EPA
United States
Environmental Protection
Agency

NATIONAL ACADEMY OF SCIENCES COMMITTEE TO REVIEW ADVANCES MADE TO THE IRIS PROCESS

February 1-2, 2018

Office of Research and Development
NCEA, IRIS

EPA
United States
Environmental Protection
Agency

INTRODUCTION AND OVERVIEW OF IMPROVEMENTS TO THE IRIS PROGRAM

Tina Bahadori* and Kris Thayer

[*Speaking]

Office of Research and Development
NCEA, IRIS

- **Created in 1985 to foster consistency in the evaluation of chemical toxicity across the Agency.**
- **IRIS assessments contribute to decisions across EPA and other health agencies.**
- **Toxicity values**
  - Noncancer: Reference Doses (RfDs) and Reference Concentrations (RfCs).
  - Cancer: Oral Slope Factors (OSFs) and Inhalation Unit Risks (IURs).
- **IRIS assessments have no direct regulatory impact until they are combined with**
  - Extent of exposure to people, cost of cleanup, available technology, etc.
  - Regulatory options.
  - Both of these are the purview of EPA's program offices.

IRIS Provides Scientific Foundation for Agency Decision Making

IRIS

- Clean Air Act (CAA)
- Safe Drinking Water Act (SDWA)
- Food Quality Protection Act (FQPA)
- Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA)
- Resource Conservation and Recovery Act (RCRA)
- Toxic Substances Control Act (TSCA)

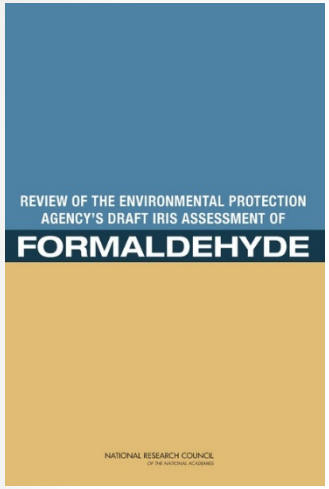
Broad Input to Support

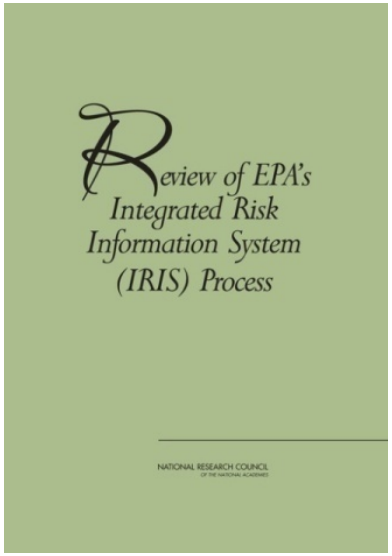
- Agency Strategic Goals
- Children's Health
- Environmental Justice

New Leadership Structure in NCEA

- **In January 2017, EPA appointed new leadership to the National Center for Environmental Assessment and to its IRIS Program.**
  - **NCEA Director:** significant experience in the chemical and energy industries, and formerly the Director of ORD's Chemical Safety for Sustainability National Research Program, Tina Bahadori brings knowledge of TSCA, innovative applications of computational toxicology, and exposure science.
  - **IRIS Program Director:** As a recognized leader in systematic review, automation, and chemical evaluations, Kris Thayer brings experience in early partner and stakeholder engagement and input, and demonstrated actions to increase capacity and transparency in assessments.
- **Improved responsiveness and accountability through Senior Leadership Team.**
- **Integrating across the spectrum of human and ecological RA practices.**

Drivers for this Study





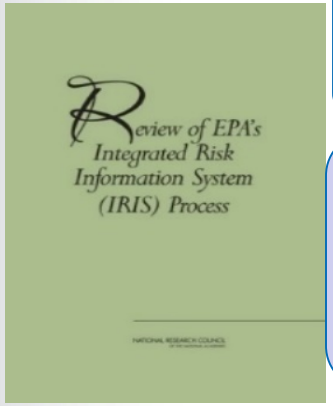
GAO
Accountability * Integrity * Reliability

https://www.gao.gov/highrisk/transforming_epa_and_toxic_chemicals/why_did_study

Fiscal Year 2017 Appropriations
https://www.congress.gov/114/crpt/srpt281/CRPT-114srpt281.pdf

NAS (2014) Overarching Statements

2014



"Overall, the committee finds that substantial improvements in the IRIS

Overall, the committee finds that substantial improvements in the IRIS process have been made

The [2011] committee recognized that its suggested changes would take several years and an extensive effort

EPA has not only responded to the recommendations made in the NRC formaldehyde report, but is well on the way to meeting the general systematic review standards

committee is confident that there is an institutional commitment to

moved forward steadily in planning for and implementing changes in each element of the assessment process

continuing commitment to improving the process, and successes to

The committee commends EPA for its substantive new approaches...the revisions will transform the IRIS Program

Program." [p.135]

Previous Phased Improvements to the IRIS Program

- **Revising the structure of assessments to enhance the clarity and transparency of presentation:**
  - Detailing the methods underlying each step of draft development (e.g., literature search strategy).
  - Restructuring the document into separate hazard identification and dose-response chapters.
  - Replacing lengthy study summaries with synthesis text, supported by standardized tables and graphs.
- **Implementing "IRIS Enhancements"**
  - An updated process for developing and reviewing assessments that increases public input and peer consultation at earlier stages of assessment development, and clarifies processes for considering new evidence and scientific issues.
- **Establishing the SAB Chemical Assessment Advisory Committee (CAAC)**
  - 5 IRIS assessments completed CAAC review since 2014.
- **Restructuring the IRIS Program to create expertise-specific workgroups and improved assessment oversight.**

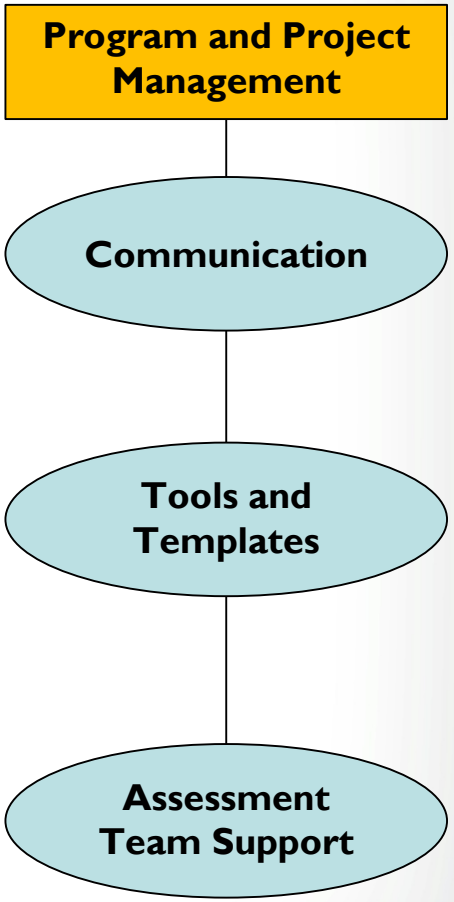
Quality Management

- **Assessment Development and Review**
  - Quality management inherent to systematic review methodology (e.g., independent screening of studies)
  - Rigorous review process includes internal, public, and external peer review
- **Scientific Support Teams**
  - Systematic review methods (Systematic Review Workgroup)
  - Systematic review support to chemical assessment teams (e.g., screening, study evaluation, data extraction, use of specialized software, etc. – train the trainer model)
  - Discipline-specific workgroups (e.g., epidemiology, PBPK, neurotoxicology, etc.)
  - Executive oversight
- **Roles and Responsibilities**
  - Assessment plans, protocols, and draft assessments indicate contributors and roles
  - Given current budget there is very limited use of contract support to conduct assessments
- **Training**
  - regular training via skill-building seminars, focused discussions, and retreats

Improved Practices for Timeliness and Resource Management

Current Program and Project Management in IRIS:

- *Centralized communication processes* for providing staff with updates on near-term priorities, template materials, and other process-oriented decisions.
- *Development and maintenance* of templates and checklists for key steps of assessment development using Microsoft SharePoint and Project as collaborative, web-based tools for assessment teams and project managers (document management and storage; scheduling support).
- *Dedicated IRIS Program staff and on-site programmatic contractor support* to facilitate continued implementation of program and project management principles.



GAO 2017 Report

- **Acknowledged the actions ORD has taken to enable the IRIS Program to produce timely, transparent, and credible assessments in support of EPA's mission.**
- **Discussions with GAO during and after the release of the 2017 High Risk Report have focused on approaches to demonstrate how management and integrity initiatives within IRIS are supporting the transformation of the program**

Summary of 2015 and 2017 GAO High Risk Criteria Ratings of the IRIS Program		
GAO High Risk Criteria	2015 Rating	2017 Rating
Leadership Commitment	Met	Met
Monitoring	Partially Met	Met
Action Plan	Partially Met	Partially Met
Demonstrated Progress	Not Met	Partially Met
Capacity	Not Met	Partially Met

- **IRIS is engaged in continual ongoing discussion with GAO regarding recommendations from the 2008, 2012, and 2013 reports.**
- **Of the seventeen recommendations issued in these three reports, as of June 2017, we have successfully closed ten recommendations and are rapidly moving to address the remaining seven.**

IRIS Multi-Year Agenda

- **Released to the public December 2015**
  - Result of a survey EPA program and regional offices for their assessment needs balanced with resource availability.
  - Other chemicals were also carried over from earlier prioritizations
  - Reflects global priorities
- **In FY 2018, reaffirm priorities; identify new or more urgent needs.**
- **Engage states.**

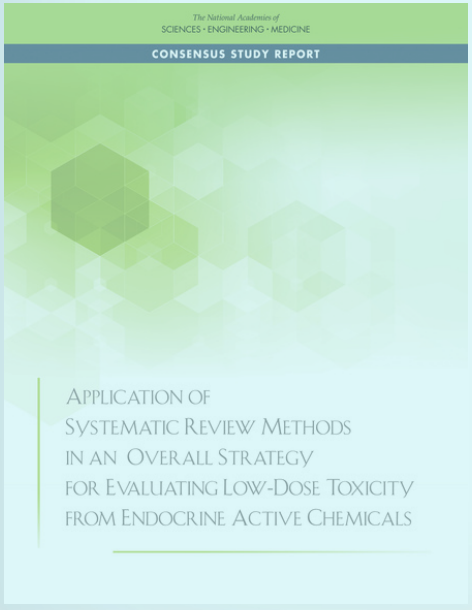
Group	Chemicals
1	Manganese
	Mercury/methylmercury
	Nitrate/nitrite
	Perfluoroalkyl compounds
	Vanadium and compounds
2	Acetaldehyde
	Ammonia (oral)
	Cadmium and compounds
	Uranium
3	Di-(2-ethylhexyl) phthalate
	Dichlorobenzene isomers
	Methyl t-butyl ether (MTBE)
	Nickel and compounds
	Styrene

A Portfolio Approach

- Moving away from a 'one-size-fits-all' approach to risk assessment towards a spectrum of assessment products to meet specific decision contexts;
- Facilitating the incorporation of new science into risk assessment and decision-making;
- Enabling assessments to be better tailored to meet needs of decision makers;
- Increasing the number of chemicals that can be evaluated for their effects on human health by utilizing constrained resources in the most efficient manner.

Leading Edge of Science – Systematic Review

NAS 2017: Reflections and Lessons Learned from the Systematic Review



"….one disadvantage in conducting a systematic review is that it can be time and resource intensive, particularly for individuals that have not previously conducted a systematic review." [p.157]

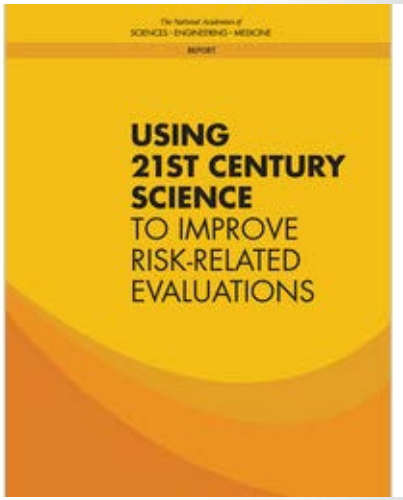
"The committee discussed at length whether it could provide EPA with advice about when a systematic review should be performed but decided it could not be more specific because that decision will depend on the availability of data and resources, the anticipated actions, the time frame for decision making, and other factors." [p.157]

"The committee also recognized that it might be advantageous for EPA to build on existing systematic reviews that are published in the peer-reviewed literature." [p.157]

"The committee recognizes that the methods and role of systematic review and meta-analysis in toxicology are evolving rapidly and EPA will need to stay abreast of these developments, strive for transparency, and use appropriate methods to address its questions." [p.157]

Leading Edge of Science – New Data Streams

Next Generation IRIS



- IRIS in the 21st Century – implement recommendations of the NAS 2017 report, Using 21st Century Science to Improve Risk-Related Evaluations;
- New Approach Methods – see poster session
- Collaborate with Tox21
  - build expert-judgment case studies that inform assessment development and fill gaps in assessments, especially for data poor chemicals;
  - inform where resources should be strategically invested to generate additional data.
- Create efficiencies – engage other agencies to share common practices, data, and tools, and more efficiently leverage resources across the federal government.
- Refresh science – MOU's with academia and other federal agencies; strategic staffing; deeper engagement with health agencies in states.

How is IRIS Evolving?

- **Increase transparency and full implementation of systematic review**
  - implement using approaches that foster consistency across the IRIS Program; many active and all new starts address systematic review-related recommendations of 2014 NAS report
- **Modernize the IRIS Program**
  - through automation and machine learning to expedite systematic review, incorporation of emerging data types
- **Modularize product lines**
  - implement a portfolio of chemical evaluation products that optimize the application of the best available science and technology. These products will allow IRIS to remain flexible and responsive to clients within the EPA as well the diverse collection of stakeholders beyond EPA, including states, tribal nations, and other federal agencies.
- **Enhance accessibility**
  - provide outreach and training to make systematic review practices ubiquitous and more accessible; enhance data sharing through publicly available software platforms for assessments developed by EPA, other federal and state agencies, industry, academia and other third-parties.

IRIS has Addressed the Major NAS 2014 Recommendations

NAS 2014 Topics	IRIS Process Improvements
General Process Issues (Chapter 2)	• Quality management pipeline implemented • Program and project management processes implemented • Frequent opportunities for stakeholder engagement
Future Directions (Chapter 8 "Lessons Learned" and "Looking Forward")	• Processes being implemented include flexibility to incorporate evolving methods in systematic review and risk assessment • Increased collaboration with federal partners and international experts prevents duplication of effort and maintains cutting edge approaches • Current research efforts and training serve to ensure that methods and staff are able to adapt to changing scientific contexts and sources of evidence, including new and emerging data types

EPA
United States
Environmental Protection
Agency

SESSION 1: SYSTEMATIC REVIEW IN THE IRIS PROGRAM - EVIDENCE IDENTIFICATION

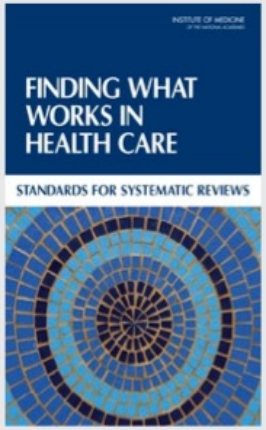
Kris Thayer*, Andrew Kraft*, April Luke, Beth Radke, Michele Taylor

[*Speaking]

Office of Research and Development
NCEA, IRIS

Systematic Review

A structured and documented process for transparent literature review[1]

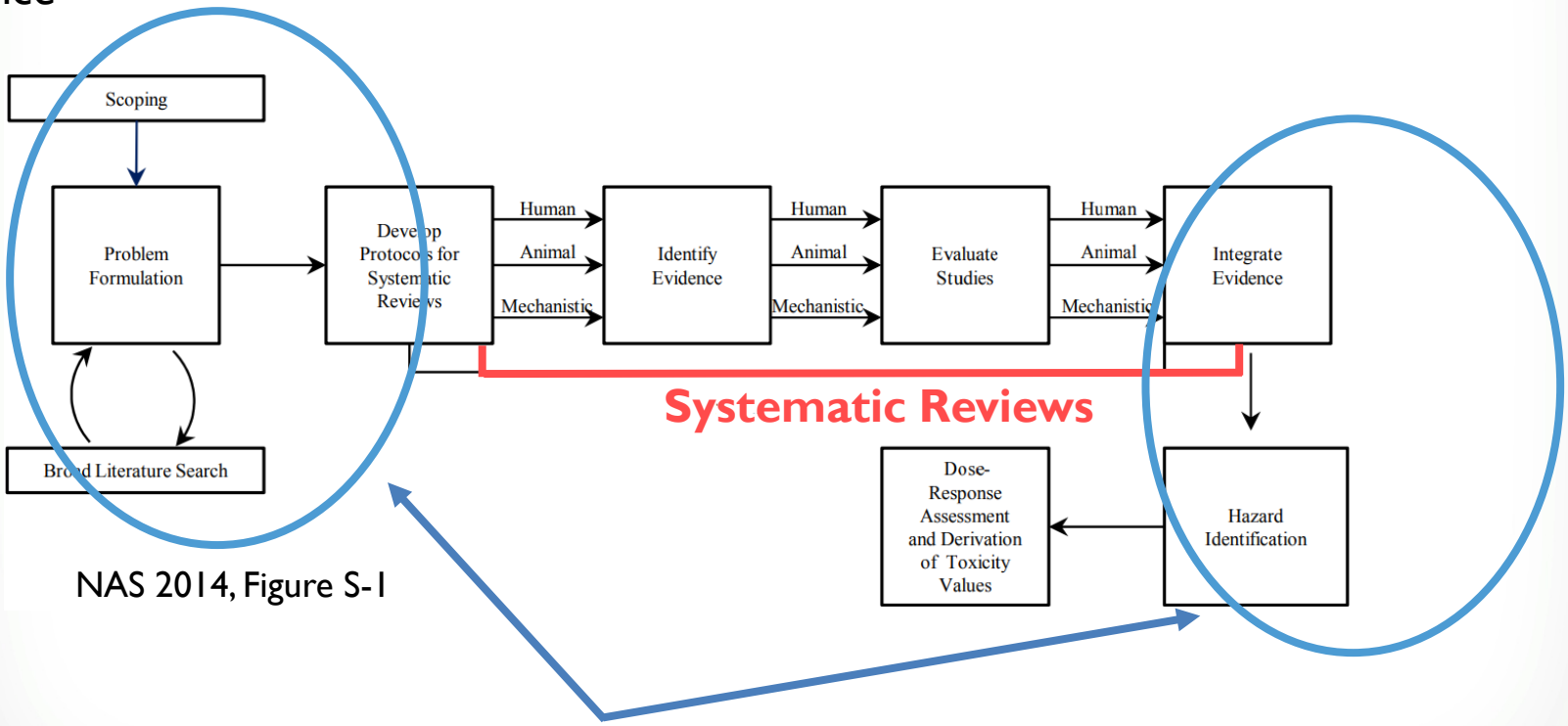


"As defined by IOM [Institute of Medicine], systematic review 'is a scientific investigation that focuses on a specific question and uses explicit, pre-specified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies.'" [p. 4] (NRC, 2014)

[1] Institute of Medicine. Finding What works in Health Care: Standards for Systematic Reviews. p.13-34. The National Academies Press. Washington, D.C. 2011

Systematic Review Elements (NAS 2014)

"In the context of IRIS, the committee has defined systematic review as including protocol development, evidence identification, evidence evaluation, and an analytic summary of the evidence"



NAS 2014, Figure S-1

IRIS also considers these phases as part of its systematic review process

EPA
United States
Environmental Protection
Agency

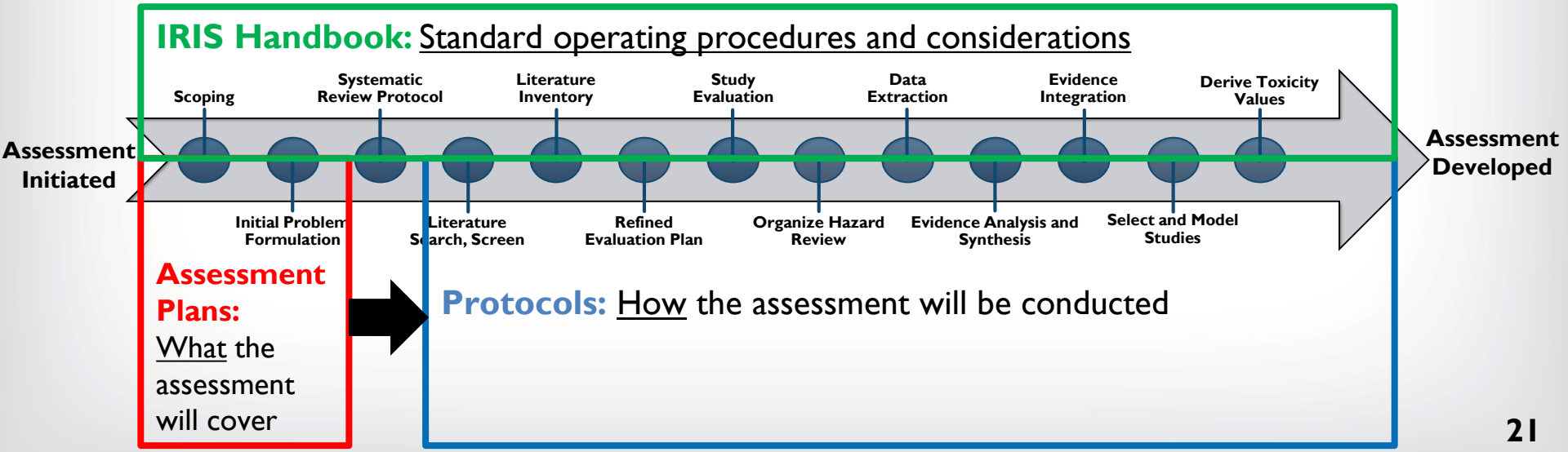
Scoping, Problem Formulation, and Protocol Development

Office of Research and Development
NCEA, IRIS

IRIS Systematic Review Documents

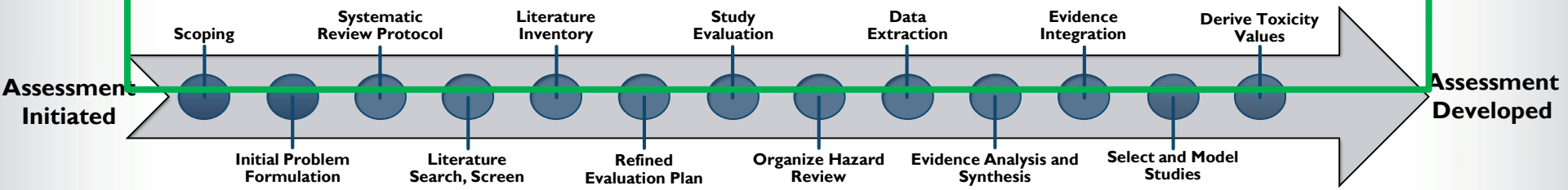
Address several NAS 2014 High Priority (Box 8-1) Recommendations

- "EPA needs to...complete documents, such as the draft handbook, that provide detailed guidance for developing IRIS assessments." (Chapter 2, General Process)
- "EPA should include protocols for all systematic reviews conducted for a specific IRIS assessment as appendixes to the assessment." (Chapter 3, Problem Formulation and Protocol Development)



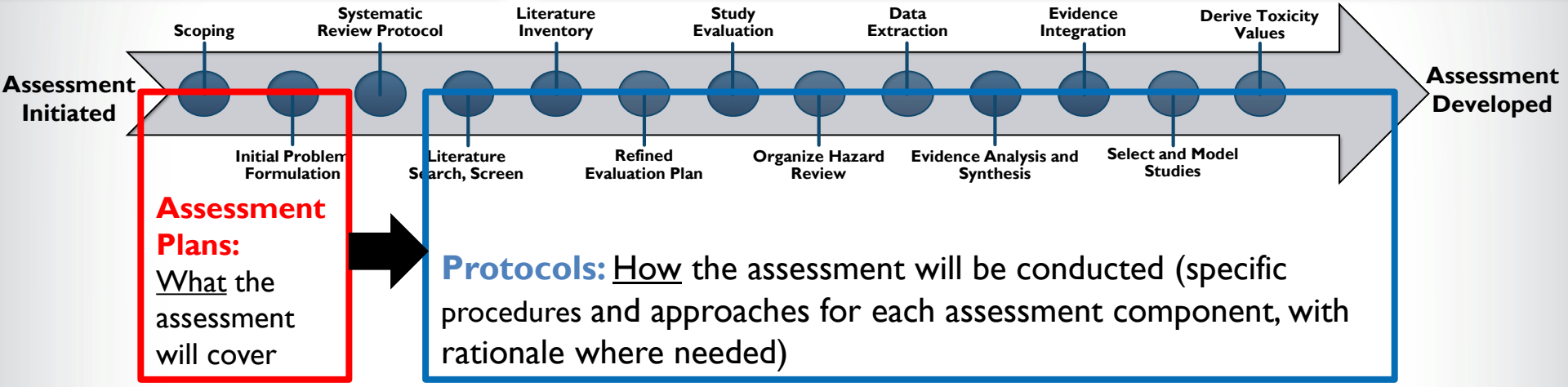
IRIS Handbook

IRIS Handbook: Approaches and considerations for applying principles of systematic review to IRIS assessments, general frameworks, and examples.



