1 mice, 30 or 60 F0 females/ dose 0, 10, 40, or 160 mg/kg-d by gavage GDs 7–16	↓ number of F1 females with viable litters: 35/35, 23/35*, 0/55*, and 0/20* ↓ F2 litter size from F1 dams (20%) at 10 mg/kg-d (no litters were produced at high doses) ↓ size or absence of F1 ovaries (weights not collected) hypoplastic ovaries with few or no follicles and corpora lutea (numerical data not reported)
Kristensen et al. (1995) NMRI mice, 9 F0 females/dose 0 or 10 mg/kg-d by gavage GDs 7–16	↓ number of F2 litters (63%) ↓ F2 litter size (30%) ↓ ovary weight (31%) in F1 females Few or no small, medium, or large follicles and corpora lutea

Standardized Presentation of Studies: Exposure-Response Array



Overview of Synthesizing and Integrating Evidence

- For each hazard, the assessment evaluates the evidence as a whole to determine whether it is reasonable to infer a causal association between exposure to the chemical and occurrence of the effect.
  - Epidemiologic evidence
  - Experimental evidence
  - Mechanistic evidence
- All results, both positive and negative, of potentially relevant studies that have been evaluated for quality (higher quality given more weight) are considered.
- **Critical weighing of the evidence - not simply tallying the number of positive and negative studies.**

Overview of Synthesizing and Integrating Evidence

- Causal inference involves scientific judgment and the considerations are nuanced and complex.
- Several health agencies (e.g., US Surgeon General, IARC, IOM, and US EPA) have developed frameworks for causal inference in epidemiologic and experimental studies.
- Each framework considers aspects of an association that suggest causation, commonly known as the Hill considerations, and elaborated on by Rothman and Greenland:
  - Strength of association
  - Consistency of association
  - Specificity of association
  - Temporal relationship
  - Biologic gradient (exposure-response relationship)
  - Biologic plausibility
  - Coherence
  - Natural experiments
  - Analogy

Overview of Synthesizing and Integrating Evidence

- Two-step process in IRIS assessments:
  1. Synthesis of evidence <u>within</u> data sets for each target organ/system–
     - duration of exposure
     - route of exposure (oral, inhalation, other)
     - type of study (epidemiologic, animal bioassays, mode of action)
     - type of endpoint
  2. Integration of evidence –
     - <u>across</u> all data sets for each target organ/system
     - <u>across</u> different target organs/systems

Overview of Synthesizing and Integrating Evidence

- Evidence is synthesized and integrated in the context of the Hill Criteria.
- Quality human data preferred because these data are more relevant in the assessment of toxicity to human health and avoid the uncertainty associated with potential interspecies differences.
- Many chemical databases contain little or no human data; assessments frequently rely on animal data.
- Animal data used under the assumption that toxicity is conserved across species; effects observed in animals would be expected to occur in humans unless data are available to indicate otherwise.
- Dose-response relationships in animal studies may be at higher doses than would be anticipated for humans.

Synthesizing and Integrating Evidence: Epidemiology Studies

- For each effect, the assessment evaluates the evidence from the epidemiologic studies as a whole.
- The objective is to determine whether and to what degree the collective evidence supports a conclusion of an association or lack of an association between an exposure and a health outcome, for which reasonable alternative explanations (e.g., reverse causation, chance, bias, or confounding) are judged to be unlikely?
- The assessment may select a standard descriptor to characterize the epidemiologic evidence of an association between exposure and an effect.
  - Sufficient epidemiologic evidence of an association consistent with causation
  - Suggestive epidemiologic evidence of an association consistent with causation
  - Inadequate epidemiologic evidence to infer a causal association
  - Epidemiologic evidence consistent with no causal association

Synthesizing and Integrating Evidence: Experimental Studies

- For each effect, the assessment evaluates the evidence from the animal experiments as a whole.
- The objective is to determine the extent to which they indicate a potential for effects in humans.
  - e.g., Consistent results across various species and strains increases confidence that similar results would occur in humans.
- In weighing evidence from multiple experiments, IRIS assessments distinguishes between:
  - Conflicting evidence – mixed positive and negative results in same sex and strain using a similar study protocol
  - Differing results – positive results and negative results in different sexes or strains or use different study protocols