- IRIS Handbook level of detail aimed for EPA staff and contractors, e.g., use of HERO, timelines for internal review steps, etc.
- Currently being updated to reflect Agency input, evolving IRIS practices as systematic review approaches are tested through implementation, and public comment received on chemical-specific protocols (e.g., chloroform)
- Evergreen to reflect future advances
- Anticipate public release in 2018

IRIS Assessment Plans and Protocols



- Chemical-specific documents
- IRIS Assessment Plans (IAPs) are problem formulation and scoping documents that include more elements of systematic review
- Protocols outline methods, including updates to the IAPs
- IAPs and protocols include proposed "modularity," targeted focus and use of existing assessments
- Templates created to promote consistency across the IRIS Program, which is implemented across NCEA divisions and geographical locations

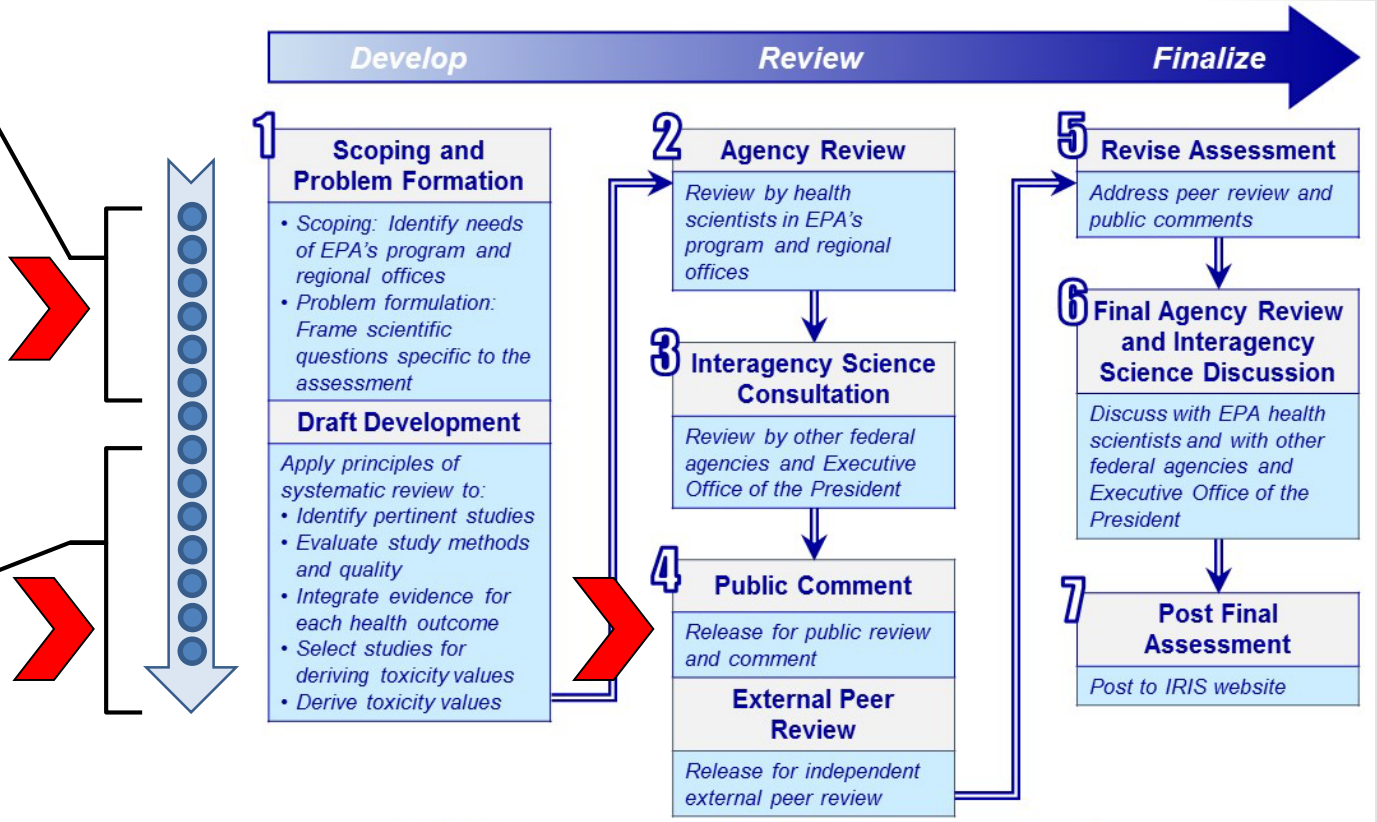
IRIS Assessment Plans, Protocols, and 7-Step IRIS Process

Early Step 1: IRIS Assessment Plans

- What the assessment covers
- 30-day public comment period + public science meeting

Mid-Step 1: Protocols

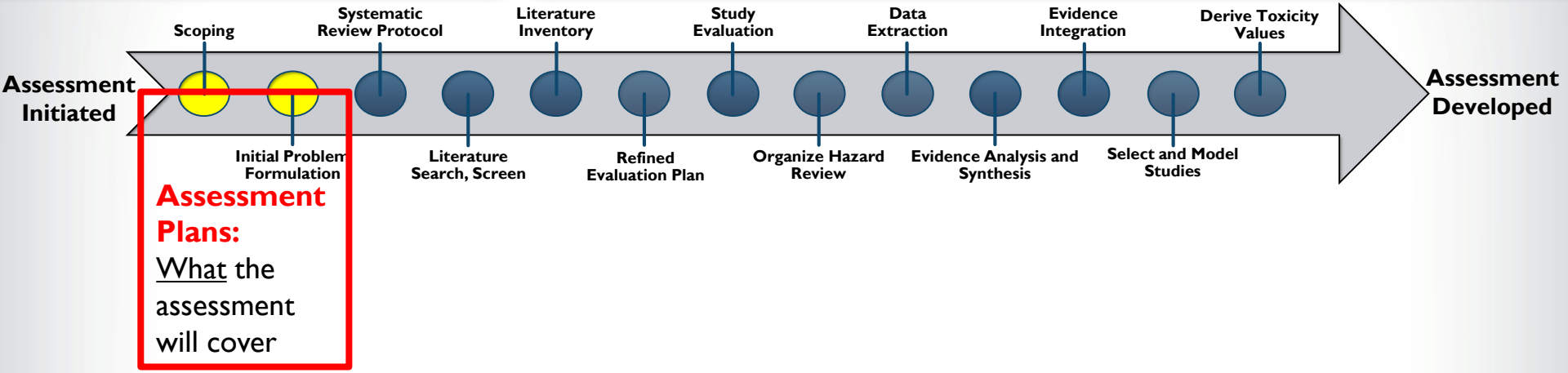
- How the assessment will be conducted
- 30-day public comment



Opportunities for Public Comment

https://www.epa.gov/iris/basic-information-about-integrated-risk-information-system#process

IRIS Assessment Plan (IAP)



- Scoping and initial problem formulation determinations
  - Background and Agency need, exposure context, objectives and specific aims, key areas of scientific complexity
  - Includes draft PECO (Populations, Exposures, Comparators, and Outcomes) criteria which outlines evidence considered most pertinent
  - Internal review of IAP fosters early and focused Agency engagement
- Released for a 30-day public comment period + public science discussion (beginning of IRIS Step 1)
  - Examples: chloroform, ethylbenzene, nitrate/nitrite (Sept 2017), uranium (Jan 2018)

IRIS Assessment Plan (IAP) Content

Table 1. EPA program and regional office interest in an as uranium

Program or regional office	Oral	Inhalation	Statutes/regulations	Anticip
Office of Land and Emergency Management	✓		CERCLA	Uranium toxicc used to make r response or re short-term ren response actio to conduct sho Superfund site costs from pot Uranium is liste under CERCLA National Priorit
Region 10[a]	✓			
OW	✓		Safe Drinking Water Act	Uranium toxicc used to inform risk determ associated with contamina found in water. The maxin level goals of 0 µg/L and m contaminant level of 30 µg were published in 2000 (65

3. OVERALL OBJECTIVE, SPECIFIC AIMS, AND DRAFT PECO (POPULATIONS, EXPOSURES, COMPARATORS, AND OUTCOMES) CRITERIA

The characteriz of uranium systematic Given the e not be a fo directly on this assess will be diss specific ain

3.1. SPECIFIC AIMS

- Building on f epidemiolog outlined in t ATSDR litera
- Conduct stuc toxicological subsequent a assessment's
- Examine wh toxicity valu kidney toxici will examine of additional uranium.
- If newer PEC considered a synthesis/in studies used using the me
- Extract data considered i
- For the ident (including bo using a narra examined by ATSDR where important new studies are *not* identified, EPA will seek to base its hazard conclusions on ATSDR's findings unless compelling reasons for further review are identified.

Table 2. Draft PECO (populations, comparators, exposures, and outcomes) criteria for the uranium assessment

PECO element	Evidence
Population[a]	*Human*: Any population and all life stages (e.g., children, general population, occupational, or high exposure from an environmental source). The following study designs will be considered most informative: controlled exposure, cohort, case-control, cross-sectional, and ecological. Note: Case reports and case series will be tracked during study screening but are not the primary focus of this assessment. They may be retrieved for full-text review and subsequent evidence synthesis if no or few more informative study designs are available. Case reports also can be used as supportive information to establish biologic plausibility for some target organs and health outcomes.
	Animal: Nonhuman mammalian animal species (whole organism) of any life stage (including preconception, in utero, lactation, peripubertal, and adult stages).
Exposure	Exposure based on administered dose or concentration, biomonitoring data (e.g., urine, blood, or other specimens), environmental, or occupational-setting measures (e.g., air, water levels), or job title or residence. Studies on natural uranium and depleted uranium will be included, studies on enriched uranium or those specific to radiation exposure from uranium will not be included. Mixture studies for animals will be included if they have an arm with a uranium compound only.
	Human and animal: Oral exposure will be examined. Other exposure routes, including dermal, inhalation, or injection, will be tracked during title and abstract as "supplemental information."
Comparator	*Human*: A comparison or reference population exposed to lower levels (or no exposure/exposure below detection levels) of uranium or to uranium for shorter periods.
	Animal: Quantitative exposure versus lower or no exposure with concurrent vehicle control group.
Outcomes	All noncancer health outcomes. In general, endpoints related to clinical diagnostic criteria, disease outcomes, histopathological examination, or other apical/phenotypic outcomes will be prioritized for evidence synthesis over outcomes such as biochemical measures.

[a] Evaluating individual mechanistic studies for uranium is not anticipated to be critical given the extent of the experimental animal evidence for noncancer outcomes and findings of earlier reviews. For mechanistic information, this assessment will primarily rely on other published authoritative sources, such as public health agency reports and expert review articles.

2.4. KEY SCIENCE ISSUES

Based on the preliminary literature survey, the following key scientific issues have been identified that warrant evaluation in this assessment.

- Uranium occurs in the environment in a variety of forms to which humans may be exposed, including metallic uranium, soluble uranium salts, and poorly soluble uranium compounds. In developing the IRIS assessment, consideration will be given to the approach used by ATSDR of providing toxicity values suitable for all soluble forms of uranium versus possible alternatives, addressing specific forms of uranium (e.g., more soluble versus poorly soluble versus insoluble species). Taking into account any new research, the assessment will develop and use a rationale for the specific categories of uranium compounds assessed.

From draft uranium IAP (2018)

IAP Can Include Literature Surveys

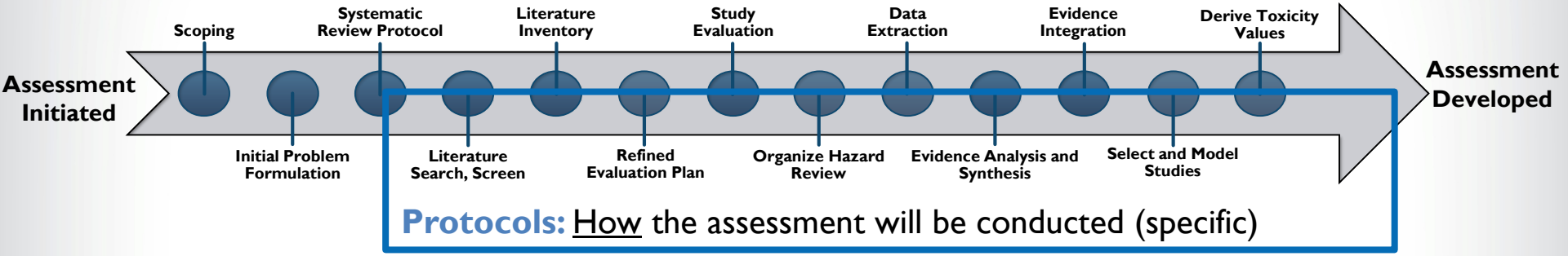
- Broad surveys to assess extent and nature of evidence, level of effort, type of expertise required
- Surveys inform decisions on targeted focus, e.g., evidence streams to consider core-PECO (versus supplemental), health outcomes likely covered in assessment
- Surveys may be developed based on other assessments, manual review of studies, or through use of specialized software applications

Nitrate/Nitrite (survey based on IARC 2010 and ATSDR 2017 assessments)

Outcomes	Human Studies				Animal Studies					
	Occupational epidemiology	General population epidemiology	Controlled exposure	Case reports and case series reports	Chronic	Subchronic	Short-term	Acute	Multi-generational	Gestational
Cancer		60			13					
Cardiovascular		1	1	3						
Dermal and ocular				1						
Developmental		14							2	6
Endocrine(thyroid)		6	1		4	3	1			
Gastrointestinal	1			7	5	1				
Hematological		25	3	10	4	6	3	1		
Hepatic					3			2		
Immunological										
Metabolic disease		8								
Musculoskeletal										
Neurological and sensory			1	6	1	1			1	
Renal					1					
Reproductive			3		2	2			1	
Respiratory										
Other					9	2	1		1	

The numbers represent the numbers of studies that investigated a particular health effect, and not the number of studies that identified a positive association with exposure.

Protocols



- Assessment specific stand-alone method documents that do not rely on the IRIS Handbook to convey methodology
- Comments received on IAP are considered when preparing the protocol (updated IAP text is included in the protocol)
- Released for 30-day public comment period (during Step 1 of IRIS Process)
- List of included, excluded, and studies tagged as supplemental will be disseminated through protocols (either during initial release or as an update)
- Protocol is iterative - Knowledge gained during implementation may result in revisions to the protocol to focus on the best available evidence. Major revisions are documented via updates, e.g., changes to specific aims or PECO

Protocol Content

3. OVERALL OBJECTIVES, SPECIFIC AIMS, AND POPULATIONS, COMPARATORS, EXPOSURE OUTCOMES (PECO) CRITERIA

The overall objective of this assessment is to identify adverse health effects and

Updated IAP text and PECO based on public comments

4. LITERATURE SEARCH AND SCREENING STRATEGIES

APPENDICES

APPENDIX A. ELECTRONIC DATABASE SEARCH STRATEGIES

5. REFINED EVALUATION PLAN

The evidence base for this assessment was relatively small and pul
assessment plan did not suggest a change was warranted to the specific ai
refined analysis plan was needed (i.e., all PECO-relevant studies will be co
assessment).

SU="CONSTRUCTION BUILDING TECHNOLOGY" OR SU="ASTRONOMY ASTROPHYSICS" OR SU="ARCHAEOLOGY" OR SU="OPERATIONS RESEARCH MANAGEMENT SCIENCE" OR SU="ANTHROPOLOGY" OR SU="SPORT SCIENCES" OR SU="ART" OR SU="PALEONTOLOGY" OR SU="TELECOMMUNICATIONS" OR SU="CHEMISTRY" OR SU="POLYMER SCIENCE" OR SU="ENGINEERING" OR SU="ENVIRONMENTAL SCIENCES ECOLOGY" OR SU="FOOD SCIENCE TECHNOLOGY" OR SU="SCIENCE TECHNOLOGY OTHER TOPICS" OR SU="BIOTECHNOLOGY APPLIED MICROBIOLOGY" OR SU="AGRICULTURE" OR SU="SPECTROSCOPY" OR SU="CRYSTALLOGRAPHY" OR SU="INTEGRATIVE COMPLEMENTARY MEDICINE" OR SU="WATER RESOURCES" OR SU="NUTRITION DIETETICS" OR SU="LIFE SCIENCES BIOMEDICINE OTHER TOPICS" OR SU="PARASITOLOGY" OR SU="THERMODYNAMICS" OR SU="OPTICS" OR SU="BIOPHYSICS" OR SU="TROPICAL MEDICINE" OR SU="VETERINARY SCIENCES" OR SU="RESEARCH EXPERIMENTAL MEDICINE" OR SU="MARINE FRESHWATER

6. STUDY EVALUATION (REPORTING, RISK OF BIAS, AND SENSITIVITY) STRATEGY

IRIS assessments evaluate each study's methods using uniform approaches for each group of similar studies s
concerns for the re
that affect the mag
study to detect a tr
animal toxicology s
supplemental mate
prominent role in t

Table 3. S

Epid
Exposure measurem
Outcome ascertainm
Participant selection
Confounding
Analysis
Selective reporting
Sensitivity

Study evalu
The study evaluati
limitations (focusi
result), considerin
null. The study eva
of the results) in th

7. DATA EXTRACTION OF STUDY METHODS AND RESULTS

Data extraction an
elements that may be colle
Choices about what data t
analyses that inform the s
following the identificatio
the data extraction workfl
extraction. Studies evalua
therefore, will not be cons
be less relevant during PE
minimal data extraction.
high confidence studies ar

The data extractio
available for download fr
[NOTE: The following bro
(preferred), Mozilla Foxfi
Internet Explorer.] Data e
independently checked by
by discussion or consultat
verified, they will be "lock
WebPlotDigitizer (http://
information from figures.

8. PHYSIOLOGICALLY BASED PHARMACOKINETIC (PBPK) MODEL IDENTIFICATION, DESCRIPTIVE SUMMARY, AND EVALUATION

PBPK (or classical pharmacokinetic [PK]) models should be used in an assessment when an applicable one exists and no equal or better alternative for dosimetric extrapolation is available. Any models used should represent current scientific knowledge and accurately translate the science into computational code in a reproducible, transparent manner. For a specific target organ/tissue, it may be possible to employ or adapt an existing PBPK model, or develop a new PBPK model or an alternate quantitative approach. Data for PBPK models may come from studies with animals or humans, and may be in vitro or in vivo in design.

8.1. IDENTIFYING PHYSIOLOGICALLY BASED PHARMACOKINETIC (PBPK) MODELS

PBPK modeling is the preferred approach for calculating a human equivalent concentration (HEC) according to the hierarchy of approaches outlined in EPA guidance (U.S. EPA, 2011a). For chloroform, metabolism is a major component of target organ toxicity, and PBPK models are available to account for interspecies differences in metabolism between rats, mice, and humans (Sasso et al., 2013; Corley et al., 1990). Chloroform is metabolized to the reactive metabolites phosgene and dichloromethyl free radical in humans and animals by cytochrome P450-dependent pathways (Gemma et al., 2003; Constan et al., 1999).

Because of the role of metabolism in the production of target organ toxicity, and the reactive

From draft chloroform protocol (2018)

Protocol Content

9. SYNTHESIS WITHIN LINES OF EVIDENCE

For each potential health effect
outcomes; or a broad hazard category),
effect evidence, ar
written to empha
the evidence inte
studies or group c
association, tempo
humans (U.S. EPA
Specificall
first be analyzed a
lack of data withir
the available mec
chloroform, a syn
evaluation of carc

9.1. SYNTHES

To assess

Table 9. Primar
syntheses[a]

Consideration	
Consistency	Repeated exist, the "differing" *Stronger h* *Stronger c*
Biological gradient (dose-response)[b]	Increases concentra or comple necessaril considere
Strength (effect magnitude) and precision	Given wha particular small effe may consi other expl errors and results acr (i.e., low p
Mechanistic evidence related to biological plausibility	Supporting effects; changes in established biol evidence strength. While a lack of strength, it may do so if findings de *Human evidence:* studies in expose *Animal evidence:* studies in exposed animals
Coherence[c]	Findings across the database that fit into a co similarity in results for related effects within a dose-dependent progression of linked effects Conversely, an observed lack of changes that subsequently) with the effect of interest coul informed by the known biological developme toxicokinetic/dynamic understanding of the c
Natural experiments	*Human evidence only:* Reductions in effect thi Although rare, such reductions can provide co
Temporality	*Human evidence only:* The exposure occurs be evaluation of exposure measures for each stu

10. INTEGRATION ACROSS LINES

For the analysis of most health outcomes, IRIS assessme
and mechanistic evidence. Depending on the assessment scope a
animal evidence, conclusions for mechanistic evidence may be b
mechanistic st
are drawn as f

- First, a
chemi
step in
cohere
- In para
the che

WITHIN STREAM CONCLUSIONS

HUMAN EVIDENCE STREAM CONCLUSION

The synthesis of evidence about health effects and mechanisms from human studies is combined (integrated) to draw a conclusion about effects within the stream

Studies and interpretation	Factors that increase confidence	Factors that decrease confidence	Summ
[Health Effect or Outcome Grouping]			
Evidence from Human Studies (Route)			
· References · Study confidence (based on evaluation of risk of bias and sensitivity) and explanation · Study design description	· Consistency · Dose-response gradient · Coherence of observed effects (apical studies) · Effect size (magnitude, severity) · Biological plausibility · Low risk of bias/ high quality · Insensitivity of null/ negative studies · Natural experiments · Temporality	· Unexplained inconsistency · Imprecision · Indirectness/ applicability · Poor study quality/ high risk of bias · Other (e.g., Single/Few Studies; small sample size) · Evidence demonstrating implausibility	*Results infor affected/ una Human evid plausibility: d data influenc judgement (precursors in*
			Could be multiple rows (e.g study confidence or popul informs results hetero
Evidence for an Effect in Animals (Route)			
· References · Study confidence (based on evaluation of risk of bias and sensitivity) and explanation · Study design description	· Consistency and Replication · Dose-response gradient · Coherence of observed effects (apical studies) · Effect size (magnitude, severity) · Biological plausibility · Low risk of bias/ high quality · Insensitivity of null/ negative studies	· Unexplained inconsistency · Imprecision · Indirectness/ applicability · Poor study quality/ high risk of bias · Other (e.g., Single/Few Studies; small sample size) · Evidence demonstrating implausibility	*Results information (gene affected/ unaffected) acro Evidence informing biolo plausibility for effects in a discuss how mechanistic influenced the within stream judgement (e.g., evidence of coherent molecular changes in animal studies)* *Could be multiple rows (e.g., by study confidence, species, or exposure duration) if this informs results heterogeneity*

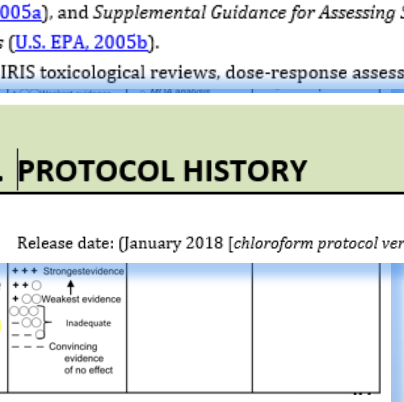


Figure 4. Evidence profile table template.

11. DOSE-RESPONSE ASSESSMENT: STUDY SELECTION AND QUANTITATIVE ANALYSIS

The previous sections of this protocol describe how systematic review principles are applied to support transparent identification of health outcomes (or hazards) associated with exposure to the chemical of interest in conjunction with evaluation of the quality of the studies considered during hazard identification. Selection of specific data for dose-response assessment and performance of the dose-response assessment is conducted after hazard identification is complete, and builds off this step in developing the complete IRIS assessment. The dataset selection process involves database- and chemical-specific biological judgments that are beyond the scope of this protocol, but are discussed in existing EPA guidance and support documents. This section of the protocol provides an overview of points to consider when conducting the dose-response assessment, particularly statistical considerations specific to dose response analysis that support quantitative risk assessment. Importantly, the considerations outlined in this protocol do not supersede existing EPA guidance. Several EPA guidance and support documents provide more detailed considerations for the development of EPA's traditional dose-response values, especially EPA's *Review of the Reference Dose and Reference Concentration Processes* (U.S. EPA, 2002), EPA's *Benchmark Dose Technical Guidance* (U.S. EPA, 2012b), *Guidelines for Carcinogen Risk Assessment* (U.S. EPA, 2005a), and *Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens* (U.S. EPA, 2005b).

For IRIS toxicological reviews, dose-response assessments are typically performed for both

12. PROTOCOL HISTORY

Release date: (January 2018 [*chloroform protocol version 1*])

From draft chloroform protocol (2018)

Publicly Available Examples

Assessment Plans

September 27-28, 2017
- Chloroform
- Nitrate/nitrites
- Ethylbenzene

January 26, 2018
- Uranium

- **Targeted focus:** chloroform, uranium, chloroprene
- **Modularity:** ethylbenzene
- **Use of existing assessments conducted by others:** nitrate/nitrate, uranium (ATSDR assessments)

- IAPs and/or protocols will be released for most in-progress assessments
  - Which document is released depends on extent of refinement in scope compared to previous public sharing and maturity of the draft assessment

Protocol

January 26, 2018
- Chloroform (includes list of included studies)

Rapid systematic review

- EPA response to the Chloroprene Request for Correction (posted January 29, 2018)

EPA
United States
Environmental Protection
Agency

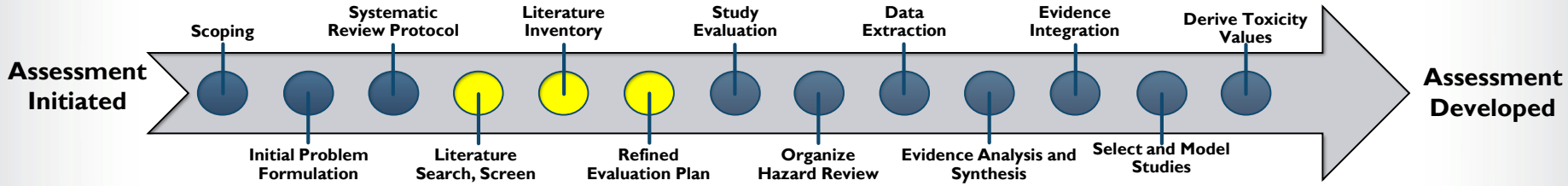
Literature Searching, Screening, and Inventories*

Office of Research and Development
NCEA, IRIS

* includes basic methodological details

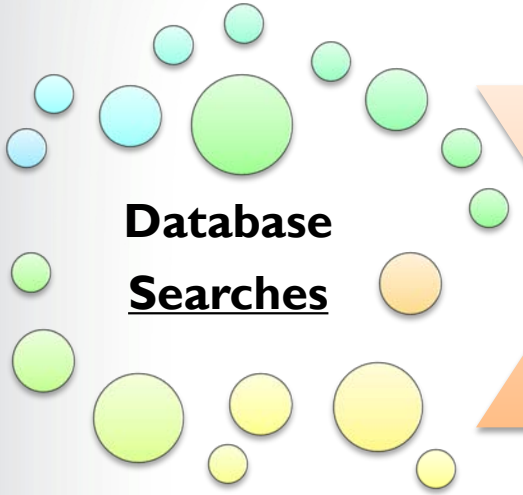
NAS 2014: High Priority (Box 8-1) Recommendations

"...include a section on evidence identification that is written in collaboration with information specialists trained in systematic reviews and that includes a search strategy for each systematic-review question being addressed in the assessment. Specifically, the protocols should provide a line-by-line **description of the search strategy, the date of the search, publication dates searched, and explicitly state the inclusion and exclusion criteria**..."



- **Protocols outline the specifics of the literature search and screening approaches, including inclusion and exclusion criteria in PECO tables**
- **Dedicated information technologists help formulate searches, and screening decisions are tracked in HERO (tagging)**
- **Manual and semi-automated approaches are being used to identify relevant studies**
- **Inventories of basic study methods organize evidence for refinement and evaluation**
- **Changes and updates are documented in the protocol**

Routine Evidence Identification Processes



Screening

1. Title/abstract
2. Full text

Inventories

Health Outcome & PBPK Studies

- Tag studies by line of evidence and outcome
- Distribute to disciplinary experts for review

Supplemental Studies

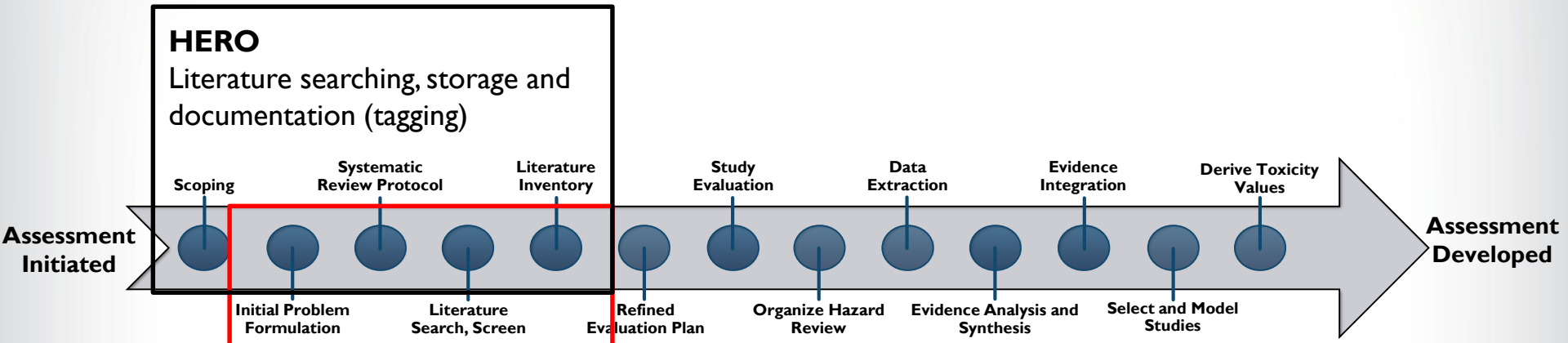
- Includes in vitro and other mechanistic evidence (e.g., non-PECO exposure route; non-PECO animal model; toxicokinetic data)

- Inventories contain basic study methods for evaluation and prioritization decisions

- Identify peer-reviewed and "gray" (unpublished) literature
- PubMed, ToxLine, and Web of Science are standard (others can be included as needed)
- Conduct regular search updates
- Details of search strategy, dates, and retrieved records are presented in protocols and assessments

- Use manual and automated approaches
- ≥ 2 screeners
- Tag studies as excluded, meeting PECO criteria, or supplemental information
- Screening decisions available in HERO
- Typically do not apply language-restrictions
- Review reference list of included studies and relevant reviews to identify studies missed from database searches
- Share list of included studies with public to further ensure all relevant studies included

Use of Specialized Software Tools for Literature Search and Screening



SWIFT Review
Problem formulation

SWIFT Review
Screening prioritization

Distiller (manual)
SWIFT Active (SWIFT is a semi-automated/ machine learning tool)
Multiple reviewer reference screening and tracking (HERO-tagging)

- **Software tools will be discussed in Session 3 and during demo session**
- **Tools are being developed and applied through testing**
- **Evergreen - new tools compatible with HERO will be added as performance is characterized**

Evidence Identification in Protocols

4. LITERATURE SEARCH AND SCREENING STRATEGIES

4.1. USE OF EXISTING ASSESSMENTS

Describe any use of existing assessments that serve as starting points for the literature search.

4.2. LITERATURE SEARCH STRATEGIES

Literature search strategies were developed using key terms and words related to the PECO statement. Developme
relevant search terms th
relevant and appropriate
previously identified pri
and (3) reviewing search
words were crafted into
specificity of the search
that 100% of the previ
database has its own sea
each database's unique s

The following databases

- PubMed
- Web of Science

4.4. SCREENING PROCESS

Studies that comply with the criteria specified in the PECO Statem
inclusion while those that do not meet these criteria will be excl
the exclusion criteria noted below will be applied. However, the
will be revie
searching.

- Reco
- Studi
these
pape
- [othe

Studi
application a

4.5. LITERATURE SURVEYS AND SUM

During title/abstract or full-text screening,
features such as
health outcome
ADME, etc.). Th
evidence for pri
assessment as s
epidemiological
inhalation, mixt
anesthetic, stud
or exposure to
supporting info
inventory of stu
for each potenti
generation and
biological proce

special topics

4.3. UNPUBLISHED DATA

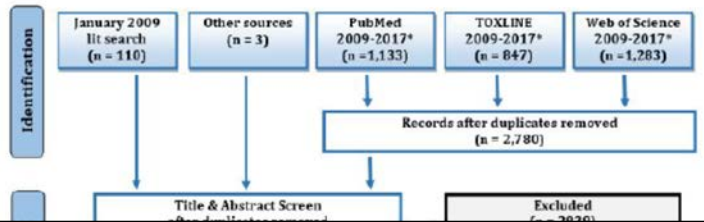
IRIS only includes publicly accessible, peer-reviewed information in its evaluations. However, it is possible that unpublished data directly relevant to the PECO statement may be identified during the course of the assessment. In this case, EPA is able to obtain external peer review if the owners of the data are willing to have the study details and results made publicly accessible. The

4.4.1. Multiple publications of the same data

Multiple publications with overlapping data for the same study (e.g., publications reporting subgroups, additional outcomes or exposures outside the scope of an evaluation, or longer follow-up) can be identified by examining author affiliations, study designs, cohort name, enrollment criteria, and enrollment dates. If necessary, study authors will be contacted to clarify any uncertainty about the independence of two or more articles. IRIS will include all publications on the study, select one study to use as the primary, and consider the others as secondary publications with annotation as being related to the primary record during data abstraction. The primary study

4.6. TRACKING STUDY INFORMATION

The main reason for exclusion at the full-text review stage was annotated and reported in a literature flow diagram (see Figure 2). Categories for exclusion included the following: (1) not relevant to PECO; (2) review, commentary, or letter with no original data; (3) conference abstract or thesis (and the criteria for including unpublished data, described above, were not met); or (4) unable to obtain full-text.



would include any specialized software tools

PECO Criteria to Identify Studies

PECO element	Evidence
Populations[a]	*Human*: Any population and lifestage (occupational or general population, including children and other sensitive populations). The following study designs will be considered most informative: controlled exposure, cohort, case-control, cross-sectional, and ecological. Note: Case reports and case series will be tracked during study screening, but are not the primary focus of this assessment. They may be retrieved for full-text review and subsequent evidence synthesis if no or few informative study designs are available. Case reports can also be used as supportive information to establish biologic plausibility for some target organs and health outcomes.
Exposures	Studies of chloroform in the context of its use as an anesthetic gas will be excluded.
	Animal: Any exposure to chloroform via inhalation. Studies employing chronic exposures or short-term, developmental-only exposures will be considered the most informative. Studies involving exposures to mixtures will be included only if they include an arm with exposure to chloroform alone. Studies utilizing chloroform as an extraction solvent to isolate specific chemical constituents will be excluded.
	Studies describing physiologically-based pharmacokinetic (PBPK) models for chloroform will be included.
Comparators	*Human*: A comparison or referent population exposed to lower levels (or no exposure/exposure below detection limits) of chloroform, or exposed to chloroform for shorter periods of time.
	Animal: A concurrent control group exposed to vehicle-only treatment.
Outcomes	All health outcomes (both cancer and noncancer). In general, endpoints related to clinical diagnostic criteria, disease outcomes, histopathological examination, or other apical/phenotypic outcomes will be prioritized for evidence synthesis over outcomes such as biochemical measures. As discussed above, based on preliminary screening work, EPA anticipates that a systematic review for health effect categories other than those identified (i.e., nasal cavity effects, nervous system effects, liver and kidney effects, immunotoxic effects, and reproductive/developmental effects) will not be undertaken unless a significant amount of new evidence is found upon review of references during the comprehensive literature search.

[a] Evidence from in vitro, in silico, and other types of mechanistic studies will be prioritized based on likelihood to impact evidence synthesis conclusions for human health. For chloroform, mechanistic studies will only be considered for evaluation if they have the potential of impacting the existing 2001 MOA analysis, or are essential for answering questions identified during the human and animal evidence syntheses.

Example from the draft chloroform protocol

Example Literature Screening Form

SUBMIT FORM and go to [ ] or Skip to Next

Forms Independently Entered by 2 Reviewers

1. **Based on Title and Abstract does the article contain relevant human, animal, or in vitro evidence?**

   ◯ Yes ◯ No ◉ No, but has supportive information
   ◯ Unclear (e.g., no abstract) **Clear Response**

2. **What kind of evidence or supportive information?**

   - ☐ human
   - ☐ animal
   - ☑ in vitro, omics, alternative model systems

3. **What kind of supportive information?**

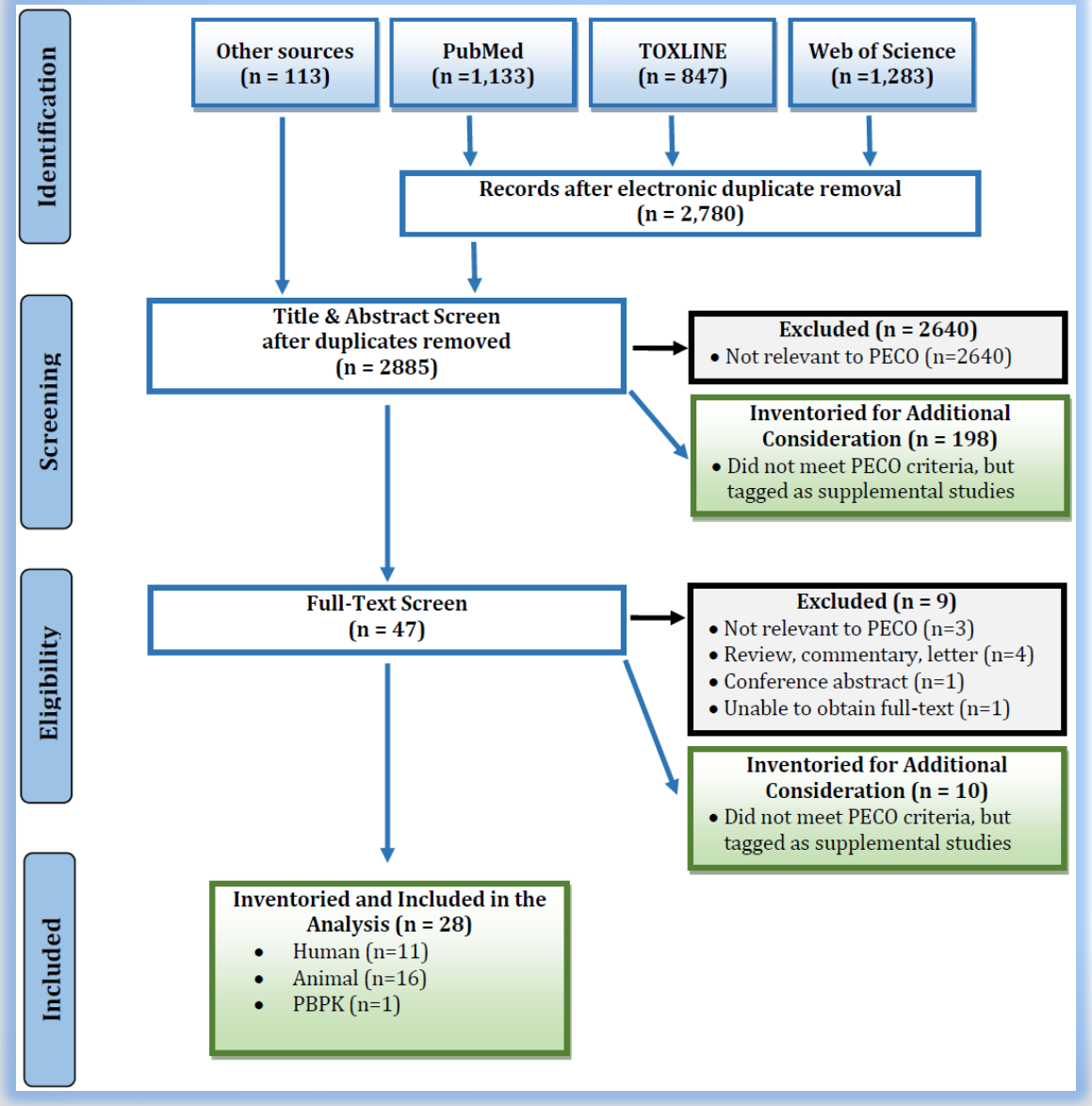
   - ☑ MOA/mech (cancer)
   - ☐ MOA/mech (non-cancer)
   - ☐ case report or poisoning
   - ☐ non-inhalation route
   - ☐ mixture
   - ☐ ADME/PBPK
   - ☐ exposure assessment
   - ☐ THM, disinfection/chlorination by-product, swimming pools
   - ☐ susceptible population
   - ☐ anesthesia/inhalent

P	Human: Any population and life stage (occupational or general population, including children and other sensitive population). The following study designs will be considered most informative: controlled exposure, cohort, case-control, cross-sectional, and ecological. Note: Case reports and case series will be tracked during study screening, but are not the primary focus of this assessment. They may be retrieved for full-text review and subsequent evidence synthesis if no or few informative study designs are available. Case reports can also be used as supportive information to establish biologic plausibility for some target organs and health outcomes. Animal: Nonhuman mammalian animal species (whole organism) of any life stage (including preconception, in utero, lactation, peripubertal, and adult stages).
E	Human: Any exposure to chloroform, including occupational exposures, via inhalation. Exposures quantified by either actual exposure measurements or occupational exposure history are preferred. Studies of chloroform in the context of its use as an anesthetic gas will be excluded Animal: Any exposure to chloroform via inhalation. Studies employing chronic exposures or short-term, developmental-only exposures will be considered the most informative. Studies involving exposures to mixtures will be included only if they include an arm with exposure to chloroform alone. Studies utilizing chloroform as an extraction solvent to isolate specific chemical constituents will be excluded. PBPK: Studies describing physiologically-based pharmacokinetic (PBPK) models for chloroform will be included.
C	Human: A comparison or reference population exposed to lower levels (or no exposure/exposure below detection limits) of chloroform, or exposed to chloroform for shorter periods of time. Animal: A concurrent control group exposed to vehicle-only treatment.
O	All health outcomes (both cancer and noncancer). In general, endpoints related to clinical diagnostic criteria, disease outcomes, histopathological examination, or other apical/phenotypic outcomes will be prioritized for evidence synthesis over outcomes such as biochemical measures. As discussed above, based on preliminary screening work, EPA anticipates that a systematic review for health effect categories other than those identified (i.e., nasal cavity effects, nervous system effects, liver and kidney effects, immunotoxic effects, and reproductive/developmental effects) will not be undertaken unless a significant amount of new evidence is found upon review of references during the comprehensive literature search.

SUBMIT FORM and go to [ ] Skip to Next

Draft example based on chloroform using Distiller

Tracking: Literature Flow Diagrams



- **Track rationale for full-text exclusions**
- **Use HERO to share repositories of included, excluded, and supplemental studies**

Example modeled on the draft chloroform protocol

Literature Inventories

Example Details Routinely Extracted (*female reproductive toxicity in animals*):

- Outcome category (e.g., fertility) and/or Specific endpoint (e.g., number of litters)
- Species (e.g., rat; alternative [nonmammalian] animal)
- Exposure duration (e.g., chronic; multi-generational; gestational)
- Exposure route (e.g., oral [gavage]; in vitro)

Assessment-Specific Extraction Details (*generic examples*):

- Exposure levels tested
- Test article details, such as purity or isomeric composition

Results are Typically <u>Not</u> Included in Inventories

Developing Extraction Forms *(all 3 lines of evidence) to be interoperable with HAWC*

Test compound	DBP	Maternal age at init. gestational exp.	NA
Species	Rat	Subject age at in vivo exposure initiation	PND21
Exposure type	In vivo	Lifestage at in vivo exposure initiation	Weaning
In vivo exposure route	Oral-gavage	Dose/concentration	0, 500 mg/kg-day
Strain of exposed test model	Sprague-Dawley	Exposure duration	Single exposure
Sex of exposed test model	M	Exposure period	Postnatal only

Refined Evaluation Plan (optional)

<u>Discipline-specific experts</u> consider whether and how to further refine or prioritize studies/outcomes for evaluation (based on study design features)

- ***Health effect studies meeting PECO criteria (e.g., organized by outcome):***
  - Considers ADME and other key science issues (supplemental studies reviewed)
  - Opportunity to discuss outcome grouping (e.g., based on known biology/MOA) and handling of key science issues during outcome-specific study evaluations
  - Studies with certain design features or specific outcomes may be selected or prioritized for evaluation and synthesis (e.g., based on exposure duration, administration, or levels tested; or endpoint specificity)
- ***Supplemental mechanistic studies (e.g., organized by test system, mechanistic event, or key characteristic [of carcinogens]) are considered iteratively:***
  - Identifies other studies on specific aim mechanistic questions (e.g., mutagenicity)
  - Organizes the available evidence to allow for pragmatic evaluations of key issues that arise during review of PECO-specific human and animal studies (Session 2)

Refinements are tracked and updated in the assessment protocol

IRIS has Addressed the Major NAS 2014 Recommendations

NAS 2014 Topics	IRIS Process Improvements
General Process Issues (Chapter 2); Problem Formulation and Protocol Development (Chapter 3)	• Draft IRIS Handbook of program SOPs is being reviewed within EPA • IAPs allow early comment on problem formulation • More frequent Agency engagement facilitates scope refinement • Assessment protocols describe methods and allow for iteration • Re-occurring staff training and template IAPs and protocols promote consistency and quality control
Evidence Identification (Chapter 4)	• Consultation with information technologists and subject experts • Adopts current systematic review best practices, including use of specialized tools • Transparent documentation (e.g., literature flow diagrams)

See Demonstrations:

- Sciome Workbench for Interactive computer-Facilitated Text mining (SWIFT Review and SWIFT Active)
- Health Assessment Workspace Collaborative (HAWC)
- Heath Effects Research Online (HERO)

EPA
United States
Environmental Protection
Agency

SESSION 2: SYSTEMATIC REVIEW IN THE IRIS PROGRAM- EVIDENCE EVALUATION

Xabier Arzuaga*, Catherine Gibbons*, Barbara Glenn*, Andrew Kraft*, Beth Radke*, Kris Thayer

[*Speaking]

Office of Research and Development
NCEA, IRIS

EPA
United States
Environmental Protection
Agency

Evaluating Individual Studies: Reporting Quality, Risk of Bias, and Sensitivity

Office of Research and Development
NCEA, IRIS

NAS 2014 High Priority (Box 8-1) Recommendations on Evidence Evaluation

"When considering any method for evaluating individual studies, EPA should **select a method that is transparent, reproducible, and scientifically defensible**. Whenever possible, there should be empirical evidence that the methodologic characteristics that are being assessed in the IRIS protocol have systematic effects on the direction or magnitude of the outcome."

"EPA should **specify the empirically based criteria it will use to assess risk of bias for each type of study design in each type of data stream.**"

"To maintain transparency, EPA should **publish its risk-of-bias assessments as part of its IRIS assessments**."

Study Evaluation – Developing an Approach

- Considered and drew from existing tools for study evaluation.
- Developed approaches for both epidemiology and toxicity studies that:
  - Addresses study sensitivity and identifies potential sources of bias.
  - Transparently presents the criteria/considerations used to consistently evaluate and judge each study/outcome.
  - Provides access to the rationale for discipline-specific decisions made during the evaluation process.
- Objective of the approach: Identify the most informative and reliable studies for evidence synthesis and integration.

PBPK Model Evaluation

Criteria	Example information
Scientific	Biological basis for the model is accurate • e.g., Predicts dose metrics expected to be relevant
	Consideration of model fidelity to the biological system strengthens the scientific basis relative to standard extrapolation (default) approaches • e.g., Can the model describe critical behavior, such as nonlinear kinetics in a relevant dose range, better than the default (i.e., $BW^{3/4}$ scaling)?
	Principle of parsimony (i.e., model complexity or biological scale should be commensurate with data available to identify parameters)
	Model describes existing PK data reasonably well, both in "shape" (e.g., matches curvature) and quantitatively (e.g., within a factor of 2–3)
	Model equations are consistent with biochemical and biological understanding
Initial technical	Well-documented model code is readily available to EPA and public
	Set of published parameters clearly identified, including origin/derivation
	Parameters do not vary unpredictably with dose • e.g., Any dose dependence in absorption constants is predictable across the dose ranges relevant for animal and human modeling
	Sensitivity and uncertainty analysis has been conducted for relevant exposure levels (local sensitivity analysis is sufficient, though global preferred) • e.g., A sound explanation should be provided when sensitivity of the dose metric to model parameters differs from what is reasonably expected

Prior to use, relevant PBPK models will:

- Be thoroughly evaluated based on scientific and technical criteria (examples to the left).
- Undergo QA/QC on model equations, parameters (including primary/secondary sources), and model code.

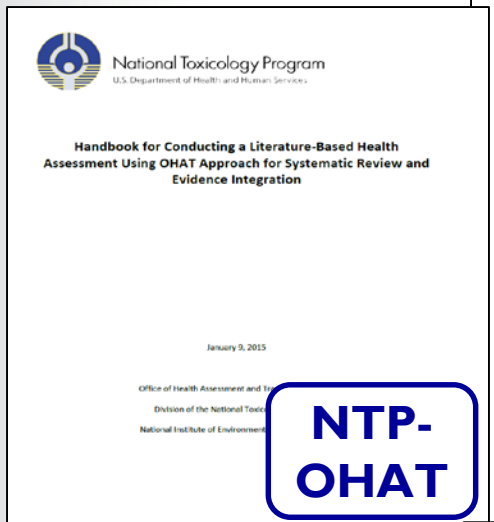
For details, please see:

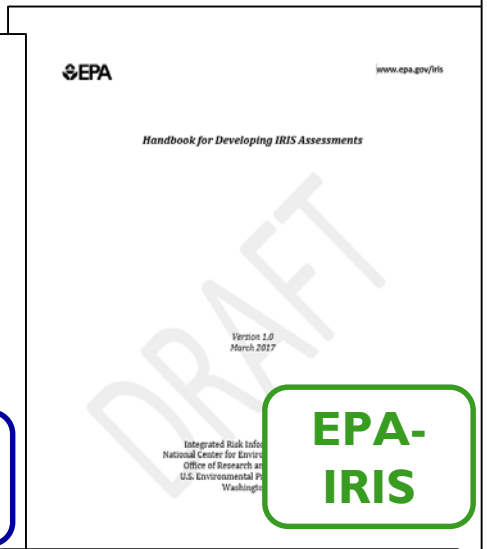
- Poster:

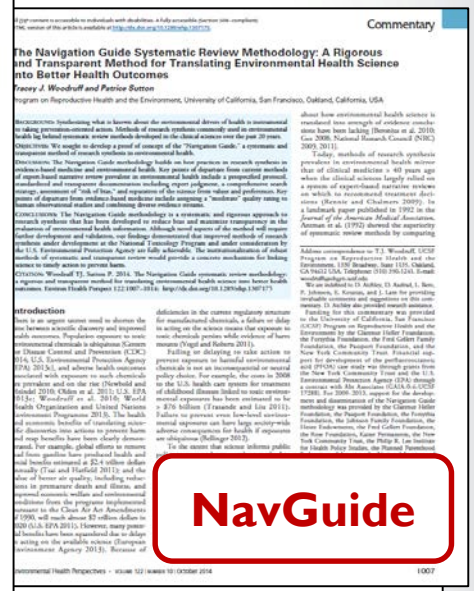
  Systematic evaluations of PBPK models for human health risk assessment
- EPA website:

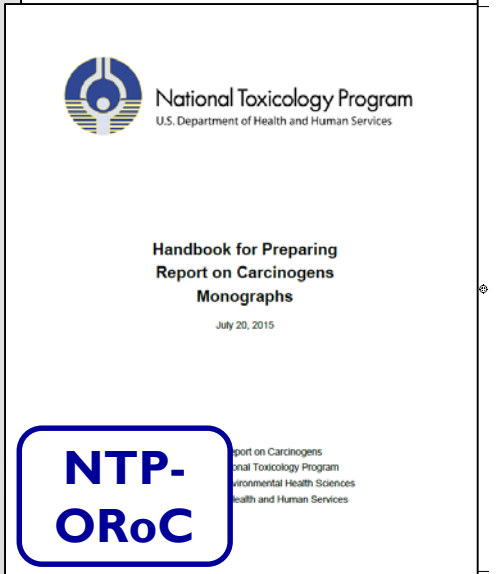
  EPA Response to the Request for Correction of the IRIS Toxicological Review of Chloroprene (2018)

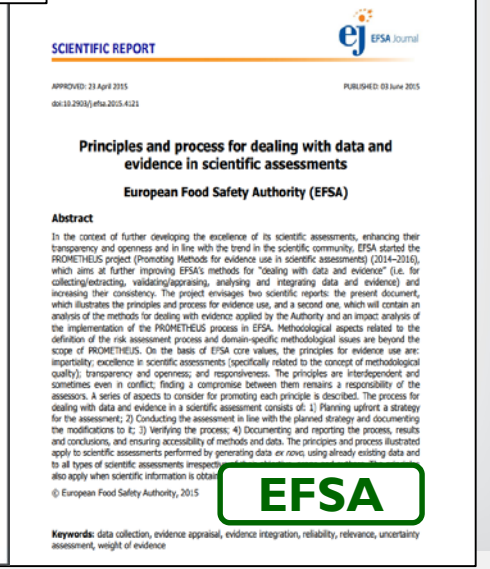
Evolving Approaches



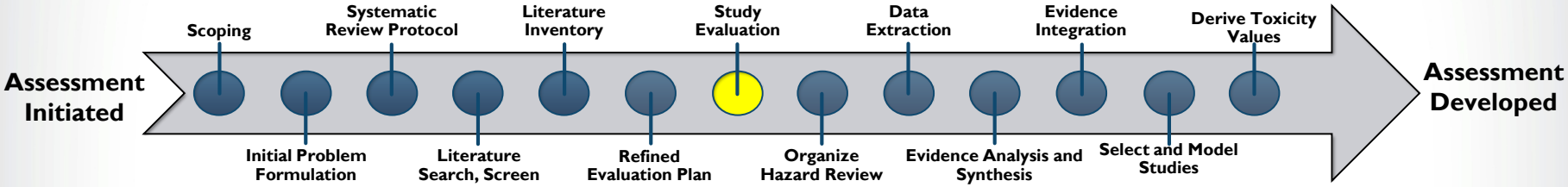








Evaluation of Individual Health Effect Studies



- General approach same for human and animal studies
- Evaluation process focused on:
  - Internal validity/bias
  - Sensitivity
  - Reporting quality

Development of Evaluation Strategies

- Questions in IRIS Protocol Template highlight general study attributes or elements to consider
- Subject-matter knowledge is used to formulate a list of issues to consider in the evaluation
- Develop a set of considerations based on exposure and outcome-specific knowledge

Study Evaluation Overview of Epidemiological and Animal Toxicity studies

Individual study level domains	
Animal	Epidemiological
Reporting Quality	Exposure measurement
Selection or Performance Bias	Outcome ascertainment
Confounding/Variable Control	Population Selection
Reporting or Attrition Bias	Confounding
Exposure Methods Sensitivity	Analysis
Outcome Measures and Results Display	Sensitivity
	Selective reporting

Domain Judgment	
++	Good
+	Adequate
-	Poor
--	Critically Deficient

Overall Study Rating
High
Medium
Low
Uninformative

Individual Domain Ratings for Epidemiological and Animal Toxicity Studies

	IRIS Judgment	How to interpret
++	Good	Appropriate study conduct relating to the domain & minor deficiencies not expected to influence results.
+	Adequate	A study that may have some limitations, but not likely to be severe or to have a notable impact on results.
-	Poor	Identified biases or deficiencies interpreted as likely to have had a notable impact on the results or prevent reliable interpretation of study findings.
--	Critically Deficient	A judgment that the study conduct relating to the domain introduced a serious flaw that is interpreted to be the primary driver of any observed effect or makes the study uninterpretable. Study is not used without exceptional justification.

Overall Study Confidence Ratings for Epidemiological and Animal Toxicity Studies

Rating	Description
High	No notable deficiencies or concerns identified; potential for bias unlikely or minimal and sensitive methodology.
Medium	Possible deficiencies or concerns noted, but resulting bias or lack of sensitivity would be unlikely to be of a notable degree.
Low	Deficiencies or concerns were noted, and the potential for substantive bias or inadequate sensitivity could have a significant impact on the study results or their interpretation.
Uninformative	Serious flaw(s) makes study results unusable

General Considerations to Evaluate Outcomes from Animal Toxicology Studies

Domain	Metric
Reporting Quality	Reporting of information necessary for study evaluation
Selection or Performance Bias	Allocation of animals to experimental groups
	Blinding of investigators, particularly during outcome assessment
Confounding/Variable Control	Control for variables across experimental groups
Reporting or Attrition Bias	Lack of selective data reporting and unaccounted for loss of animals
Exposure Methods Sensitivity	Characterization of the exposure to the compound of interest
	Utility of the exposure design for the endpoint of interest
Outcome Measures and Results Display	Sensitivity and specificity of the endpoint evaluations
	Usability and transparency of the presented data

Epidemiology Study Evaluation

- Approach based on the Cochrane Risk of Bias in Nonrandomized Studies of Interventions (ROBINS-I)[1], modified for environmental and occupational exposures
- Start by considering an "ideal" study for each domain, identifying "critical deficiencies", then developing criteria to define other levels of confidence
- Emphasis is on discerning bias that would produce a substantive change in the estimated effect estimate.

[1]Sterne, Hernan, et al. ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. BMJ 2016; 355:i4919.

Epidemiology Evaluation Domains

Domain	Core Question
Exposure measurement	Does the exposure measure reliably distinguish between levels of exposure in an appropriate time window?
Outcome ascertainment	Does the outcome measure reliably distinguish the presence or absence (or degree of severity) of the outcome?
Population selection	Is there evidence that selection into or out of the study (or analysis sample) was jointly related to exposure and outcome?
Confounding	Is confounding of the effect of the exposure likely?
Analysis	Does the analysis strategy and presentation convey the necessary familiarity with the data and assumptions?
Sensitivity	Are there concerns for study sensitivity?

Example of Considerations by Domains

Domain	Core Question
Exposure measurement	Does the exposure measure reliably distinguish between levels of exposure in an appropriate time window?