Synthesizing and Integrating Evidence: Mechanistic Studies

- Mechanistic data contribute to the hazard evaluation of empirical evidence from human and animals by informing the:
  - Biological plausibility of a causal interpretation in human studies.
  - Biological plausibility of animal data are generalizable to humans.
  - Susceptibility of particular populations or lifestages.
- The focus of the analysis is to describe, if possible, adverse outcome pathways (AOPs) that lead to a health effect; the analysis encompasses toxicokinetic and toxicodynamic processes.
- For each endpoint, the evaluation of AOPs begins by identifying information that may help identify the toxic moiety and the target site, and how the toxic agent is delivered to that site. Including:
  - Observational and experimental evidence specific to the chemical and endpoint.
  - Studies of metabolites or compounds that are structurally similar or act through similar mechanisms.
  - Information on how the chemical may disrupt normal biological processes, interactions with background aging/disease processes or with other chemicals, and factors affecting biological susceptibility.

Synthesizing and Integration Evidence: Mechanistic Studies

- For each effect, the assessment discusses the available information that may help identify key events in the hypothesized modes of action (MOA).
  - Key events: empirically observable, necessary precursor steps or biologic markers of such steps.
  - Mode of action: a series of key events involving interaction with cells, operational and anatomic changes, and resulting in disease.
- The assessment addresses several questions about each hypothesized MOA.
  - Is the hypothesized MOA sufficiently support in test animals?
  - Is the hypothesized MOA relevant to humans?
  - Which populations or lifestages can be particularly susceptible to the hypothesized MOA?
- Information on mode of action is not required for a conclusion that exposure to a chemical is related to an effect.

Synthesizing and Integrating Evidence: Characterization

- The hazards associated (or potentially associated) with a chemical are identified by evaluating the totality of evidence.
- The human, animal, and mechanistic evidence that is pertinent to causation are synthesized and integrated to reach conclusions about the hazards.
- The assessment develops a narrative that describes this integration and concludes with a statement regarding hazard.
- For cancer hazards, the following standard descriptors combine epidemiologic, experimental, and mechanistic evidence of carcinogenicity to conclude a chemical is:
  - Carcinogenic to humans
  - Likely to be carcinogenic to humans
  - Suggestive evidence of carcinogenic potential
  - Inadequate information to assess carcinogenic potential
  - Not likely to be carcinogenic to humans

Hazard Identification: Summary

- Hazard identification involves a process of identifying evidence, evaluating evidence, and synthesizing and integrating across human, animal and mechanistic evidence.

- This process results in characterization of the overall evidence, and conclusions about the hazards associated with exposure to a chemical.

- In IRIS assessments, Hazard Identification includes:

  - A chapter in the new document template, with subsections based on organ/system-specific hazards.
  - Evidence tables to present key study findings that support how hazards are identified.
  - Exposure-response arrays used as visual tools to inform the hazard characterization.
  - A summary of evidence and a hazard conclusion statement provided at the end of each organ/system-specific subsection.
  - A summary of all of the potential hazards and a rationale for those hazards that are carried forward for dose-response analysis provided at end of the Hazard Identification chapter.

New Developments in Response to the NRC Recommendations

- A comprehensive and transparent review and documentation process for:
  - Literature searches
  - Literature screening and selection of studies
  - Study quality evaluation of individual studies
- New document structure focuses on Hazard Identification; separated from Dose-Response Analysis.
- Standardized presentation of data in evidence tables and exposure-response arrays <u>for each</u> endpoint **and** <u>across</u> endpoints.
- More descriptive and transparent text to synthesize and integrate data –
  - What is the available evidence?
  - Is there evidence to conclude that an association exists between an exposure and a hazard?
  - Across the whole data set, what potential hazards have the most/least credible evidence?
- Strengthened, more integrated and transparent discussion of the weight of the available evidence supporting hazard identification.