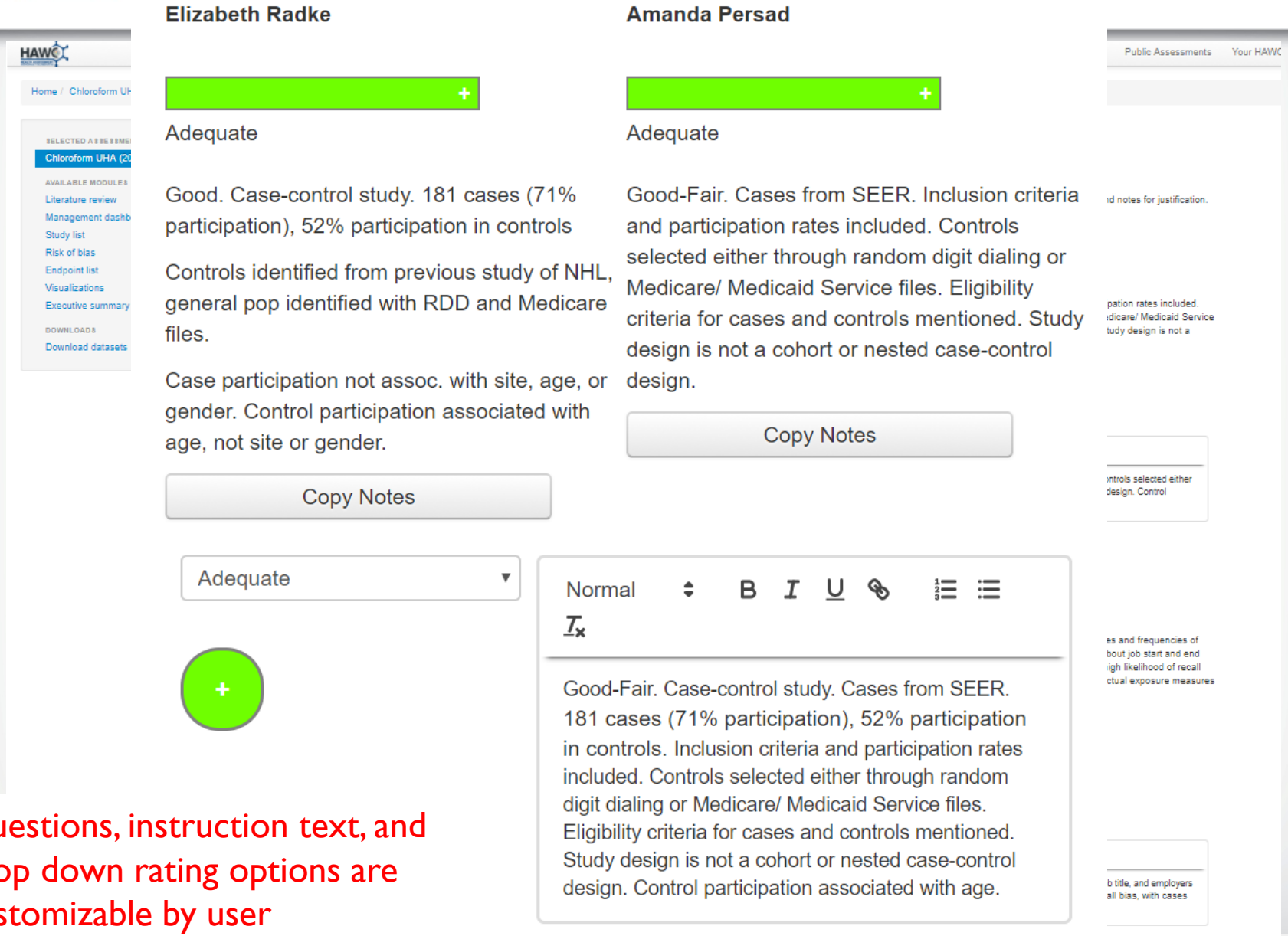
Examples of Prompting Questions:

- Does the exposure measure capture the variability in exposure among the participants, considering intensity, frequency, and duration of exposure?
- Does the exposure measure reflect a relevant time window?
- Was exposure measurement likely to be affected by knowledge of outcome or by presence of the outcome (i.e., reverse causality)?

Examples of Follow-up Questions:

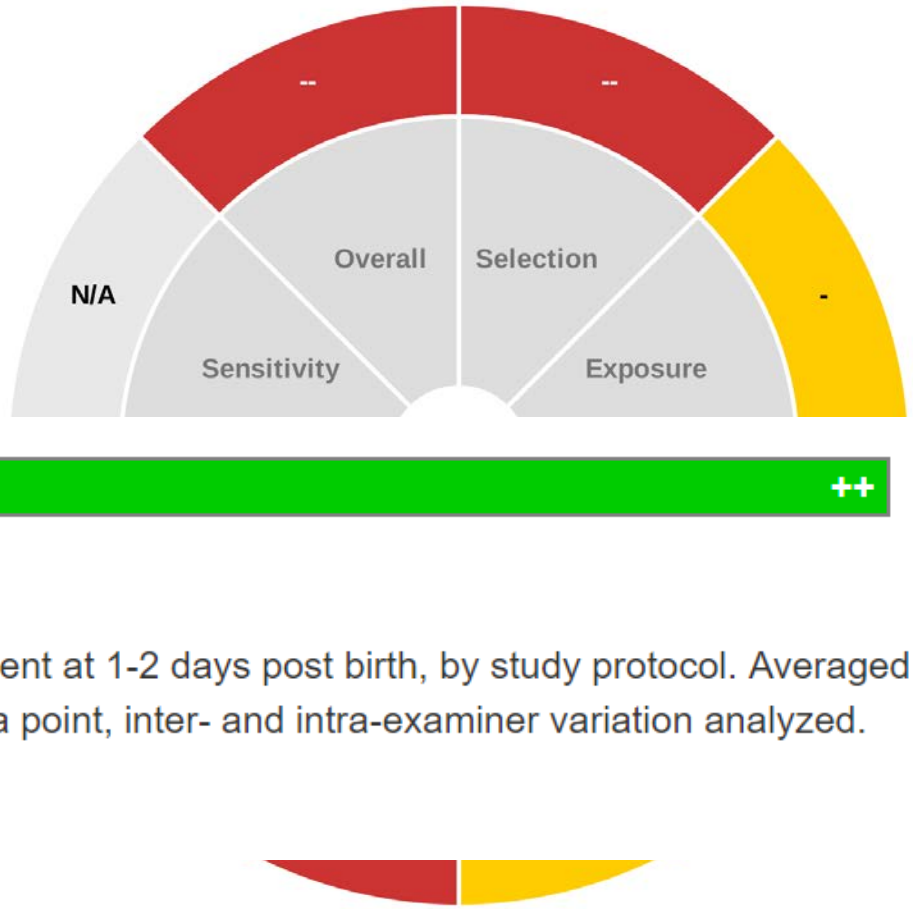
- Is the degree of exposure misclassification likely to vary by exposure level?
- If there is a concern about the potential for bias, what is the predicted direction of the bias on the effect estimate?

Study Evaluation: Final Review in HAWC



Questions, instruction text, and drop down rating options are customizable by user

Individual Studies in HAWC



++

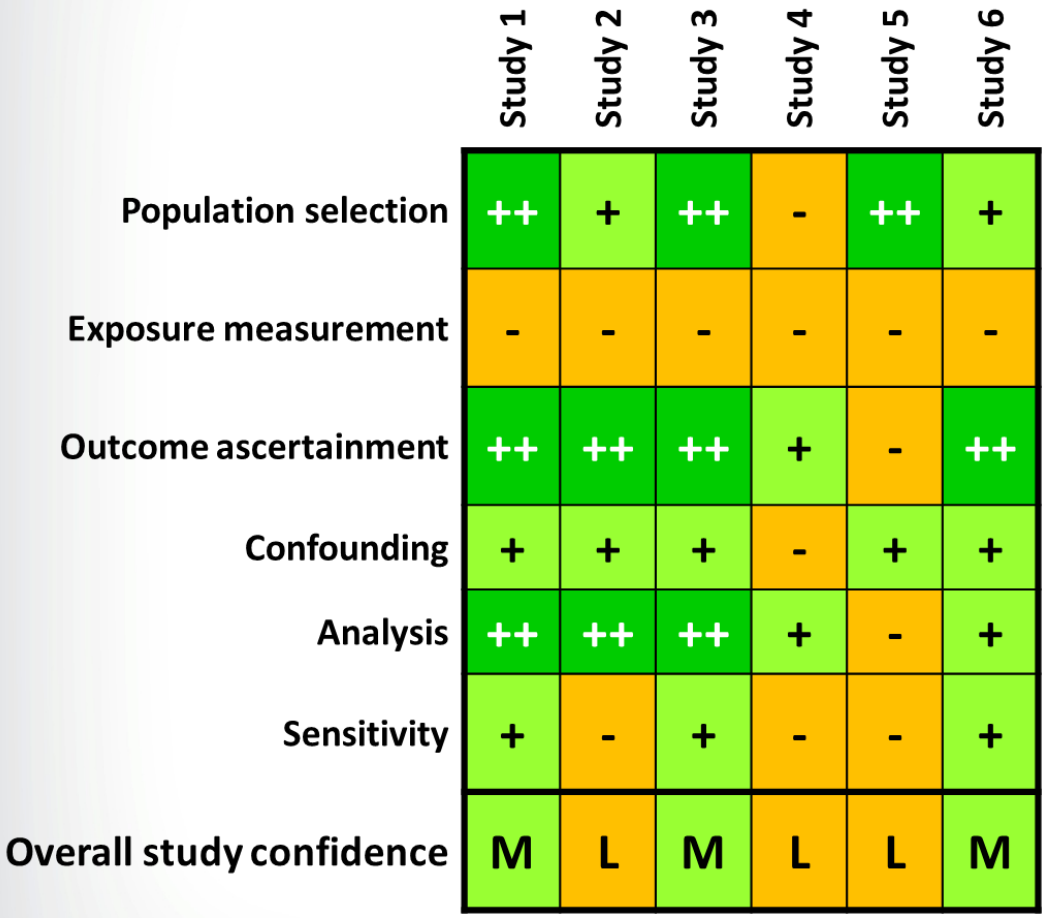
Good

Very Good. AGD measurement at 1-2 days post birth, by study protocol. Averaged 3 measurements for each data point, inter- and intra-examiner variation analyzed.

Medium confidence

Uninformative

Study Evaluation Summary in HAWC



Legend

Domain judgement		Overall study rating	
++	Good	H	High confidence
+	Adequate	M	Medium confidence
NR	Not reported		
-	Poor	L	Low confidence
--	Critically deficient	U	Uninformative
N/A	Not applicable		

Publicly available examples

- Initial and iterative improvements to study evaluation
  - Ammonia, Inhalation (final 2016)
  - RDX (peer review draft 2016)
  - TBA (peer review draft 2017)
  - ETBE (peer review draft 2017)

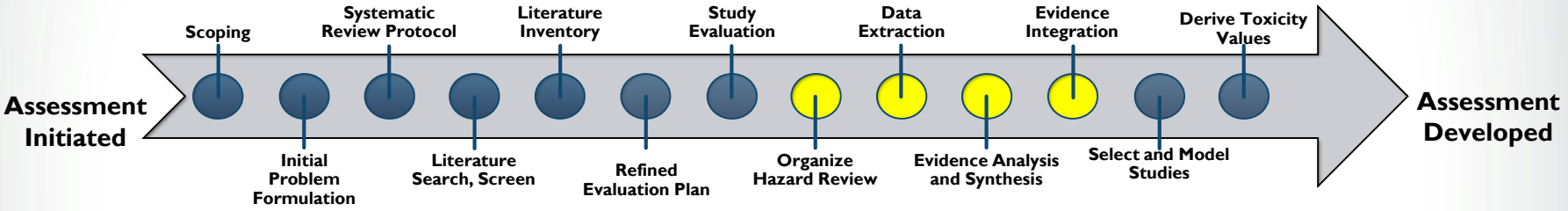
- Current methods for study evaluation
  - Chloroform protocol (2018)
  - EPA Response to Chloroprene Request for Correction (2018)

EPA
United States
Environmental Protection
Agency

Evaluating Confidence in a Body of Evidence: Evidence Synthesis and Integration to Reach Hazard Conclusions

Office of Research and Development
NCEA, IRIS

Synthesis and Integration of Evidence Linking Exposure and Health Effects: Purpose



Synthesis: To describe the types of information within each line of evidence (human, animal and mechanistic), and to analyze and present study results regarding a given health effect to facilitate integration judgments.

- Decisions about the organization of the synthesis made prior to data extraction
- Narratives, but not study by study summaries
- Highlight information that informs the hazard evaluation

Integration: To develop judgments regarding strength of evidence for a health effect across lines of evidence

- A two-step process involving transparent and structured approaches for drawing summary conclusions across lines of evidence

NAS 2014: Relevant Comments and Recommendations

The NAS 2014 report discusses the complexities with organizing analyses around mechanism, noting that, "The history of science is replete with solid causal conclusions in advance of solid mechanistic understanding." (NRC, 2014, p. 90).

- The current approach focuses first on the available human and animal studies on health effects, incorporating mechanistic information at various stages of assessment development to clarify identified gaps in understanding (e.g., human relevance of animal-model data).

"The risk-of-bias assessment of individual studies should be carried forward and incorporated into the evaluation of evidence among data streams." (NAS 2014 Recommendation, Box 8-1)

- The results of the evaluation of individual studies is a critical component of the current evidence synthesis processes and integration frameworks.

NAS 2014: Relevant High Priority (Box 8-1) Recommendations

"**EPA should continue to improve its evidence-integration process incrementally and enhance the transparency of its process**. It should either maintain its current guided-expert-judgment process but make its application more transparent or adopt a structured (or GRADE-like) process...the committee does not offer a preference but suggests that EPA consider which approach best fits..."

"EPA should expand its ability to perform quantitative modeling of evidence integration."

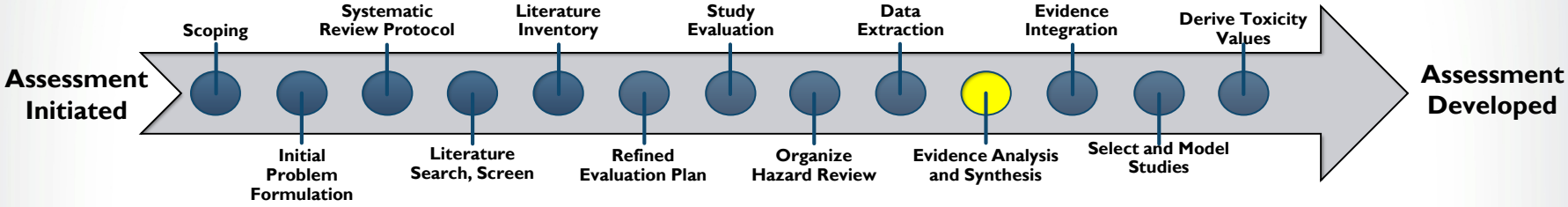
- The current approach continues to use a guided expert judgment process, but structured sets of categorical criteria for decision-making within that process are more explicitly defined.
- The current frameworks, and documentation of decisions within these frameworks, enhance transparency, reproducibility, and comparability across health effects and assessments; these approaches are evolving within NCEA and across the field.
- Current research activities include quantitative methods to integrate evidence across streams (e.g., Bayesian approaches; see Session 4)

Synthesizing Evidence on Health Effects – Organization and Structure

Some questions about the evidence

- What outcomes are relevant to each health hazard domain and at what level (e.g., health effect or subgroupings) should synthesis occur?
- What populations were studied (e.g., general population, occupations, life stages, species, etc.) and do responses vary?
- Can study results be described across varying exposure patterns, levels, duration or intensity?
- Are there differences in the confidence in study results for different outcomes, populations, or exposure?
- Does toxicokinetic information explain differences in responses across route of exposure, other aspects of exposure, species, or life stages?
- How might dose response relationships be presented (specific study results or across study results)?

Scientific Judgment in Analysis and Synthesis of Evidence



- Synthesis of evidence is more than counting the number of "positive" and "negative" studies
- Must systematically consider the influence of bias and sensitivity when describing study results and synthesizing evidence
- Synthesis should primarily be based on studies of medium and high confidence (when available)
- Analysis should try to draw conclusions about the strength of evidence from findings across collections of studies

Synthesis Considerations for Determining Strength of Evidence

Epidemiology evidence	Animal toxicology evidence
Study evaluation conclusions (risk of bias, sensitivity) are incorporated into analyses of each of the following considerations (adapted Hill considerations):	

Consistency

Effect magnitude/ precision

Biological gradient/ dose-response

Coherence

Natural experiments

Temporality

Rare, but important to highlight

- concerns about bias; smaller effect size is not discounted outright

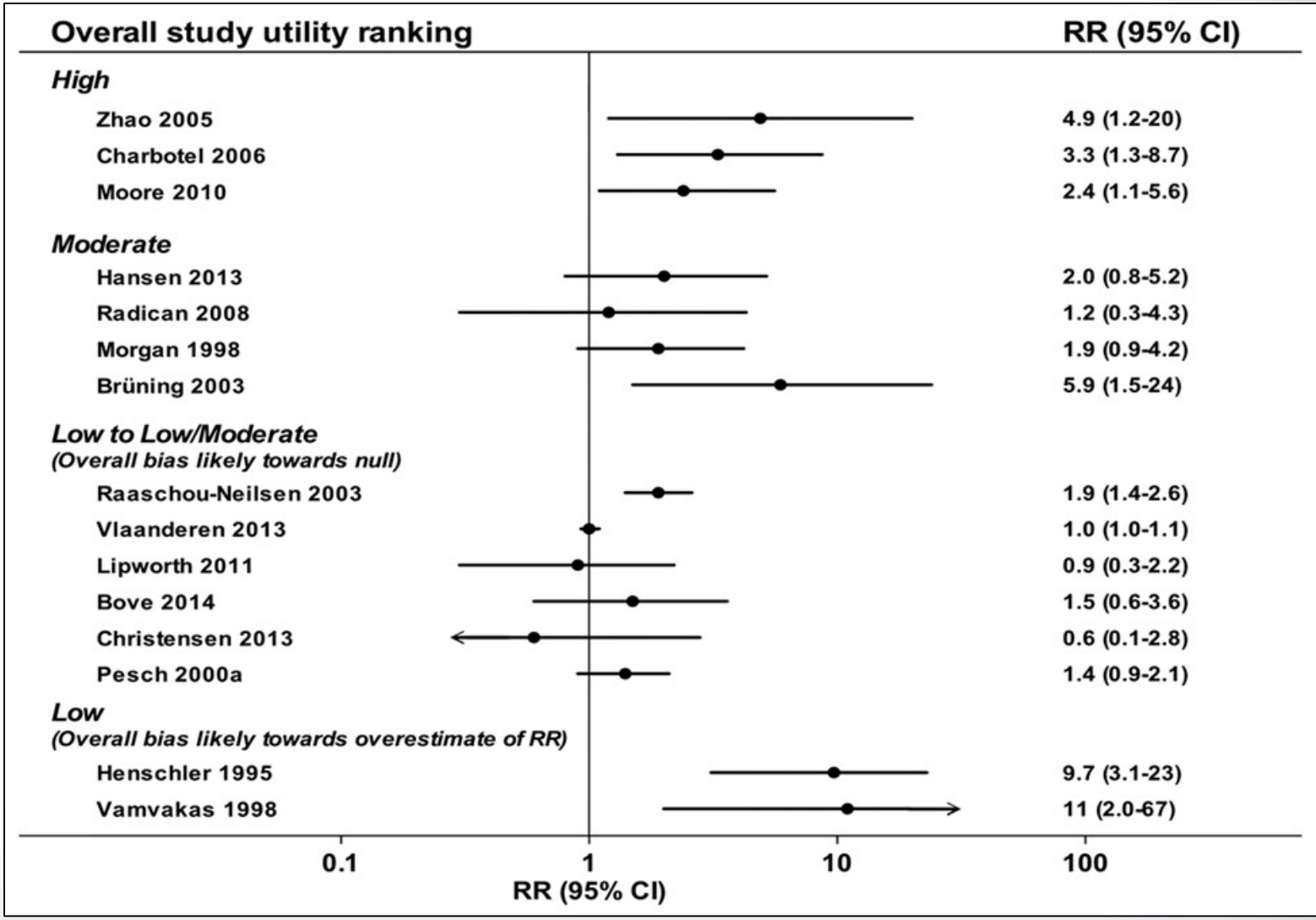
Timing of exposure relative to development of outcomes is assessed during study evaluation phase

- concerns about chance

- Informative human and animal hea
  analyzed and synthesized separate
- Mechanistic evidence is synthesize
  the human and animal health effect evidence

Synthesis Examples: Epidemiology

TCE and kidney cancer: stratification by utility

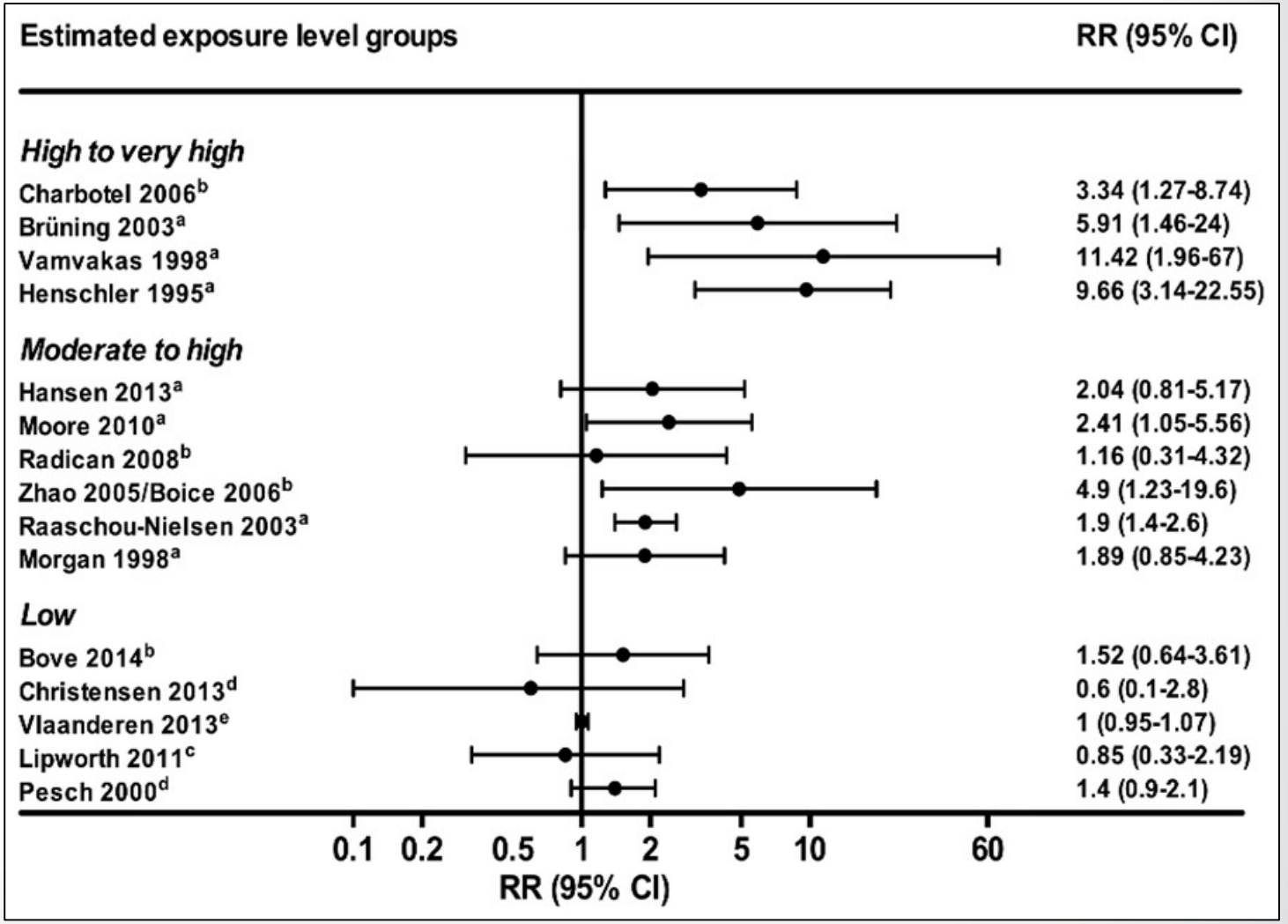


Highest exposure level graphed for each study

RoC Monograph on Trichloroethylene. January 2015. https://ntp.niehs.nih.gov/go/797306

Synthesis Examples: Epidemiology

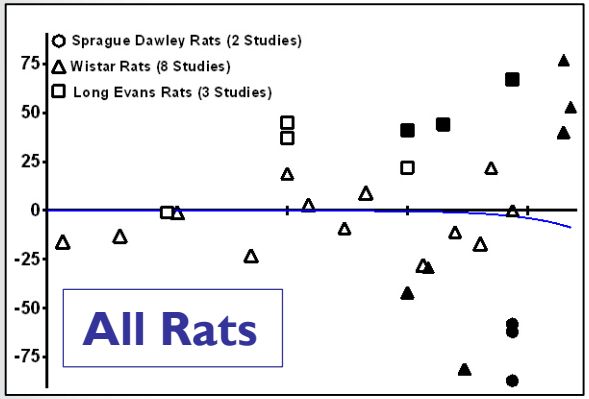
TCE and Kidney Cancer: stratification by exposure level

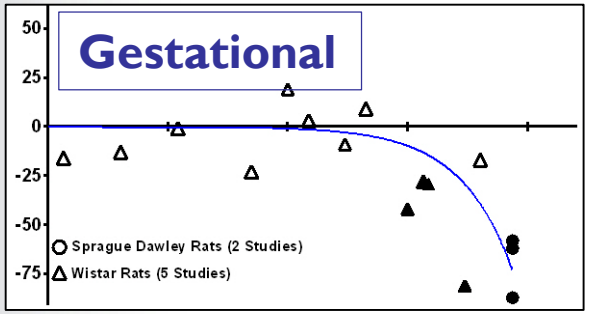


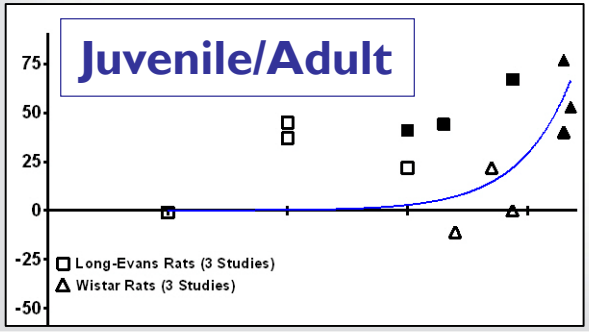
EPA. 2011. Toxicological Review of Trichloroethylene

Synthesis Examples: Animal Toxicology

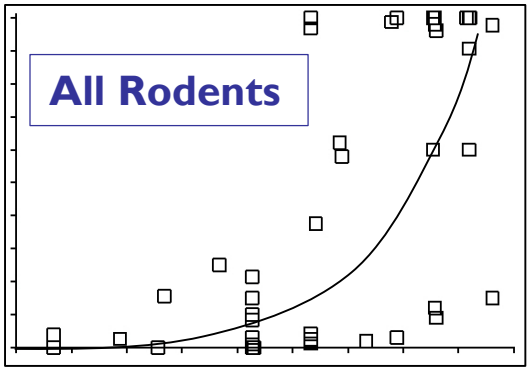
Hormone Level

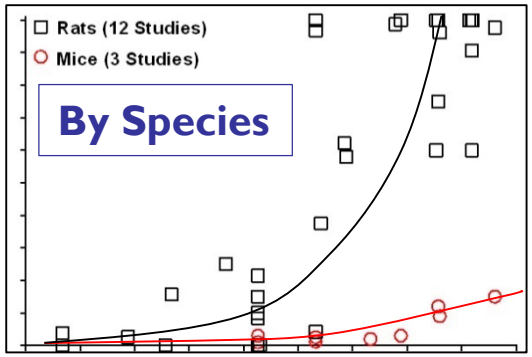




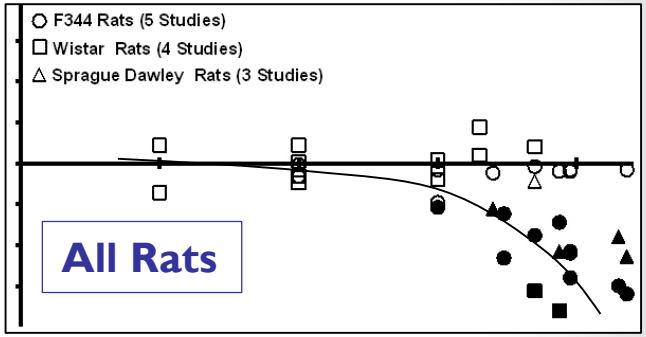


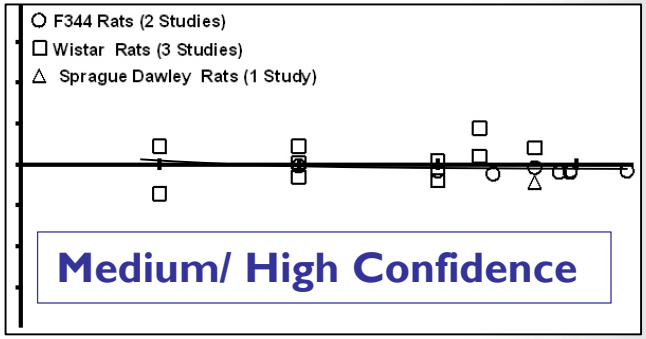
Pathology Incidence

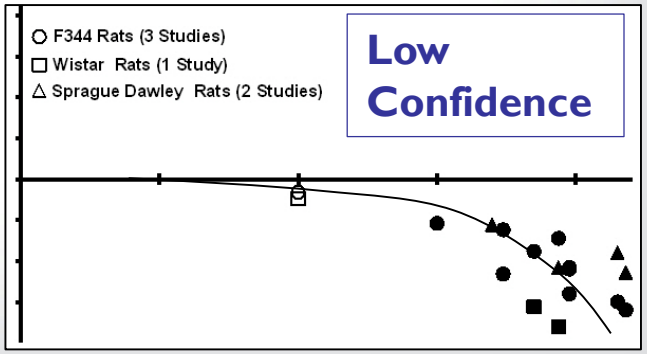




Behavioral Function







Mechanistic Evidence

"Mechanistic data represent a wide variety of studies not intended to identify an adverse outcome." (NRC, 2014)

- When evaluating mechanistic evidence, the scope is larger than "*in vitro*" data
- **Mechanistic inventories collected at earlier stages may include:**
  - *In vivo* (cellular, biochemical, molecular)
  - *In vitro* or *ex vivo* (human or animal tissues or cells)
  - Non-animal or non-mammalian alternative animal models
  - Big data ('omics or high-throughput assays)
  - "Intervention" studies (pharmacologic, environmental, genetic)

"…there might be hundreds of *in vitro* and other mechanistic studies of a given chemical…" (NRC, 2014)

"For a given chemical, multiple mechanisms might be involved in a given end point, and it might not be evident how different mechanisms interact in different species to cause the adverse outcome." (NRC, 2014)

Systematic review of mechanistic information requires a different approach

"When human data are nonexistent, are mixed, or consistently show no association and an animal study finds a positive association, the importance of mechanistic data is increased..." (NRC, 2014)

To narrow the scope of the analyses of mechanistic information, IRIS applies an iterative approach to identifying key mechanistic questions at various stages of the systematic review

- *Problem formulation* identifies predefined analyses (e.g., when a mutagenic MOA is indicated)
- *Literature inventory* allows identification of studies on an organ system that human and animal studies meeting the PECO criteria have not examined
- *Human and animal evidence syntheses* may flag impactful qualitative and quantitative analyses

Human and animal evidence syntheses may flag impactful mechanistic analyses

- Identify precursor events for apical toxicity endpoints
- Inform susceptibility (species, strain, or sex differences; at-risk populations or lifestages)
- Inform human relevance of animal data (note: the level of analysis will vary depending on the impact of the animal evidence)
- Provide biological plausibility (i.e., to human or animal health effect data when evidence is weak or critical uncertainties are identified)
- Establish mechanistic relationships (or lack thereof) across sets of potentially related endpoints/outcomes to inform the consideration of coherence during evidence integration
- Aid extrapolation (high-to-low dose; short-to-long duration; route-to-route)
- Improve dose-response modeling and quantification of uncertainties

Mechanistic Analysis Focused on Specific Questions

Examples of when these analyses have been triggered in recent IRIS Assessments:

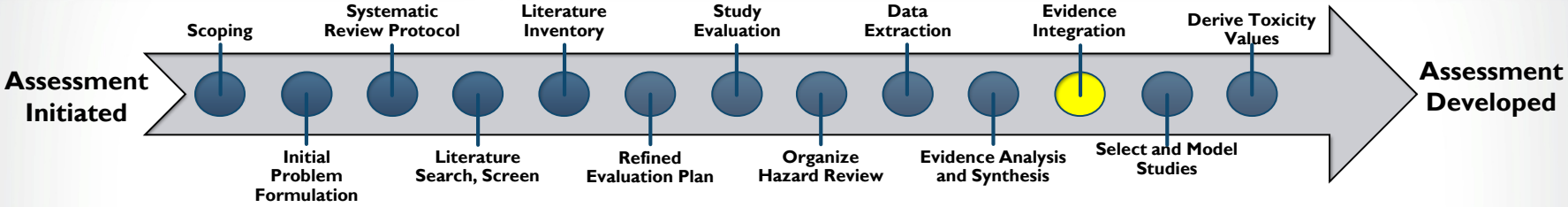
- Benzo[a]pyrene (2017): The descriptor "carcinogenic to humans" was supported by strong mechanistic evidence that established the biological plausibility of the animal findings occurring in humans, despite lack of human exposure data
  - Key precursors (BPDE-DNA adducts) were identified in humans exposed to PAH mixtures that are specific to B[a]P, form mutational spectra unique to B[a]P, and are associated with cancer in humans
- Dichloromethane (2011): The cancer risk estimate was specifically derived for a susceptible subpopulation (GSTT1+/+) identified by the mechanistic evaluation
  - Differing results *in vivo* were explainable by species and tissue differences in the availability of GST
  - PBPK modeling addressed the variability in this population
- **Documentation and transparency is key for future mechanistic analyses**

Focused mechanistic evaluations

"Several criteria should be considered in assessing in vitro toxicology studies for risk of bias and toxicologic relevance. Relevance should be determined in several domains, including cell systems used, exposure concentrations, metabolic capacity, and the relationship between a measured in vitro response and a clinically relevant outcome measure. **Few tools are available for assessing risk of bias in in vitro studies. Because of the nascent status of this field, the committee can provide only provisional recommendations for EPA to consider...**EPA should carry out, support, or encourage research on the development and evaluation of empirically based instruments for assessing bias in…mechanistic studies." (NRC, 2014)

- **Prioritize studies of relevant endpoints and associated assays by toxicologic relevance (e.g., model systems; dose range; sensitivity and specificity of assay)**
- **Conduct individual study evaluations on the most impactful studies**
- **EPA is exploring the use of existing tools, including adaptations of IRIS study evaluation tools**
- **Organizational frameworks (e.g., EPA's MOA framework using modified Hill considerations; visual AOP-like constructs) are useful for organizing and documenting these analyses transparently to convey conclusions for evidence integration**

Moving from Synthesis to Integration



Outputs of Evidence Synthesis

Results of Human Health Effect Study Synthesis

Results of Animal Health Effect Study Synthesis

Results of Synthesis of Mechanistic Evidence Informing the Human and Animal Syntheses

Evidence Integration

Transparent and Structured Processes for Drawing Summary Conclusions Across Lines of Evidence

Evidence Integration Involves a Sequential, Two-Step Process

- **Evidence synthesis interpretations for each consideration relating to causality are combined across lines of evidence using transparent, structured frameworks**

Step 1: "Within-Stream" Integration	Step 2: "Across-Stream" Integration
Judge the Strength of the Evidence from the: • **Human Evidence Stream** • **Animal Evidence Stream**	Draw Overall Evidence Integration Conclusions based on: • Combined **Human** and **Animal Evidence Streams**
Human health effect study synthesis conclusions for each consideration are integrated in light of mechanistic evidence in exposed humans or human cells (or other human models)	The judgments regarding the strength of the human and animal evidence streams are integrated in light of evidence on the human relevance of the findings in animals, susceptibility, and the coherence of the findings across evidence streams.
Characterize the Strength of the Evidence for an Effect in Animals (**Animal Evidence Stream Judgment**)	
Animal health effect study synthesis conclusions for each consideration are integrated in light of mechanistic evidence in exposed animals or animal cells (or other relevant models)	

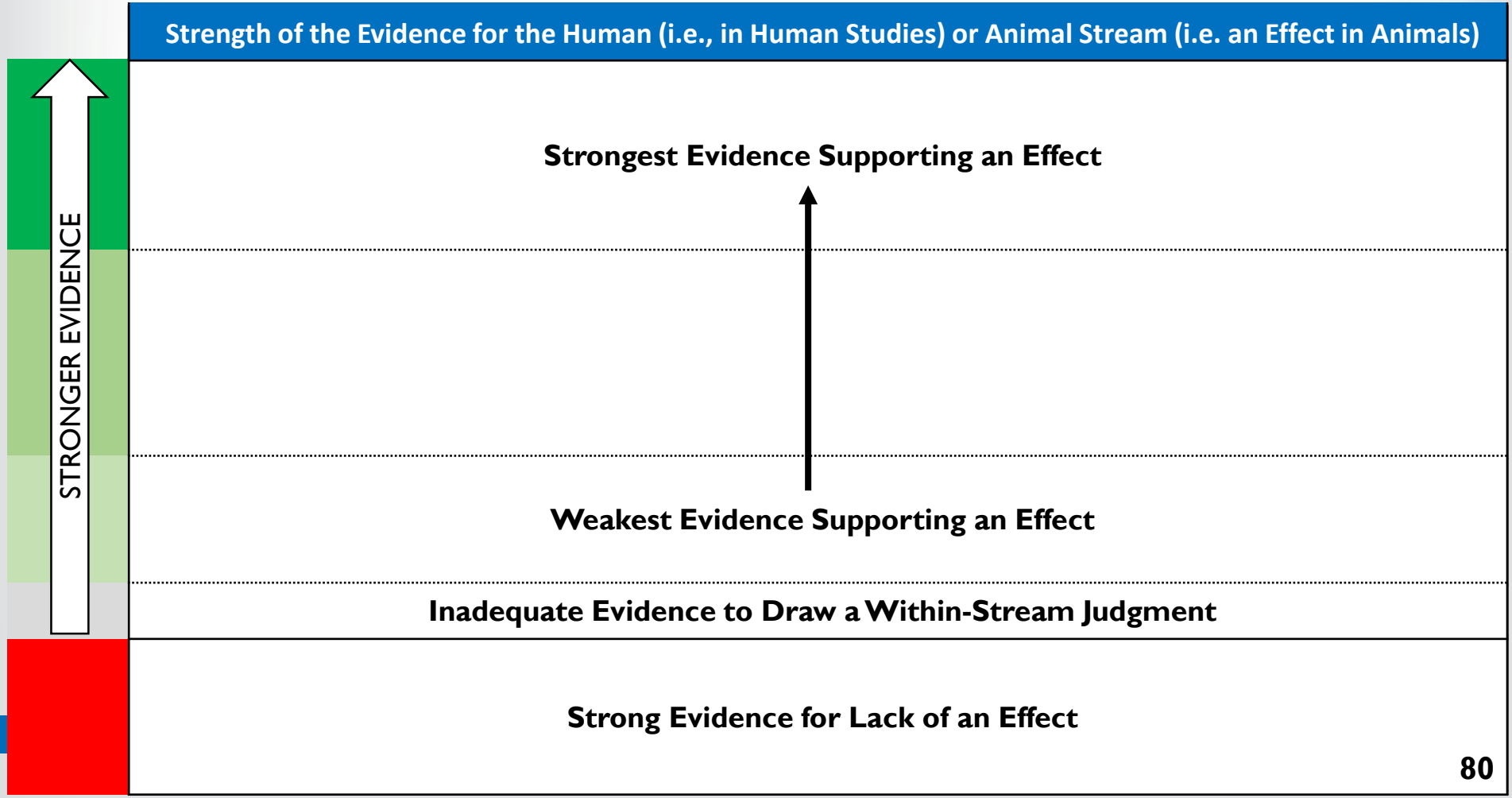
Within-Stream (Human; Animal Stream) Evidence Judgment Considerations

	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
Dose-response	• Simple or complex (nonlinear) relationships 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 reduces evidence strength	
	• Informed by mechanistic evidence on the biological development of the health effect or toxicokinetic/dynamic knowledge of the chemical or related chemicals	
Mechanistic Evidence on Biological Plausibility	• 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	

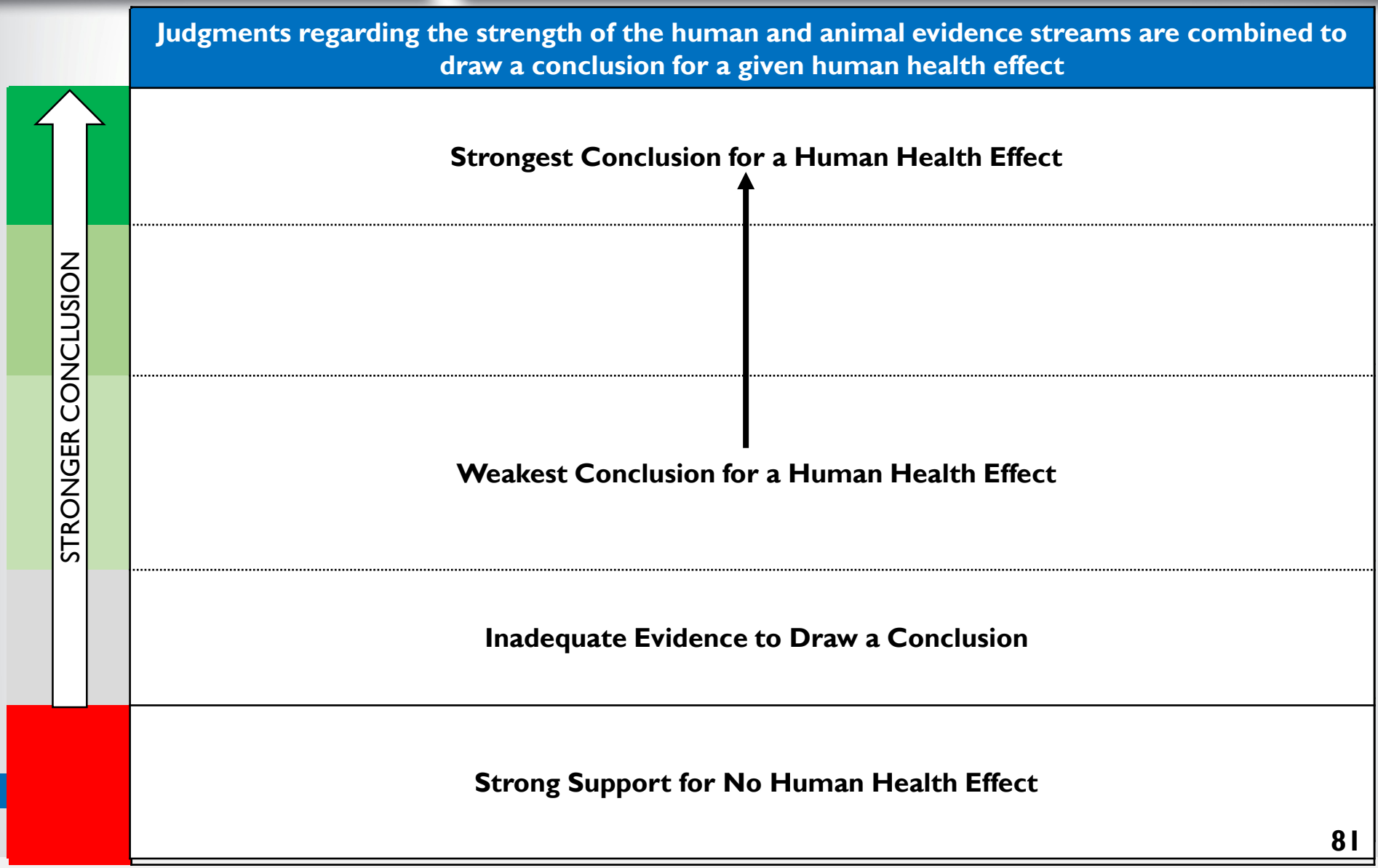
Light blue rows highlight mechanistic inferences; "temporality" and "natural experiments" not shown

Step 1: Framework for Within-Stream Evidence Judgments

The Hill-based considerations are applied to judge the strength of the evidence from human studies and, separately, the evidence for an effect in animals



Step 2: Framework for Overall Evidence Integration Conclusions



Evidence Profile Table: Supports the Evidence Integration Narrative

"the weight of evidence descriptions need to indicate the various determinants of weight... to be able to understand what elements (such as consistency) were emphasized" [NRC, 2011]; "No matter what method is used to integrate the different kinds of evidence available for an IRIS assessment, using a template for the evidence-integration narrative could help to make IRIS assessments more transparent." [NRC, 2014]

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)					*Human relevance of findings in animals* • *Cross-stream coherence (i.e. for both health effect-specific and mechanistic data)* • *Other inferences:* o *Information on susceptibility* o *MOA analysis inferences: precursors, cross-species inferences of toxicokinetics, or quantitative implications* o *Relevant information from other sources (e.g., read across; other, potentially related health hazards)*	*Describe conclusion(s) and primary basis for the integration of all available evidence (across human, animal, and mechanistic):* + + + Strongest conclusion + + ◯ ↑ + ◯◯ Weakest conclusion ◯◯◯ – ◯◯ – – ◯ Inadequate – – – Strong support for no human health effect Summarize the models and range of dose levels upon which the conclusions were primarily reliant
• *References* • *Study confidence (based on evaluation of risk of bias and sensitivity) and explanation* • *Study design description*	• *Consistency* • *Dose-response gradient* • *Coherence of observed effects (apical studies)* • *Effect size (magnitude, severity)* • *Biological plausibility* • *Low risk of bias/ high quality* • *Insensitivity of null/ negative studies* • *Natural experiments* • *Temporality*	• *Unexplained inconsistency* • *Imprecision* • *Indirectness/ applicability* • *Poor study quality/ high risk of bias* • *Other (e.g., Single/Few Studies; small sample size)* • *Evidence demonstrating implausibility*	• *Results information (general endpoints affected/ unaffected) across studies* • *Human mechanistic evidence informing biological plausibility: discuss how data influenced the within stream judgment (e.g., evidence of precursors in exposed humans).* *Could be multiple rows (e.g., grouped by study confidence or population) if this informs results heterogeneity*	*Describe strength of the evidence from human studies, and primary basis:* + + + Strongest evidence + + ◯ ↑ + ◯◯ Weakest evidence ◯◯◯ – ◯◯ – – ◯ Inadequate – – – Strong evidence for no effect		
Evidence for an Effect in Animals (Route)						
• *References* • *Study confidence (based on evaluation of risk of bias and sensitivity) and explanation* • *Study design description*	• *Consistency and Replication* • *Dose-response gradient* • *Coherence of observed effects (apical studies)* • *Effect size (magnitude, severity)* • *Biological plausibility* • *Low risk of bias/ high quality* • *Insensitivity of null/ negative studies*	• *Unexplained inconsistency* • *Imprecision* • *Indirectness/ applicability* • *Poor study quality/ high risk of bias* • *Other (e.g., Single/Few Studies; small sample size)* • *Evidence demonstrating implausibility*	• *Results information (general endpoints affected/ unaffected) across studies* • *Animal mechanistic evidence informing biological plausibility for effects in animals: discuss how mechanistic data influenced the within stream judgment (e.g., evidence of coherent molecular changes in animal studies)* *Could be multiple rows (e.g., by study confidence, species, or exposure duration) if this informs results heterogeneity*	*Describe strength of the evidence for an effect in animals, and primary basis:* + + + Strongest evidence + + ◯ ↑ + ◯◯ Weakest evidence ◯◯◯ – ◯◯ – – ◯ Inadequate – – – Strong evidence for no effect		

Evidence Integration Conclusions

- **For Cancer**, conclusions on the integrated evidence for each cancer type (or grouping) are evaluated in the context of MOA information to develop an evidence integration narrative that includes a descriptor for carcinogenicity:
  - **carcinogenic** to humans; **likely** to be carcinogenic to humans; **suggestive** evidence of carcinogenic potential; **inadequate** information to assess carcinogenic potential; or **not likely** to be carcinogenic to humans
- **For Noncancer Effects**, frameworks for evaluating the integrated evidence have been developed to add structure and transparency to the evidence integration narrative(s), which include(s) the relevant exposure context.
  - IRIS has not yet incorporated standardized descriptors for noncancer effects
  - The NAS recommended incremental improvements in this area, including recommendations to "Develop uniform language to describe strength of evidence on noncancer effects" [p. 92, 2014]
  - The specific way in which these conclusions are summarized is currently being tested and discussed within EPA

IRIS has Addressed the Major NAS 2014 Recommendations

NAS 2014 Topics	IRIS Process Improvements
Evidence Evaluation (Chapter 5)	• Individual studies are evaluated for reporting quality, risk of bias, and sensitivity • Decisions and supporting rationale are clearly documented • Study evaluations impact subsequent assessment decisions
Evidence Integration for Hazard Identification (Chapter 6)	• Structured frameworks provide transparency in expert judgments across human, animal, and mechanistic studies (based on Hill) • Standardized templates documenting key evidence integration decisions have been developed (evidence profile tables)

See Posters and Demonstrations:

- Male reproductive toxicity in studies of phthalates (4 posters on a case study for each of the 3 lines of evidence and the overall evidence integration)
- Combining data within species (poster on meta-analytical approaches)
- PBPK model evaluation for human health assessments (poster)
- Health Assessment Workspace Collaborative (demonstration)

EPA
United States
Environmental Protection
Agency

SESSION 3: DEVELOPMENT AND APPLICATION OF SPECIALIZED TOOLS FOR SYSTEMATIC REVIEW

Kris Thayer*, Michele Taylor*, Amina Wilkins, Xabier Arzuaga

[*Speaking]

Office of Research and Development
NCEA, IRIS

NAS 2014: Chapter 8 "Looking Forward"

"[EPA] need to consider developing a strategic plan for continuous updating of the IRIS methodology... For example, such a strategic plan should address:

- Applying advances in data retrieval and text-mining

"The committee also found that the proposed format for the assessments should enhance "user friendliness" and transparency. The evidence tables and data displays in the new documents are moving to the standard practice for systematic reviews." [p. 136]

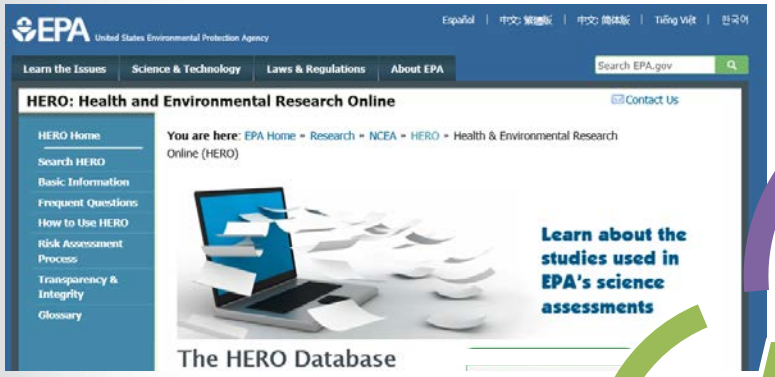
Current Application of Systematic Review Software

- Specialized software tools make the process more efficient
  - Time and cost savings, improved data management, increased transparency
- NOT all systematic review software tools are intended to automate/semi-automate the process, e.g., HAWC helps manage information content
  - Currently, automation tools are most advanced for evidence identification
- Prefer free tools when possible to help address needs of a potentially large community of users in environmental and biomedical sciences
- Incorporate tools after confirming acceptable performance and interoperability with HERO
  - A toolbox approach, not a "one and only" tool model
- Organized multiple IRIS staff training sessions in 2017 and created a support team ("train the trainers" model)

Research Activities

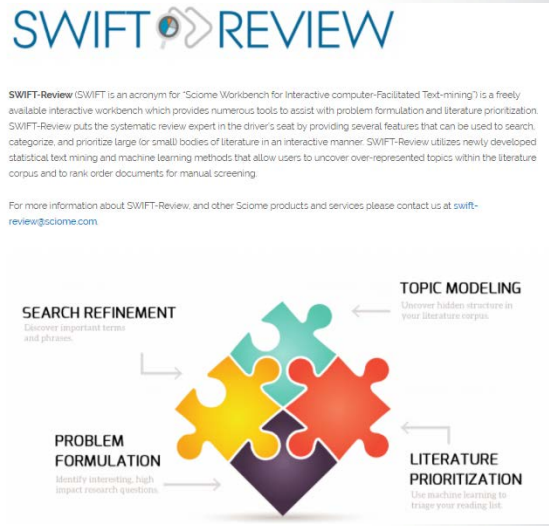
- Developing tools to help automate beyond evidence identification is a long-term research commitment
  - Major hurdle is lack of training/test sets for model development
  - Better performance expected for more structured content (e.g., animal bioassay compared to epidemiological studies)
- Any progress on semi-automation could result in large time and cost savings
- In 2017, NCEA created an interagency agreement with NTP to leverage resources
  - Current activities focus on creating test/training sets and model development for basic content of animal studies (e.g., test chemical, species, dose levels, randomization, etc.).
  - Other parts of EPA can also utilize interagency agreement
- Innovation challenges may be required to identify solutions for capturing complex content, i.e., table content, information spread across multiple sentences and paragraphs

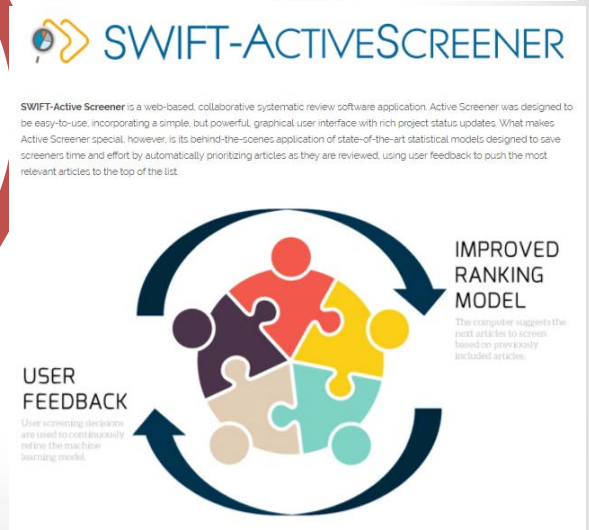
Suite of Systematic Review Software Tools – Upcoming Demonstrations



ARE

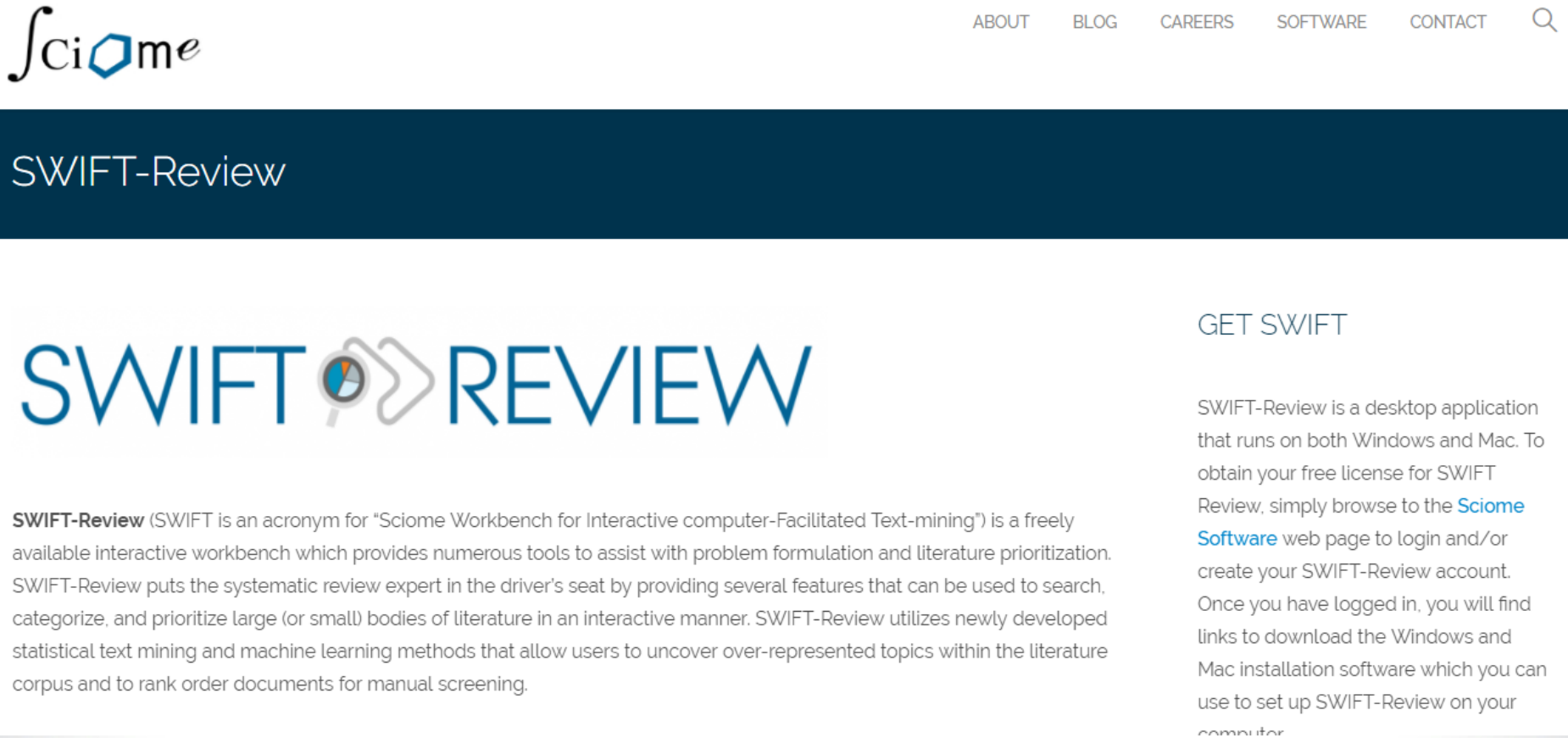
INTEROPERABLE





HAWC
HEALTH ASSESSMENT WORKSPACE COLLABORATIVE

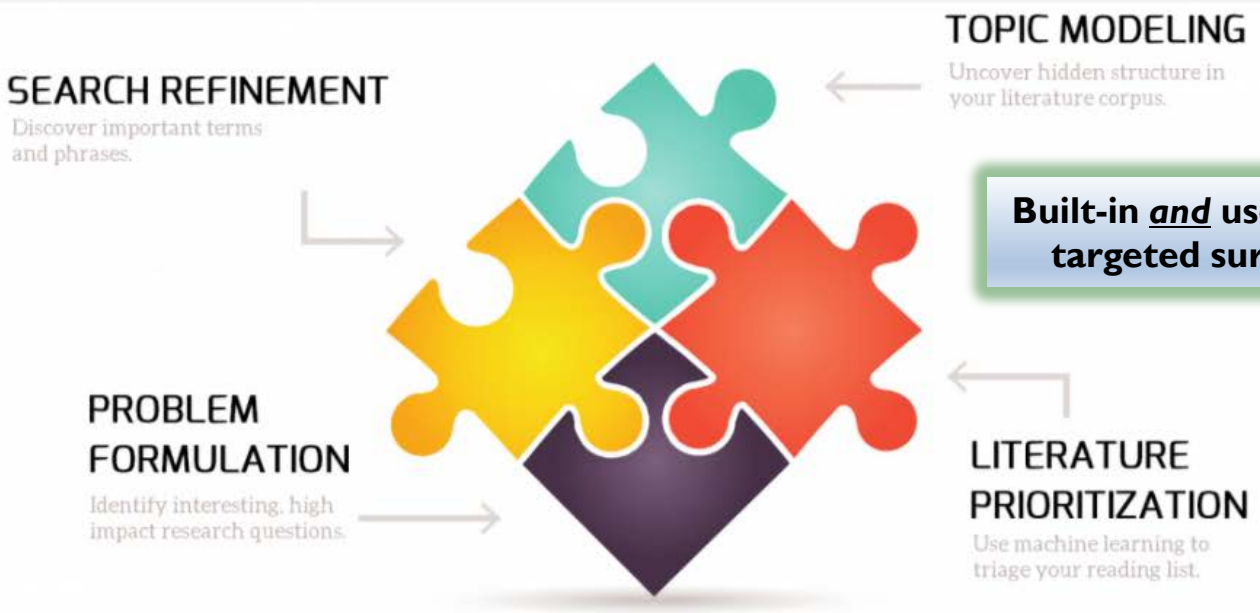
SWIFT Review: Scoping and Problem Formulation



https://www.sciome.com/swift-review/

Increased Efficiency During Scoping and Problem Formulation

Can be used to screen studies according to the PECO statement



Built-in *and* user-defined search queries allow targeted surveys of the literature corpus

Machine learning prioritizes relevant literature, reducing the screening burden by at least 50%

"Tags" Facilitate Searching During Problem Formulation

SEARCH REFINEMENT

Discover important terms and phrases.

PROBLEM FORMULATION

Identify interesting, high impact research questions.

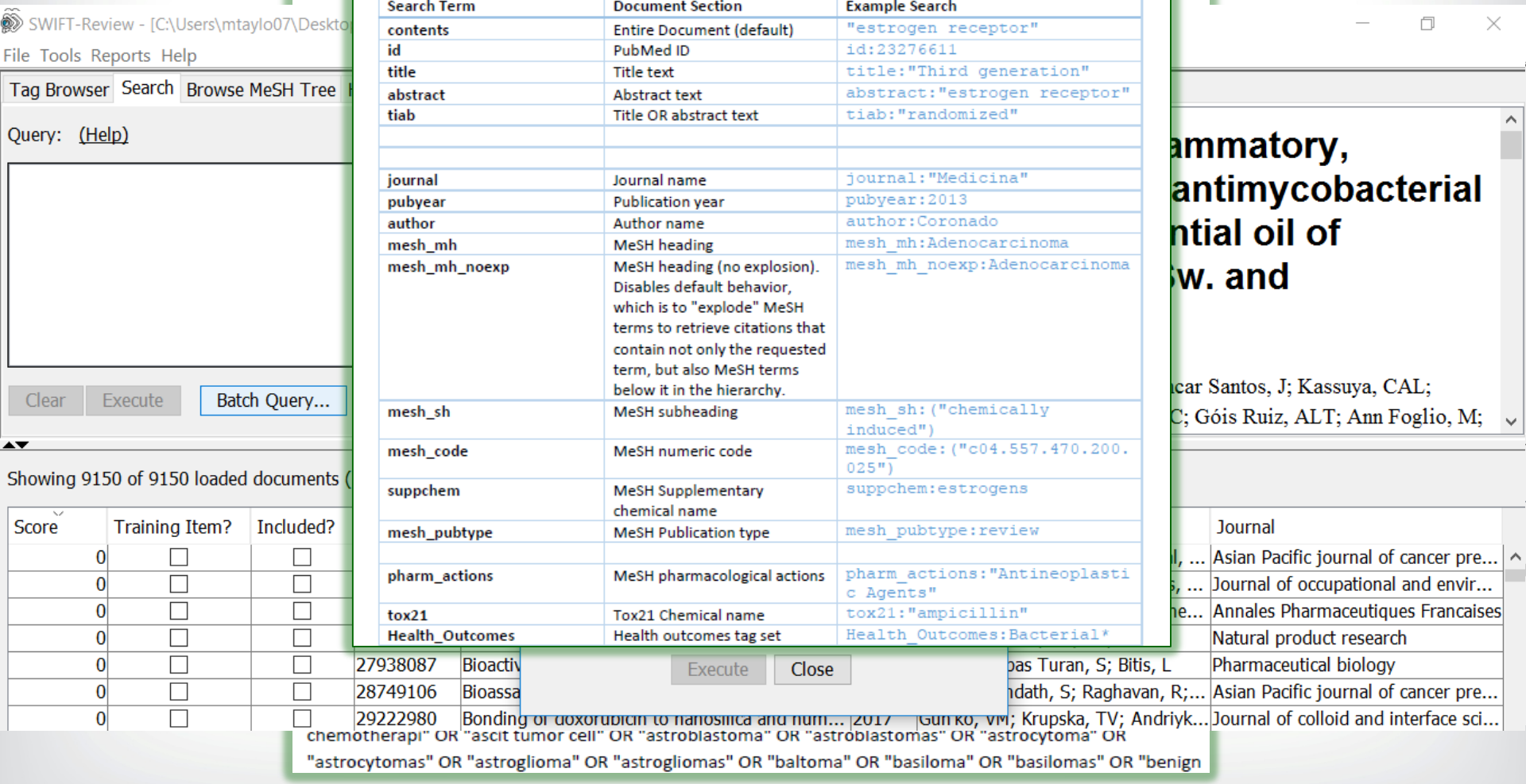
MeSH Terms and Bibliographic Data: Documents originating from PubMed bring along their associated Medline tags, including MeSH Ter

SWIFT-Review **tags** are labels assigned to bibliographic documents that are organized into **tag categories**. For example, the tag category "Health Outcomes" includes the following tags: "Cancer," "Cardiovascular," and "Neurological." When used with the **Tag Browser** or **Search** functionalities, tags facilitate increased efficiency during scoping and problem formulation by quickly finding the documents you're interested in.

<u>Tags can be assigned both manually and automatically using a variety of mechanisms:</u>

of relevance to environmental health researchers, such as the nearly 10,000 Tox21 chemicals

Built-in and User-Defined Search Strategies



The following table lists the (case sensitive) terms available to use with fielded search:

Search Term	Document Section	Example Search
contents	Entire Document (default)	"estrogen receptor"
id	PubMed ID	id:23276611
title	Title text	title:"Third generation"
abstract	Abstract text	abstract:"estrogen receptor"
tiab	Title OR abstract text	tiab:"randomized"
journal	Journal name	journal:"Medicina"
pubyear	Publication year	pubyear:2013
author	Author name	author:Coronado
mesh_mh	MeSH heading	mesh_mh:Adenocarcinoma
mesh_mh_noexp	MeSH heading (no explosion). Disables default behavior, which is to "explode" MeSH terms to retrieve citations that contain not only the requested term, but also MeSH terms below it in the hierarchy.	mesh_mh_noexp:Adenocarcinoma
mesh_sh	MeSH subheading	mesh_sh:("chemically induced")
mesh_code	MeSH numeric code	mesh_code:("c04.557.470.200.025")
suppchem	MeSH Supplementary chemical name	suppchem:estrogens
mesh_pubtype	MeSH Publication type	mesh_pubtype:review
pharm_actions	MeSH pharmacological actions	pharm_actions:"Antineoplastic Agents"
tox21	Tox21 Chemical name	tox21:"ampicillin"
Health_Outcomes	Health outcomes tag set	Health_Outcomes:Bacterial*

Tag Browser Search by Health Outcome

SWIFT-Review - [C:\Users\mtaylo07\Desktop\SWIFT Review\NAS Session 3.stp]

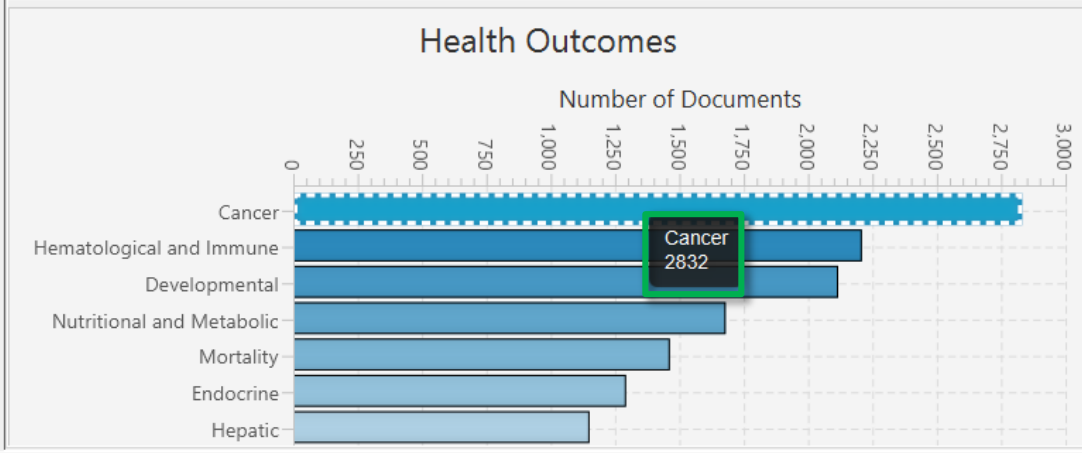
File Tools Reports Help

Tag Browser | Search | Browse MeSH Tree | Heatmap Browser | Prioritized Lists

Health Outcomes

Tag	Code(s)	Count
Cancer		2832
Hematological and Immune		2209
[No Tag]		2130
Developmental		2117
Nutritional and Metabolic		1680
Mortality		1463
Endocrine		1293
Hepatic		1151
Respiratory		1051
Gastrointestinal		1032

Document Preview | Pie Chart | Bar Chart



Showing 2832 of 9150 loaded documents (1 selected; 27 total included; 50 total training docs.)

Score	Training Item?	Included?	RefID	Title	Year	Authors	Journal
0.95	☑	☑	h1022338	Identification of genotoxic compounds in drink...	1998	Le Curieux, F.; Erb, F.; Marzin, D.	
0.95	☐	☐	h1021972	Mutagenicity study of carbon tetrachloride and...	1998	Sasaki, T.; Suzuki, M.; Noda, K.; ...	Journal of Toxicological Sciences
0.95	☐	☐	h630464	Advances in research on carcinogenic and gen...	1993	Daniel, F. B.; Meier, , J. R.; Deang...	Annali dell'Istituto superiore di sa...
0.95	☐	☐	h1024901	Cytosine attack by free radicals arising from br...	1993	Castro, G. D.; Castro, J. A.	Teratogenesis, Carcinogenesis, an...
0.95	☑	☑	h1010308	International Commission for Protection Again...	1992	Lohman, P. H. M.; Mendelsohn, M...	Mutation Research: Fundamental a...
0.95	☑	☑	h1024875	International Commission for Protection Again...	1992	Mendelsohn, M. L.; Moore, D. H.; ...	Mutation Research
0.95	☐	☐	h1024972	Rational approach to the quantification of gen...	1992	Benigni, R.	Environmental and Molecular Muta...
0.95	☐	☐	h194881	An association between mutagenicity of the Ar...	1991	Roldán-Arjona, T.; García-Pedraja...	Mutagenesis
0.95	☐	☐	h1023024	Volatile Organohalogen Compounds from the ...	1991	Rosenberg, C.; Nylund, L.; Aalto, ...	Chemosphere

Tag Browser Search by Health Outcome

SWIFT-Review - [C:\Users\mtaylo07\Desktop\SWIFT Review\NAS Session 3.stp]

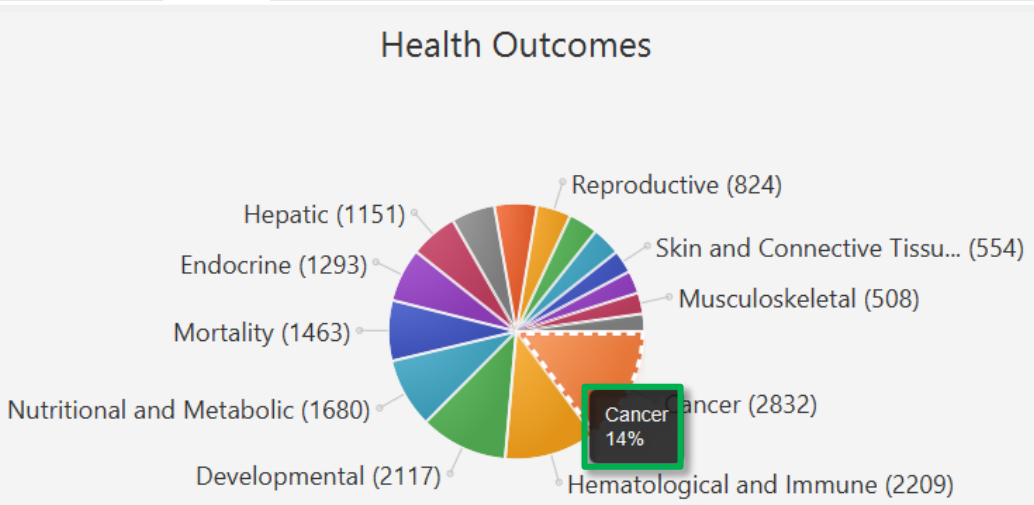
File Tools Reports Help

Tag Browser | Search | Browse MeSH Tree | Heatmap Browser | Prioritized Lists

Health Outcomes

Tag	Code(s)	Count
Cancer		2832
Hematological and Immune		2209
[No Tag]		2130
Developmental		2117
Nutritional and Metabolic		1680
Mortality		1463
Endocrine		1293
Hepatic		1151
Respiratory		1051
Gastrointestinal		1032
Reproductive		824
Renal		713
Neurological		698
Skin and Connective Tissue		554

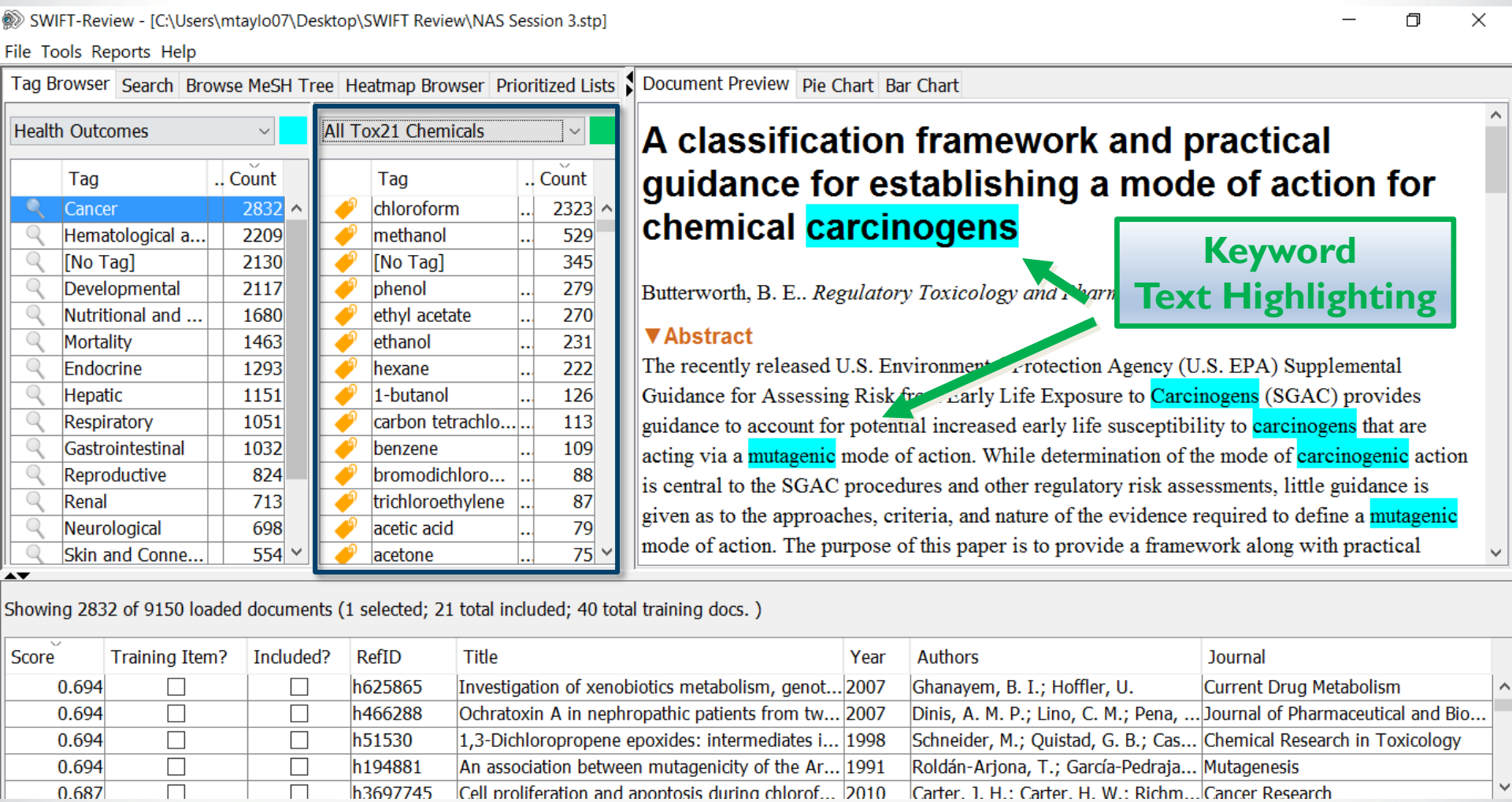
Document Preview | Pie Chart | Bar Chart



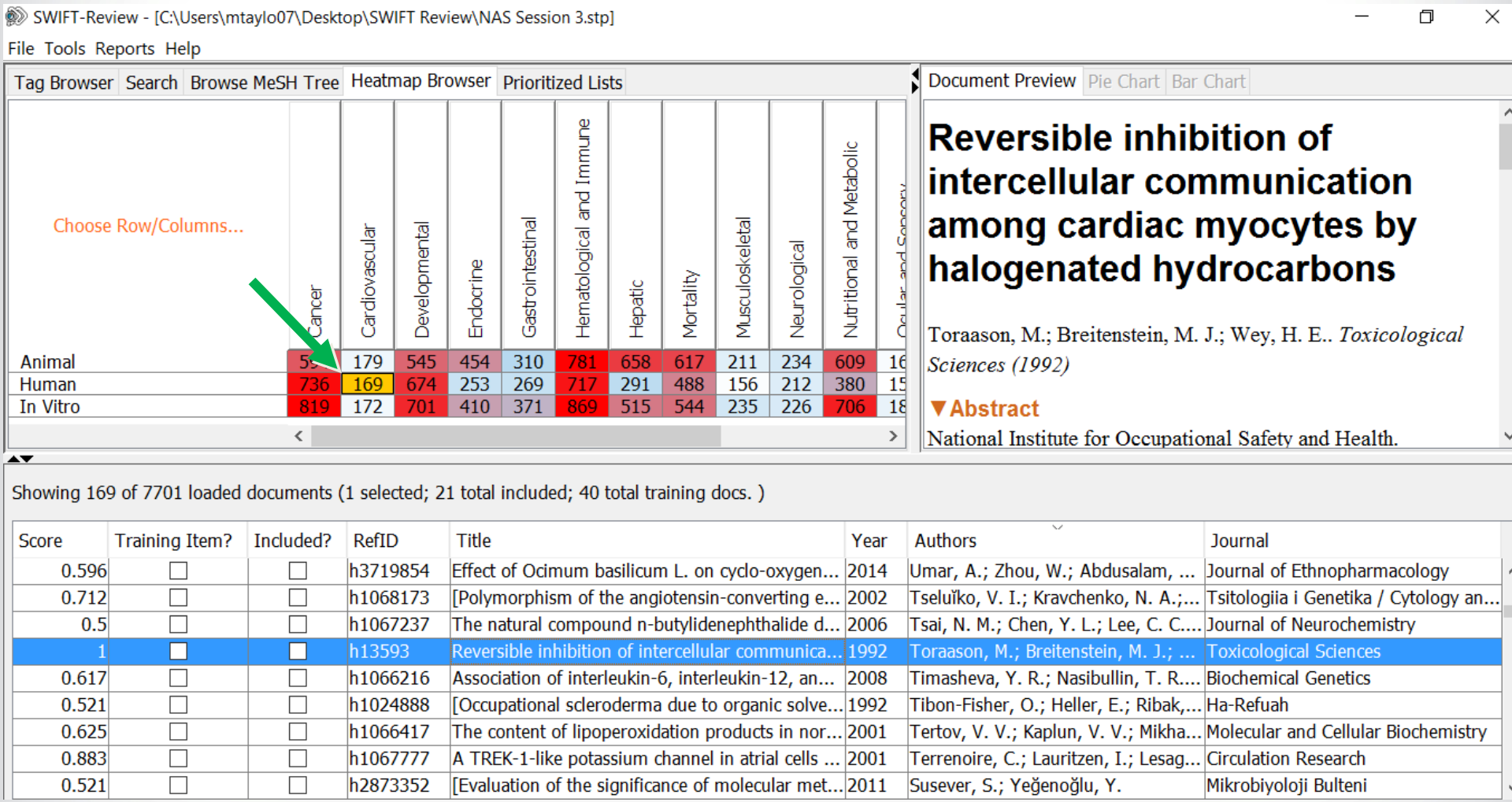
Showing 2832 of 9150 loaded documents (1 selected; 21 total included; 40 total training docs.)

Score	Training Item?	Included?	RefID	Title	Year	Authors	Journal
0.95	☐	☐	h1021972	Mutagenicity study of carbon tetrachloride and...	1998	Sasaki, T.; Suzuki, M.; Noda, K.; ...	Journal of Toxicological Sciences
0.95	☐	☐	h630464	Advances in research on carcinogenic and gen...	1993	Daniel, F. B.; Meier, , J. R.; Deang...	Annali dell'Istituto superiore di sa...
0.95	☐	☐	h1024901	Cytosine attack by free radicals arising from br...	1993	Castro, G. D.; Castro, J. A.	Teratogenesis, Carcinogenesis, an...
0.95	☑	☑	h1010308	International Commission for Protection Again...	1992	Lohman, P. H. M.; Mendelsohn, M...	Mutation Research: Fundamental a...
0.95	☑	☑	h1024875	International Commission for Protection Again...	1992	Mendelsohn, M. L.; Moore, D. H.; ...	Mutation Research

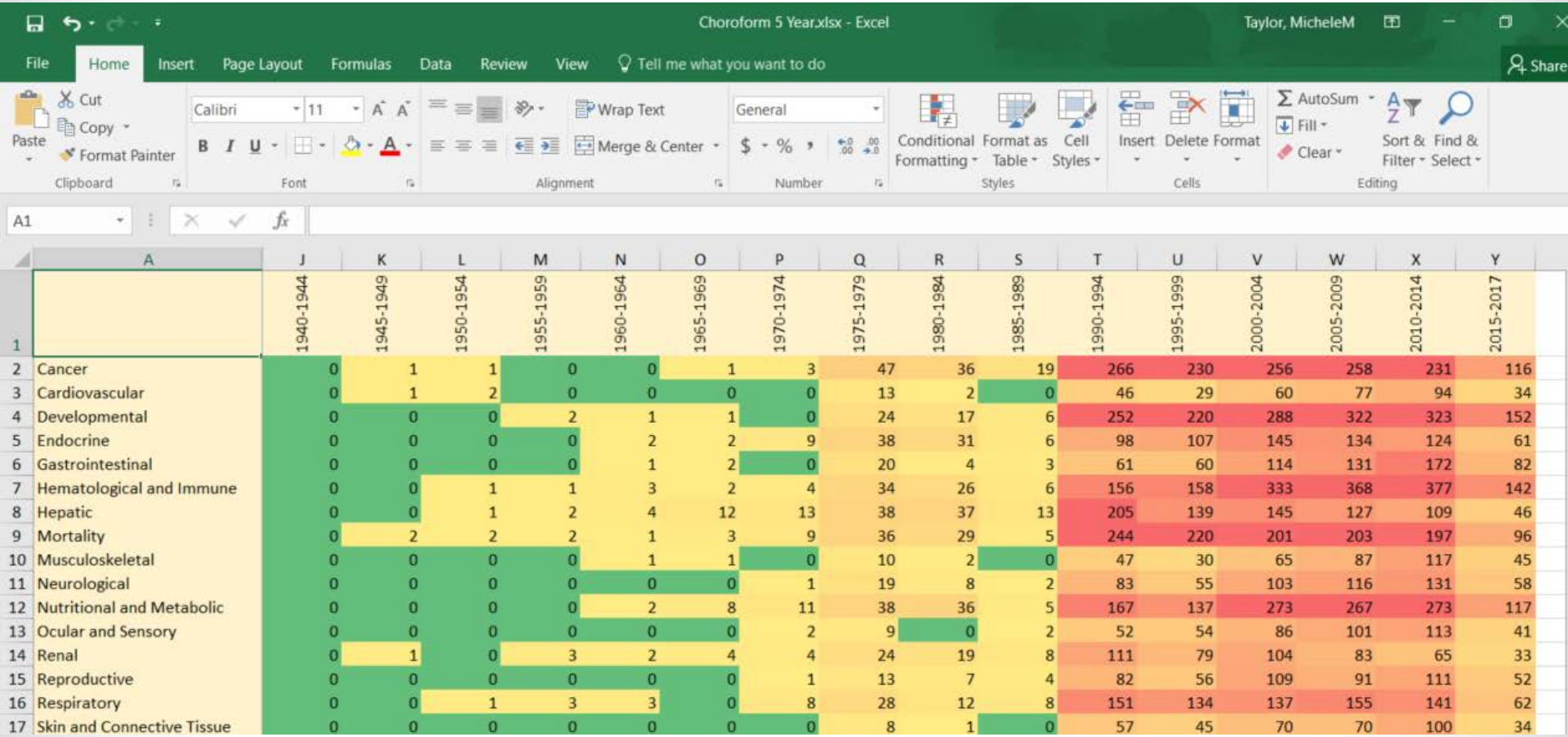
Searching Additional Tag Categories Tox21 Chemicals



Interactive Displays Reveal Patterns of Available Evidence



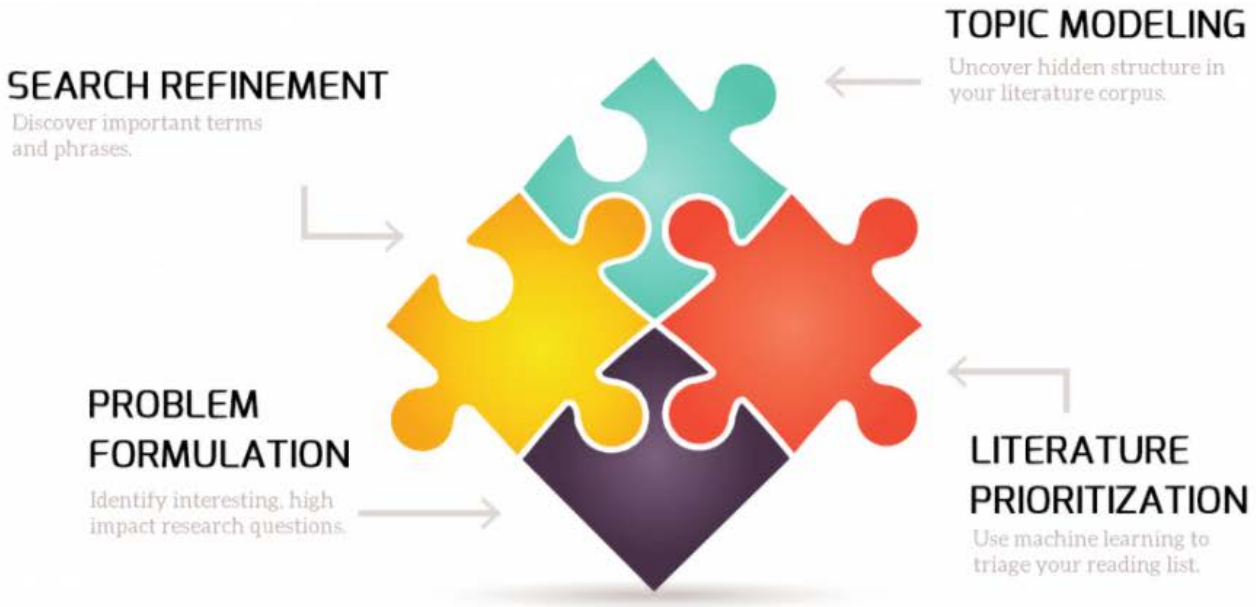
Publication Year by Health Outcome



	A	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1		1940-1944	1945-1949	1950-1954	1955-1959	1960-1964	1965-1969	1970-1974	1975-1979	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2009	2010-2014	2015-2017
2	Cancer	0	1	1	0	0	1	3	47	36	19	266	230	256	258	231	116
3	Cardiovascular	0	1	2	0	0	0	0	13	2	0	46	29	60	77	94	34
4	Developmental	0	0	0	2	1	1	0	24	17	6	252	220	288	322	323	152
5	Endocrine	0	0	0	0	2	2	9	38	31	6	98	107	145	134	124	61
6	Gastrointestinal	0	0	0	0	1	2	0	20	4	3	61	60	114	131	172	82
7	Hematological and Immune	0	0	1	1	3	2	4	34	26	6	156	158	333	368	377	142
8	Hepatic	0	0	1	2	4	12	13	38	37	13	205	139	145	127	109	46
9	Mortality	0	2	2	2	1	3	9	36	29	5	244	220	201	203	197	96
10	Musculoskeletal	0	0	0	0	1	1	0	10	2	0	47	30	65	87	117	45
11	Neurological	0	0	0	0	0	0	1	19	8	2	83	55	103	116	131	58
12	Nutritional and Metabolic	0	0	0	0	2	8	11	38	36	5	167	137	273	267	273	117
13	Ocular and Sensory	0	0	0	0	0	0	2	9	0	2	52	54	86	101	113	41
14	Renal	0	1	0	3	2	4	4	24	19	8	111	79	104	83	65	33
15	Reproductive	0	0	0	0	0	0	1	13	7	4	82	56	109	91	111	52
16	Respiratory	0	0	1	3	3	0	8	28	12	8	151	134	137	155	141	62
17	Skin and Connective Tissue	0	0	0	0	0	0	0	8	1	0	57	45	70	70	100	34

Priority Ranking Reduces Screening Burden

SciCome



Automated Priority Ranking Reduces Screening Burden

SciCome

TOPIC MODELING

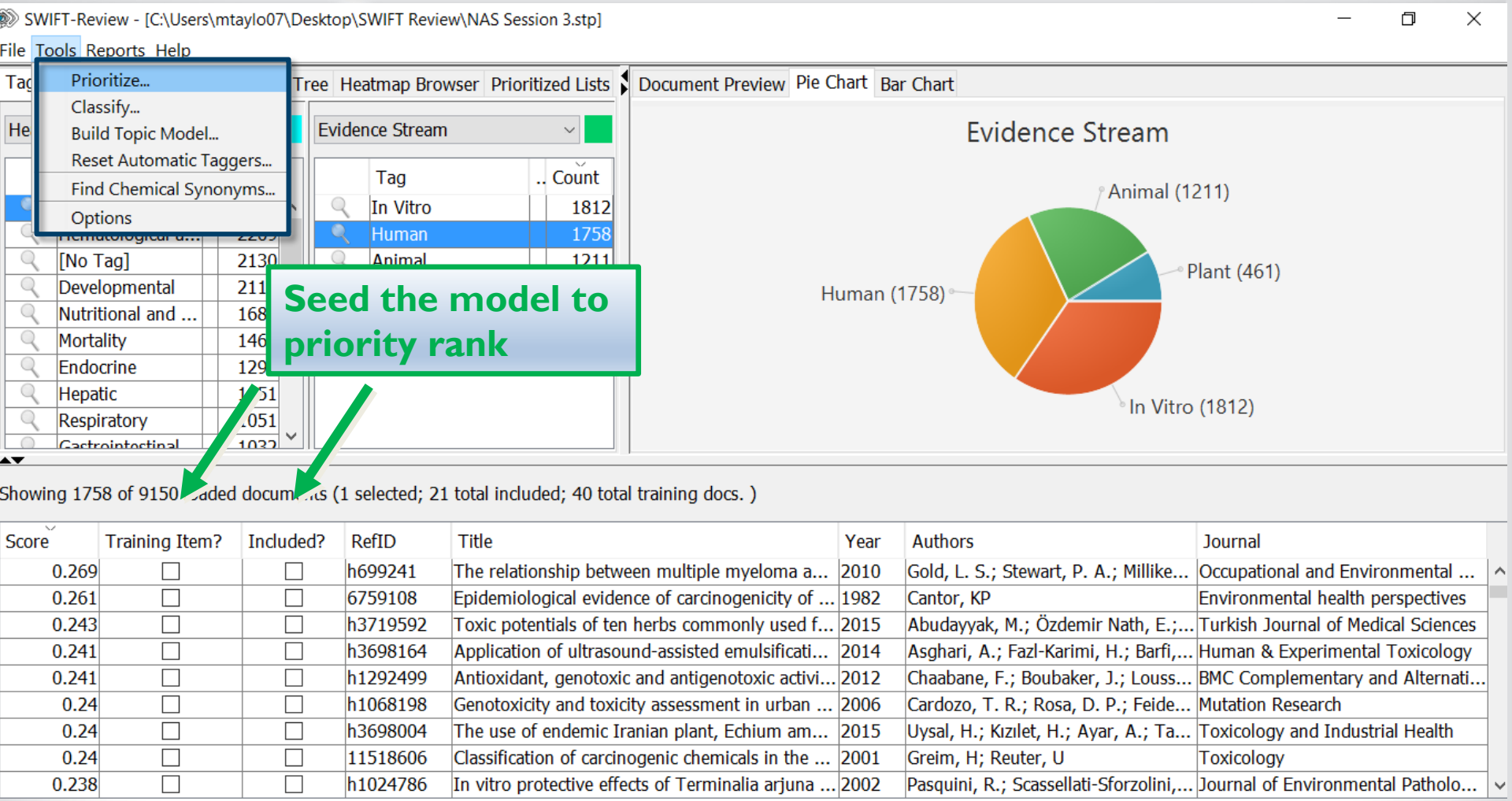
Uncover hidden structure in your literature corpus.

LITERATURE PRIORITIZATION

Use machine learning to triage your reading list.

Topic modeling is a statistical methodology (Latent Dirichlet Allocation or LDA) that automatically computes then categorizes documents according to pre-defined topics. Users can also customize their own topic model by choosing Tools > Build Topic Model

Incorporate human curated training sets or manually annotate **"included"** and **"excluded"** training **"seeds"** to automatically priority rank the remaining documents.



Score	Training Item?	Included?	RefID	Title	Year	Authors	Journal
0.269	☐	☐	h699241	The relationship between multiple myeloma a...	2010	Gold, L. S.; Stewart, P. A.; Millike...	Occupational and Environmental ...
0.261	☐	☐	6759108	Epidemiological evidence of carcinogenicity of ...	1982	Cantor, KP	Environmental health perspectives
0.243	☐	☐	h3719592	Toxic potentials of ten herbs commonly used f...	2015	Abudayyak, M.; Özdemir Nath, E.;...	Turkish Journal of Medical Sciences
0.241	☐	☐	h3698164	Application of ultrasound-assisted emulsificati...	2014	Asghari, A.; Fazl-Karimi, H.; Barfi,...	Human & Experimental Toxicology
0.241	☐	☐	h1292499	Antioxidant, genotoxic and antigenotoxic activi...	2012	Chaabane, F.; Boubaker, J.; Louss...	BMC Complementary and Alternati...
0.24	☐	☐	h1068198	Genotoxicity and toxicity assessment in urban ...	2006	Cardozo, T. R.; Rosa, D. P.; Feide...	Mutation Research
0.24	☐	☐	h3698004	The use of endemic Iranian plant, Echium am...	2015	Uysal, H.; Kızılet, H.; Ayar, A.; Ta...	Toxicology and Industrial Health
0.24	☐	☐	11518606	Classification of carcinogenic chemicals in the ...	2001	Greim, H; Reuter, U	Toxicology
0.238	☐	☐	h1024786	In vitro protective effects of Terminalia arjuna ...	2002	Pasquini, R.; Scassellati-Sforzolini,...	Journal of Environmental Patholo...

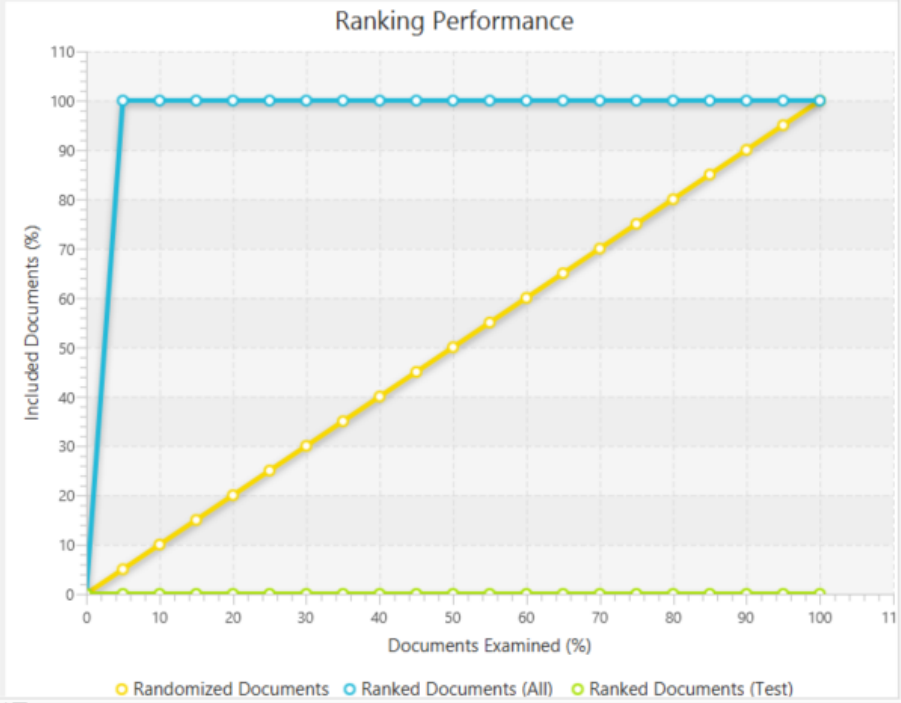
Priority Ranking Improves Literature Screening Efficiency

SWIFT-Review - [C:\Users\mtaylo07\Desktop\SWIFT Review\NAS Session 3.stp]

File Tools Reports Help

Tag Browser | Search | Browse MeSH Tree | Heatmap Browser | Prioritized Lists

Sep 25, 2017 1:29:04 PM



Document Preview | Pie Chart | Bar Chart

Chloroform

IARC,. *(1999)*

▼Abstract

Exposure data. Occupational exposure to chloroform may occur during its production and use as a solvent and chemical intermediate. The general population may be exposed as a result of its presence in chlorinated drinking-water, ambient air and some foods. Human carcinogenicity data. Two cohort studies of cancer and drinking-water quality were carried out in the United States. One conducted in Maryland showed excess mortality from cancers of the liver and breast in association with water chlorination, while that conducted in Iowa showed increased risks for cancers of the colon and lung and skin melanoma associated with chloroform concentrations in drinking-water. Eight case-control studies have been reported on bladder cancer in relation to chlorinated drinking-water in the United States. Significant results were obtained in five studies, but there was little consistency in the risk pattern in subgroups defined by sex or surrogate measures of chloroform intake. Significant increasing trends in the risk for bladder cancer were seen in two studies. The study in Colorado showed increasing risk with years of exposure to chlorinated water; the study in Iowa showed increasing risk with lifetime intake of trihalomethanes (from drinking-water), but only in men and not in women. Seven case-control studies addressed the risk for cancers of the large bowel in association with consumption of chlorinated water. In two of these studies, lifetime exposure to trihalomethanes was assessed. Two studies showed significant associations with rectal cancer. Overall, however, the results were inconsistent with regard to the subsite of the large bowel and sex,

Showing 7701 of 7701 loaded documents (1 selected; 20 total included; 40 total training docs.)

Score	Training Item?	Included?	RefID	Title	Year	Authors	Journal
0.906	☑	☑	h1068322	Chloroform	1991	Meldrum, M.	
0.903	☑	☑	h1065954	Chloroform	1999	IARC,	
0.903	☑	☑	h1022997	Chloroform	1994	IPCS,	

Automated Priority Ranking

sciome

ABOUT BLOG CAREERS SOFTWARE CONTACT

SWIFT-Active Screener

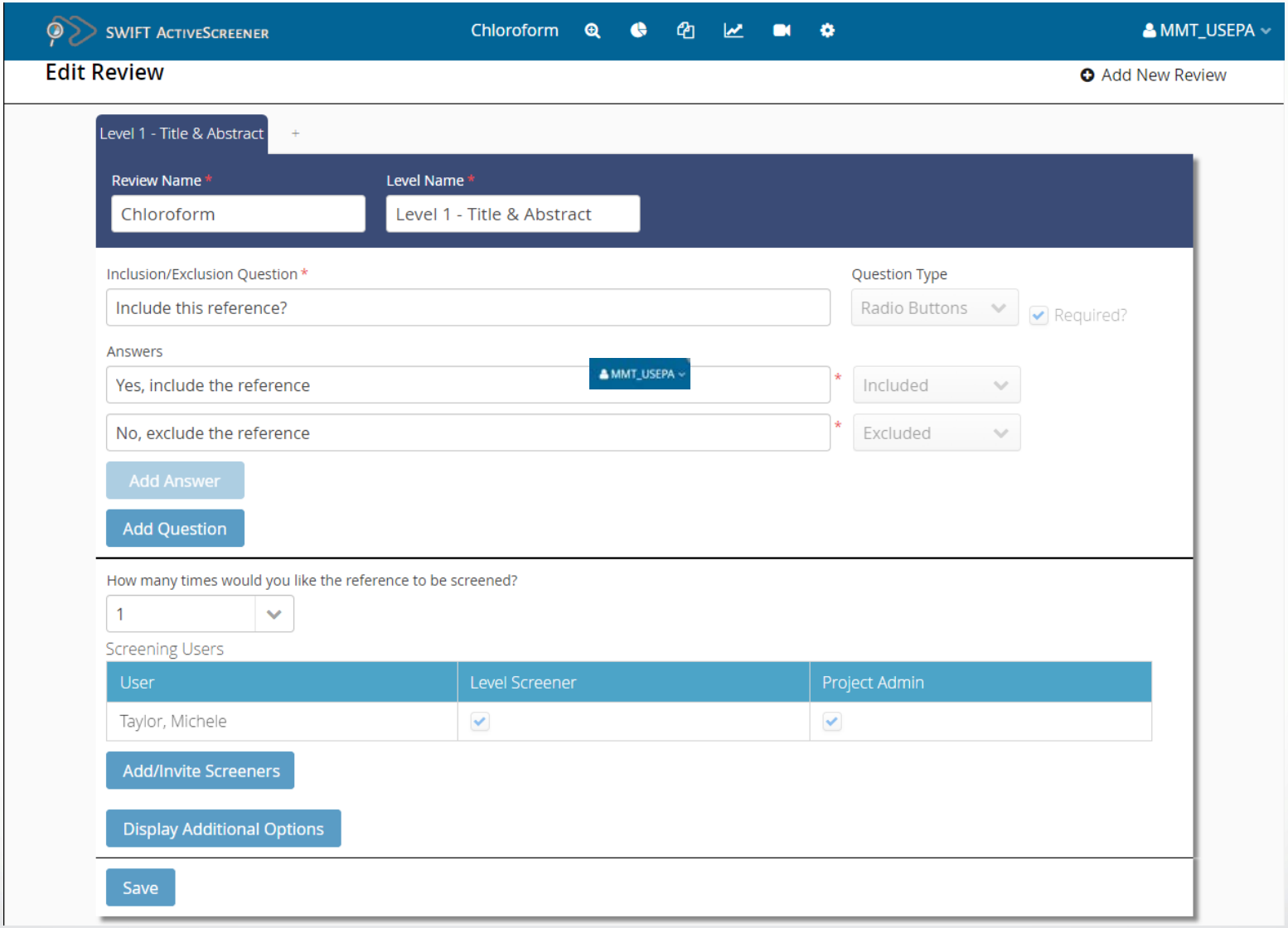
SWIFT-ActiveScreener

SWIFT-Active Screener is a web-based, collaborative systematic review software application. Active Screener was designed to be easy-to-use, incorporating a simple, but powerful, graphical user interface with rich project status updates. What makes Active Screener special, however, is its behind-the-scenes application of state-of-the-art statistical models designed to save screeners time and effort by automatically prioritizing articles as they are reviewed, using user feedback to push the most relevant articles to the top of the list.

SWIFT Active Screener Capabilities - Improved Ranking Model

- Web-based, real-time, collaborative, systematic review software application
- State-of-the-art statistical models prioritize articles ***as they are being reviewed***
- Experience suggests screening burden is reduced by at least 50% (likely more)
- Algorithm improves from screener input without training "seeds" further increasing efficiency (more efficient than implementing a "seed studies" only model)
- Option to "seed" studies if relevant on/off topic literature has been identified
- Incorporates a graphical user interface to provide project status updates
- User-defined screening levels
  - Level 1: Title and Abstract
  - Level 2: Full text screening
  - Level 3: Conflict Resolution

Customize Inclusion/Exclusion Criteria According to the PECO Statement



User Input Improves the Algorithm to Priority Rank While Screening

SWIFT ACTIVESCREENER | Chloroform | MMT_USEPA

Screening References

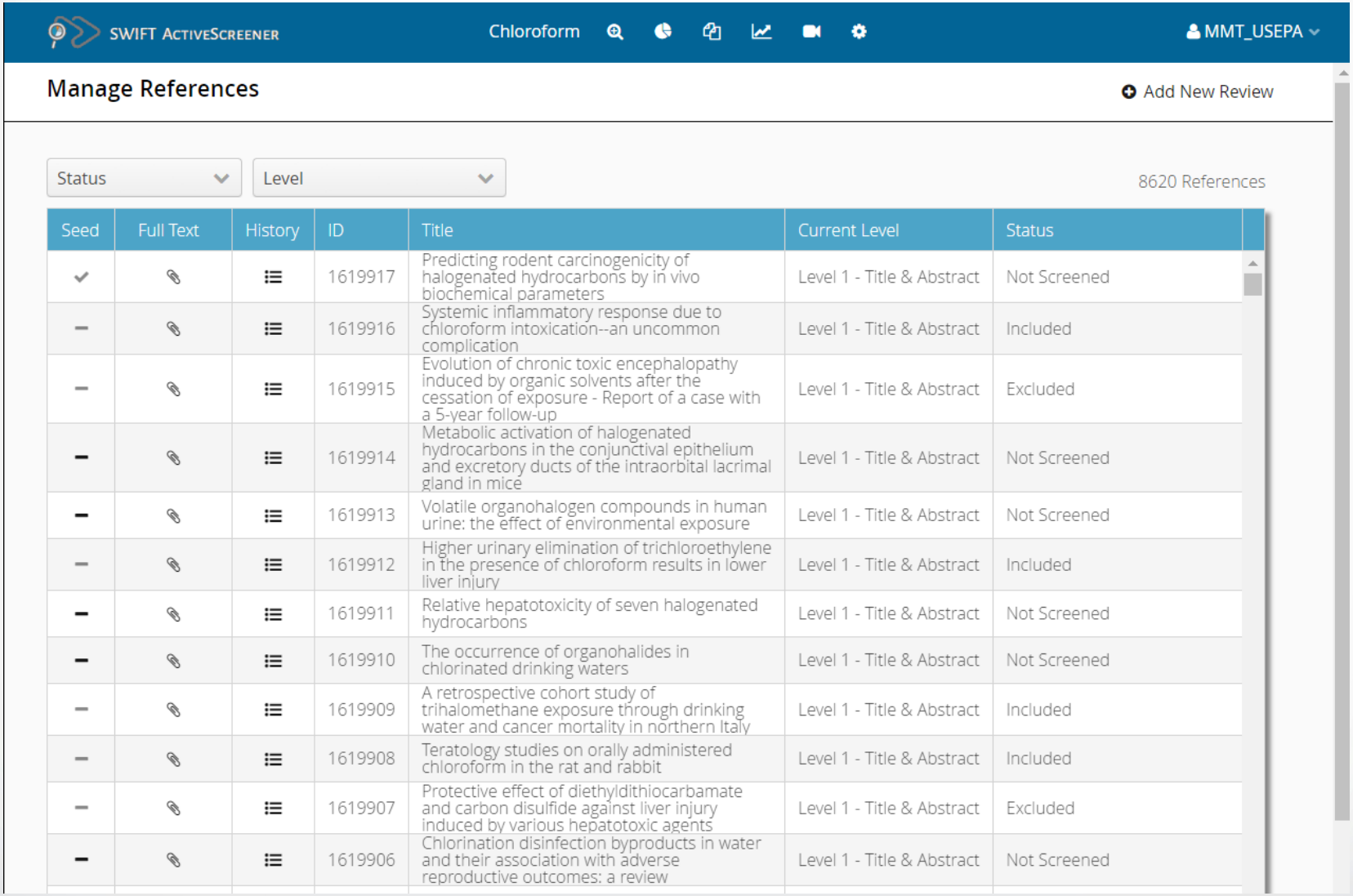
Add New Review

0.3%

Level 1 - Title & Abstract

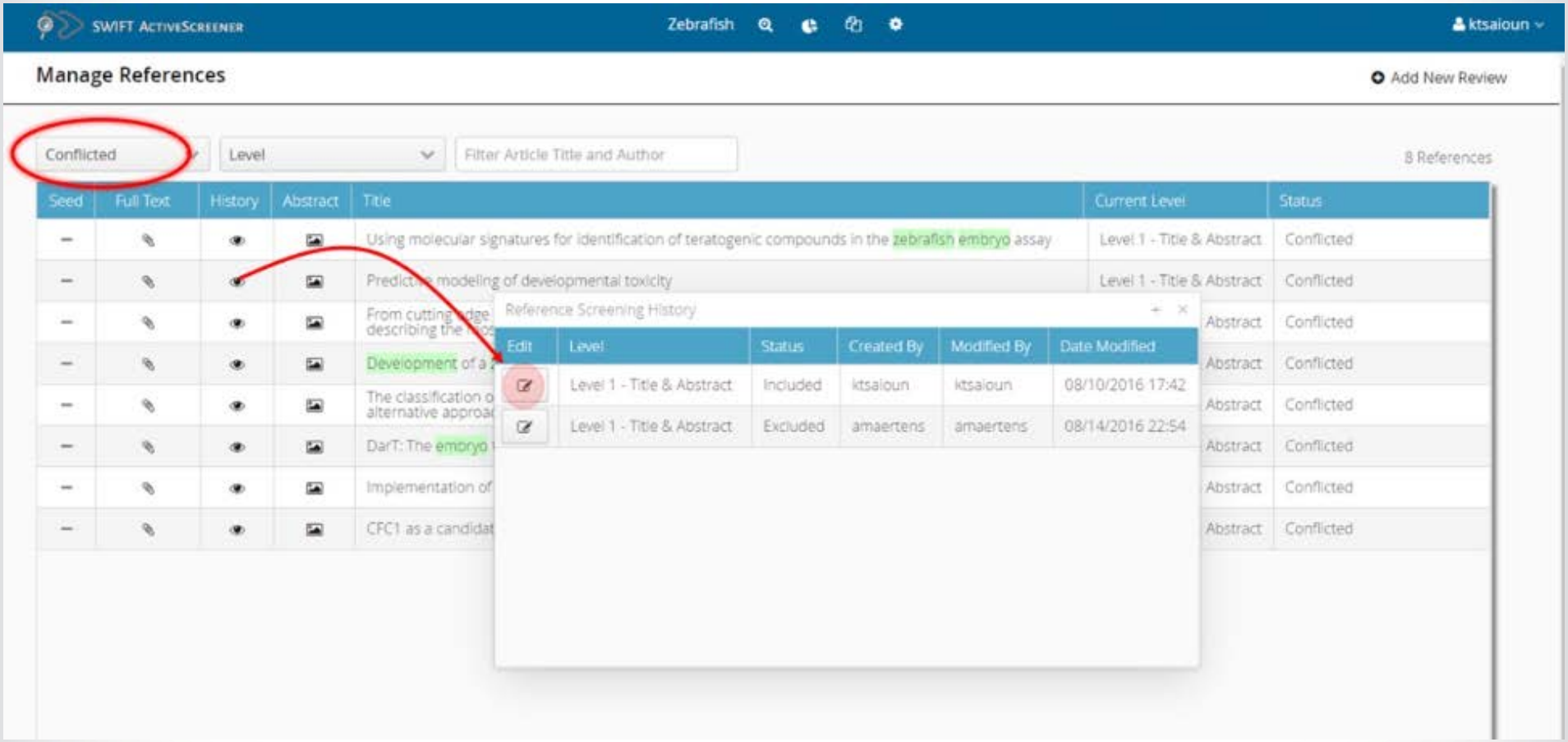
Include	Exclude	Detailed Screen	ID	Title
✓	⊖	🔍	1619902	Hepatotoxicity and lethality of halomethanes in Mongolian gerbils pretreated with chlordecone, phenobarbital or mirex
✓	⊖	🔍	1619890	Central nervous effect and blood level regressions on exposure time paralleled in solvents (toluene, carbon tetrachloride and chloroform)
✓	⊖	🔍	1619889	DNA damage as a consequence of chloroform-induced cytotoxicity in male F344 rat and B6C3F1 mouse hepatocytes in vitro
✓	⊖	🔍	1619865	Suppression of pulmonary host defenses and enhanced susceptibility to respiratory bacterial infection in mice following inhalation exposure to trichloroethylene and chloroform
✓	⊖	🔍	1619857	Carcinoembryonic antigen, alpha-fetoprotein, and prostate-specific antigen in the sera of industrial workers exposed to phenol, formaldehyde, urea, and mixed vapors
✓	⊖	🔍	1619856	Ranking of chemicals for carcinogenic potecya comparative study of 13 carcinogenic chemicals and an examination of some of the issues involved
✓	⊖	🔍	1619852	Metabolism of haloforms to carbon monoxide: II in vivo studies
✓	⊖	🔍	1619846	A painless burn: systemic toxicity after dermal exposure to chloroform
✓	⊖	🔍	1619844	U.S. Environmental Protection Agency's revised cancer guidelines for carcinogen risk assessment

"Seed" studies when Relevant On/Off Topic Literature is Identified



Seed	Full Text	History	ID	Title	Current Level	Status
✓			1619917	Predicting rodent carcinogenicity of halogenated hydrocarbons by in vivo biochemical parameters	Level 1 - Title & Abstract	Not Screened
–			1619916	Systemic inflammatory response due to chloroform intoxication--an uncommon complication	Level 1 - Title & Abstract	Included
–			1619915	Evolution of chronic toxic encephalopathy induced by organic solvents after the cessation of exposure - Report of a case with a 5-year follow-up	Level 1 - Title & Abstract	Excluded
–			1619914	Metabolic activation of halogenated hydrocarbons in the conjunctival epithelium and excretory ducts of the intraorbital lacrimal gland in mice	Level 1 - Title & Abstract	Not Screened
–			1619913	Volatile organohalogen compounds in human urine: the effect of environmental exposure	Level 1 - Title & Abstract	Not Screened
–			1619912	Higher urinary elimination of trichloroethylene in the presence of chloroform results in lower liver injury	Level 1 - Title & Abstract	Included
–			1619911	Relative hepatotoxicity of seven halogenated hydrocarbons	Level 1 - Title & Abstract	Not Screened
–			1619910	The occurrence of organohalides in chlorinated drinking waters	Level 1 - Title & Abstract	Not Screened
–			1619909	A retrospective cohort study of trihalomethane exposure through drinking water and cancer mortality in northern Italy	Level 1 - Title & Abstract	Included
–			1619908	Teratology studies on orally administered chloroform in the rat and rabbit	Level 1 - Title & Abstract	Included
–			1619907	Protective effect of diethyldithiocarbamate and carbon disulfide against liver injury induced by various hepatotoxic agents	Level 1 - Title & Abstract	Excluded
–			1619906	Chlorination disinfection byproducts in water and their association with adverse reproductive outcomes: a review	Level 1 - Title & Abstract	Not Screened

Manage References with Conflict Resolution – Track and Archive Changes



SWIFT ACTIVESCREENER Chloroform MMT_USEPA

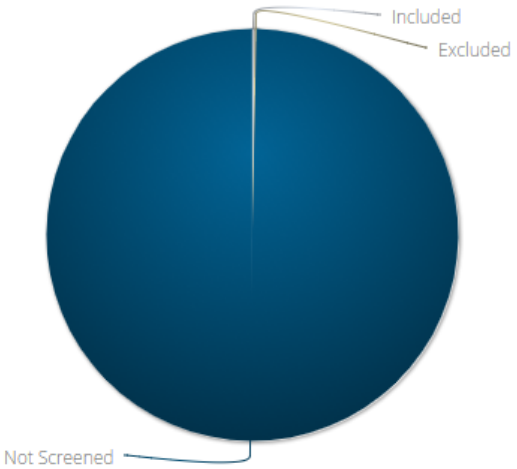
Review Summary

Add New Review

Viewing Project As Taylor, Michele

Chloroform | Level 1 - Title & Abstract

User's Screening Status for Level 1 - Title & Abstract



References: 8620

Screened: 29 Not Screened: 8591

User's Screening Progress for Level 1 - Title & Abstract

ActiveScreener model built at: 09/25/2017 15:54

Reviewers

Taylor, Michele

SWIFT Active: Data Integration

- **Active Screener integrates with systematic review tools already in use:**
  - Accepts imports from bibliographic databases and reference curation platforms including *SWIFT Review*, EndNote, Mendeley, Zotaro, and PubMed
  - Results from screening in Active Screener can be exported in standard data formats compatible with applications including *HAWC* and Excel, EndNote, Mendeley, and Zotaro

Current Users

National Institute of Environmental Health Sciences (NIEHS)

United States Environmental Protection Agency (EPA)

The Endocrine Disruption Exchange (TEDX)

Johns Hopkins Evidence-Based Toxicology Collaboration (EBTC)

United States Department of Agriculture (USDA)

HAWC: Study Evaluation, Extraction, Visualization and Data Sharing

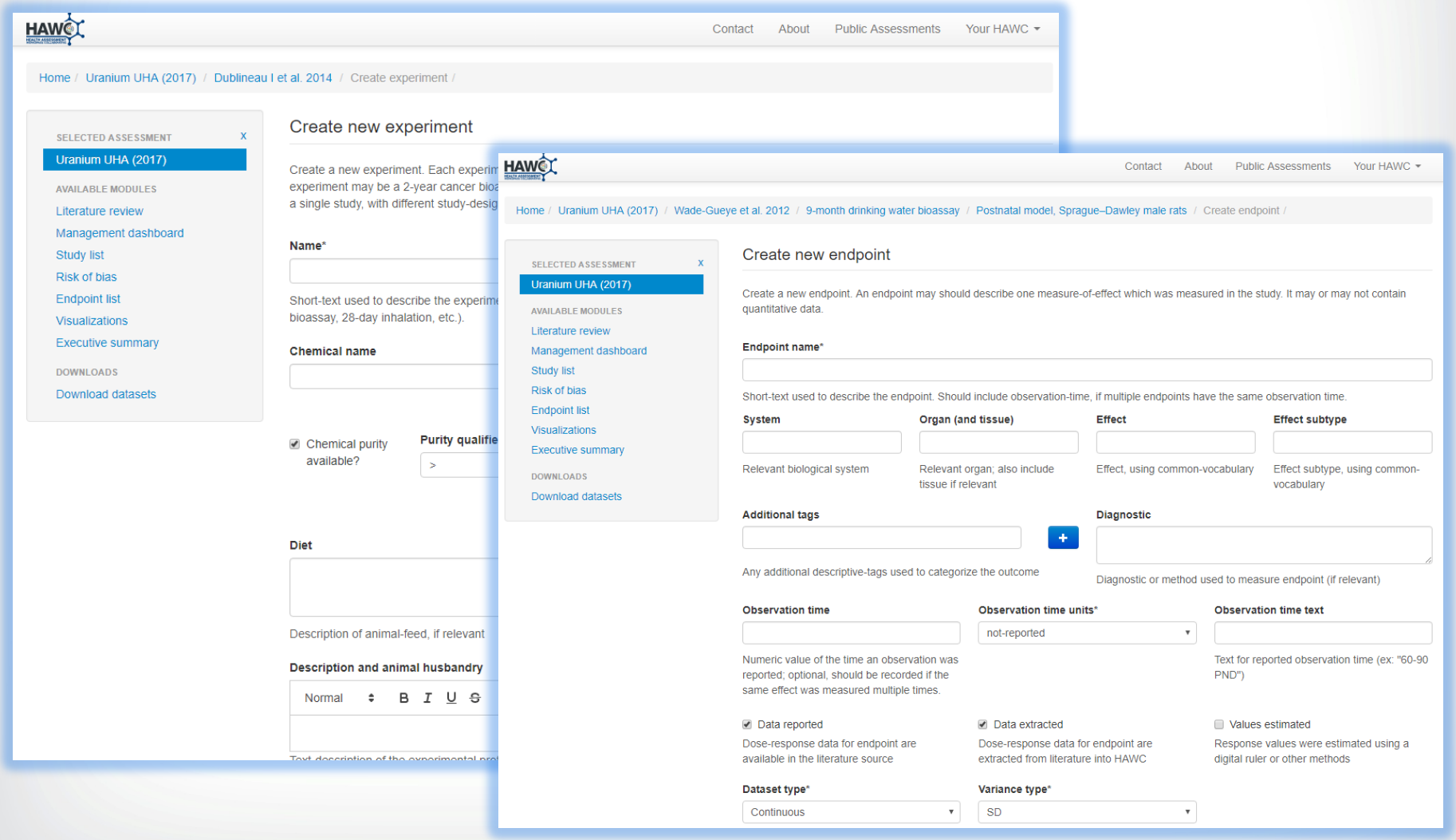
https://hawcproject.org/

HAWC Capabilities

- **Free and open source**
- **Developed at UNC by Andy Shapiro* with Ivan Rusyn**
- **Literature search and initial screening**
- **Animal bioassay, epidemiological, and in vitro structured study methods/data extraction and visualization**
- **Interactive "click to see more" graphics**
- **Risk of bias and sensitivity evaluation**
- **Modular to work with other tools and maximize flexibility for users**
- **Works best in Google Chrome (preferred), Mozilla Firefox, and Safari**

*current affiliation is National Institute of Environmental Health Sciences/National Toxicology Program (NIEHS/NTP)

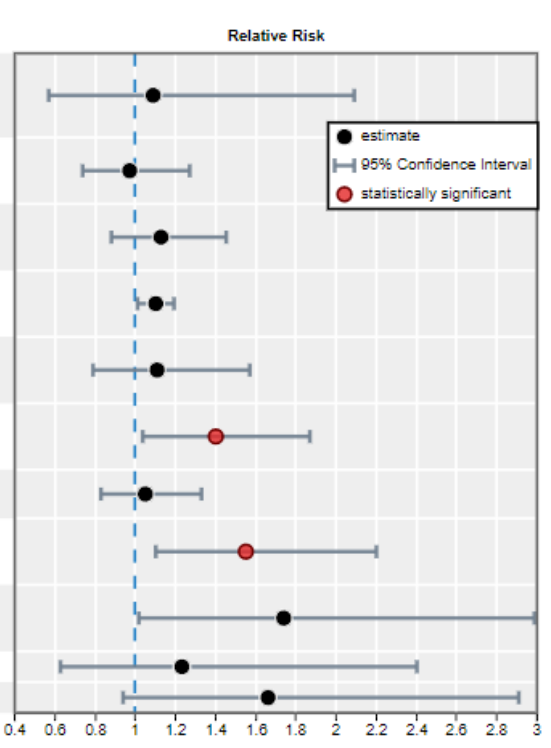
HAWC: Summarizing Animal Bioassays



Epidemiology: Click to See More Display

Example from Chloroform

Study	Population	Design	Exposure	Outcome	Comments	Exposure Contrast (EWPM)	adjOR*
Brender et al. 2014	USA (TX 1996-2008) case-control, 9 (60,613 case-mothers; 244,927 control-mothers)	Case-control	maternal residential proximity (EWPM based on residential distances to industrial sources and TRI data)	anencephaly	3,985 (1.6%) controls and 10 (1.7%) cases in >0 group	0 vs >0	1.09
				heart defect (conotruncal)	3,985 (1.6%) controls and 60 (1.8%) cases in >0 group	0 vs >0	0.97
				heart defect (obstructive)	3,985 (1.6%) controls and 70 (2.1%) cases in >0 group	0 vs >0	1.13
				heart defect (septal)	3,985 (1.6%) controls and 739 (1.7%) cases in >0 group	0 vs >0	1.1
				limb deficiency (any)	3,985 (1.6%) controls and 37 (1.9%) cases in >0 group	0 vs >0	1.11
				neural tube defects	3,985 (1.6%) controls and 52 (2.4%) cases in >0 group	0 vs >0	1.4
				oral cleft (any)	3,985 (1.6%) controls and 79 (2%) cases in >0 group	0 vs >0	1.05
				spina bifida	3,985 (1.6%) controls and 37 (2.8%) cases in >0 group	0 vs >0	1.55
					p-trend = 0.027; 9-14 (0.7-1.1%) exposed cases per group	0 vs 0.01-42.27	1.74
						0 vs 42.28-1490.26	1.23
						0 vs >1490.26	1.66



Visualizing Epidemiology Evidence

Example from Chloroform

Study	Population	Design	Exposure	Outcome	Comments	Exposure Contrast	adjOR*
Brender et al. 2014	USA (TX 1996-2008) case-control, ♀♂ (60,613 case-mothers; 244,927 control-mothers)	Case-control	maternal residential proximity (EWPM based on residential distances to industrial sources and TRI data)	anencephaly	3,985 10 (1. group		
				heart defect (conotruncal)	3,985 60 (1. group		
				heart defect (obstructive)	3,985 70 (2. group		
				heart defect (septal)	3,985 739 (group		
				limb deficiency (any)	3,985 37 (1. group		
				neural tube defects	3,985 52 (2. group		
				oral cleft (any)	3,985 79 (2		
				spina bifida	3,985 37 (2. group		
					p-tren (0.7-1 per gr		

Brender et al. 2014 / USA (TX 1996-2008) case-control, ♀♂ (60,613 case-mothers; 244,927 control-mothers) / birth defects (registry)

limb deficiency (any) | spina bifida | heart defect (septal)

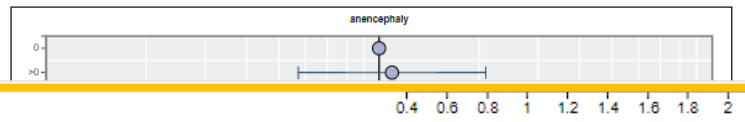
Results	anencephaly
Comparison set	0 vs >0
Data location	Table 2
Population description	3,985 (1.6%) controls and 10 (1.7%) cases in >0 group
Metric Description	adjOR
Adjustment factors	• birth year • geographic • maternal age • race/ethnicity
Dose response	not-applicable
Statistical power	not reported or calculated
Prevalence incidence	3,985 (1.6%) controls and 10 (1.7%) cases in >0 group
Comments	3,985 (1.6%) controls and 10 (1.7%) cases in >0 group

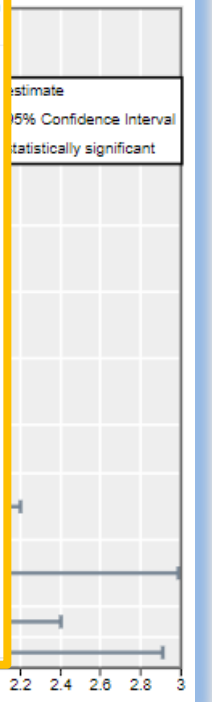
Results by group

Group	N	Estimate	95% confidence intervals	*p*-value
0	-	1	-	-
>0[a]	3995	1.09	0.57 - 2.09	n.s.

[a] Main finding as selected by HAWC assessment authors (not-supportive).

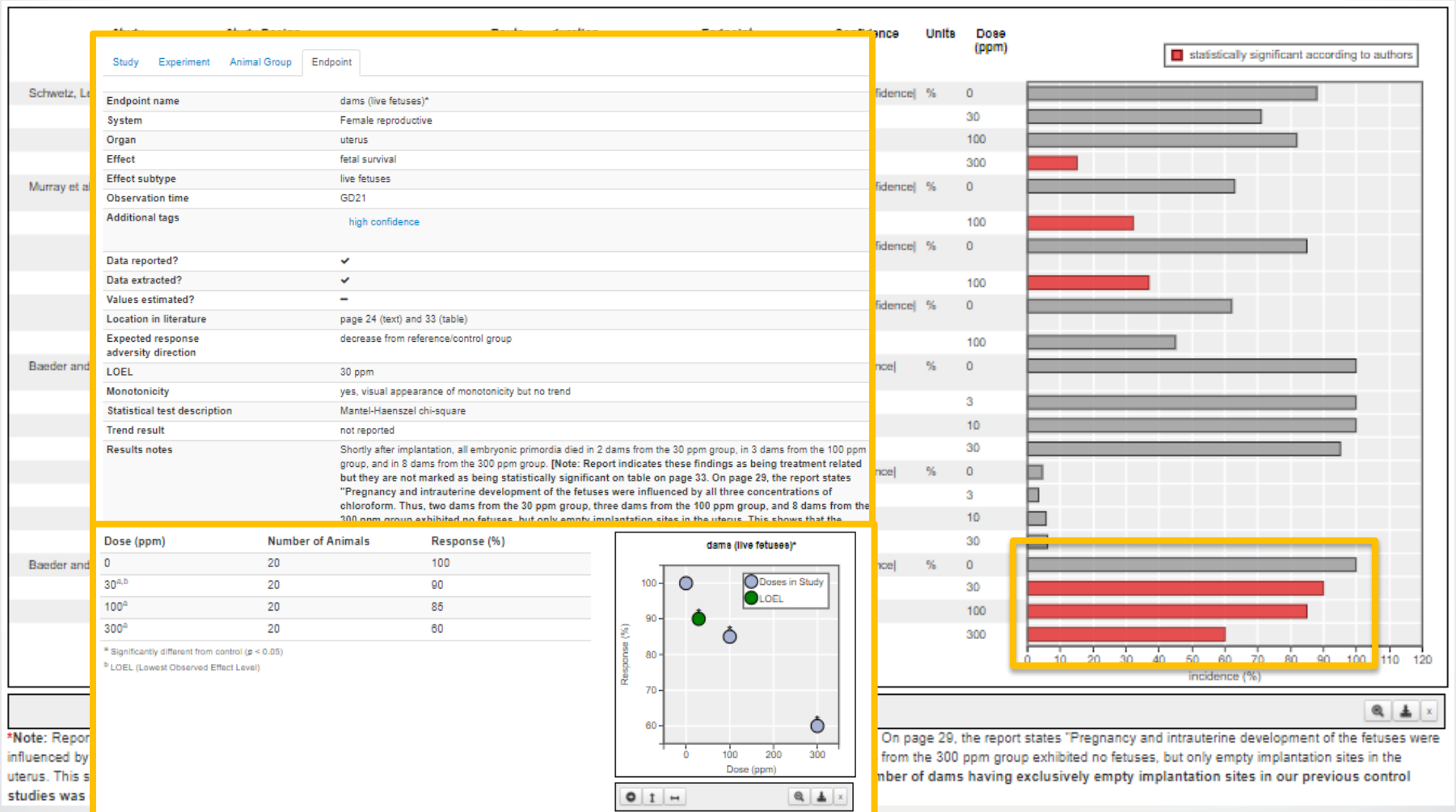
Forest plot





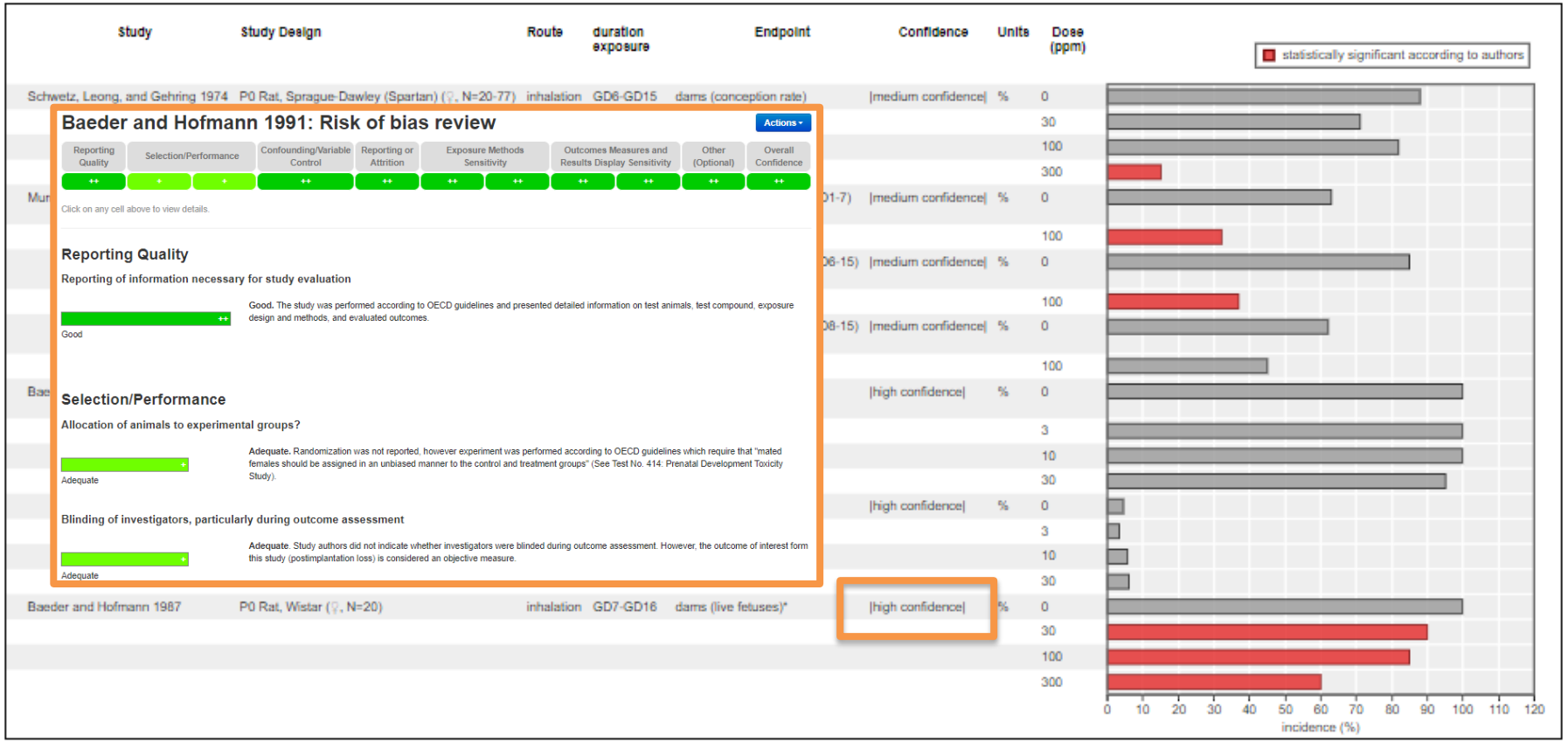
Visualizing Animal Evidence

Chloroform Fetal Survival



Visualizing Animal Evidence

Chloroform Fetal Survival

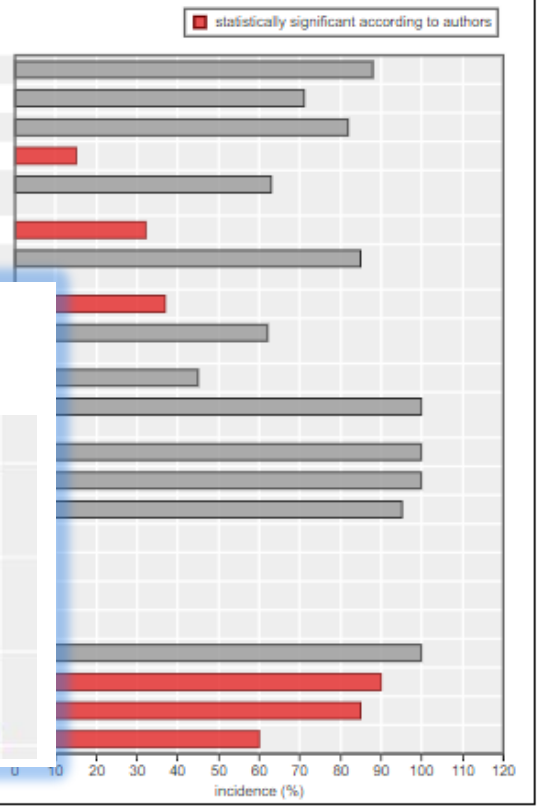


*Note: Report indicates these findings as being treatment related but they are not marked as being statistically significant on table on page 33. On page 29, the report states "Pregnancy and intrauterine development of the fetuses were influenced by all three concentrations of chloroform. Thus, two dams from the 30 ppm group, three dams from the 100 ppm group, and 8 dams from the 300 ppm group exhibited no fetuses, but only empty implantation sites in the uterus. This shows that the embryos died shortly after implantation. In the control group, all dams carried live fetuses til delivery. **Since the number of dams having exclusively empty implantation sites in our previous control studies was only a single dam of a total of 1,275 animals, this condition must be due to the exposure to chloroform.**"

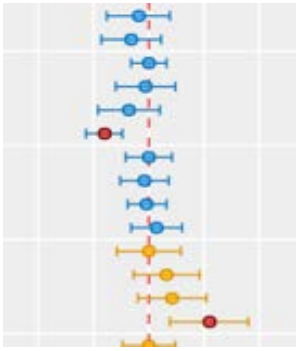
Multiple Formats to Present Results

Chloroform Fetal Survival

Study	Study Design	Route	duration exposure	Endpoint	Confidence	Units	Dose (ppm)
Schwetz, Leong, and Gehring 1974	P0 Rat, Sprague-Dawley (Spartan) (♀, N=20-77)	inhalation	GD6-GD15	dams (conception rate)	[medium confidence]	%	0
							30
							100
							300
Murray et al. 1979	P0 Mouse, Cf-1 (♀, N=34-35)	inhalation	GD1-GD7, 7 hr/day	dams (conception rate, GD1-7)	[medium confidence]	%	0
							100
			GD6-GD15, 7 hr/day	dams (conception rate, GD6-15)	[medium confidence]	%	0
	P0 Mouse, Cf-1 (♀, N=40)	inhalati					
Baeder and Hofmann 1991	P0 Rat, Wistar (♀, N=20)	inhalati					
Baeder and Hofmann 1987	P0 Rat, Wistar (♀, N=20)	inhalati					

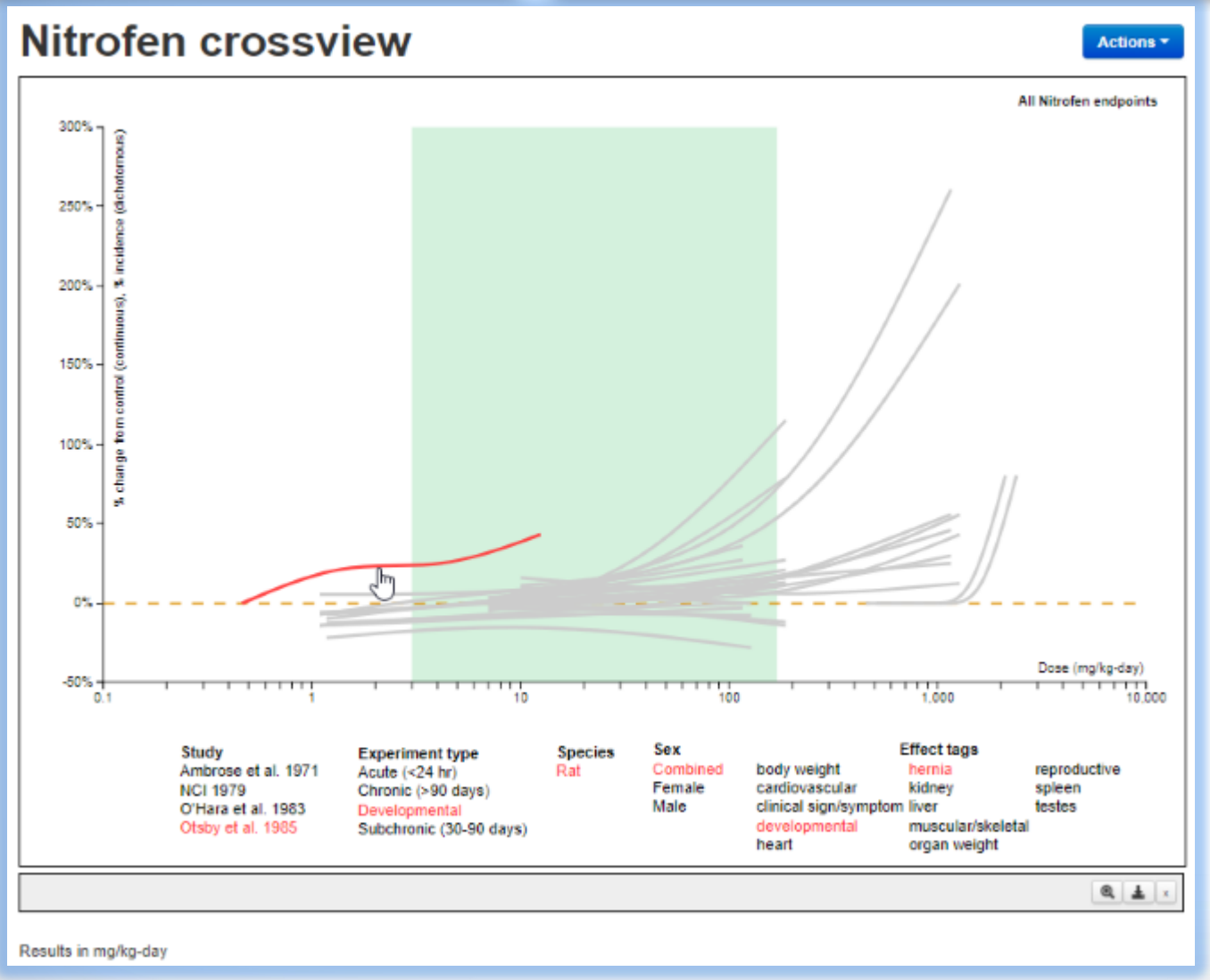


Animal data can be expressed as effect size, e.g., percent control

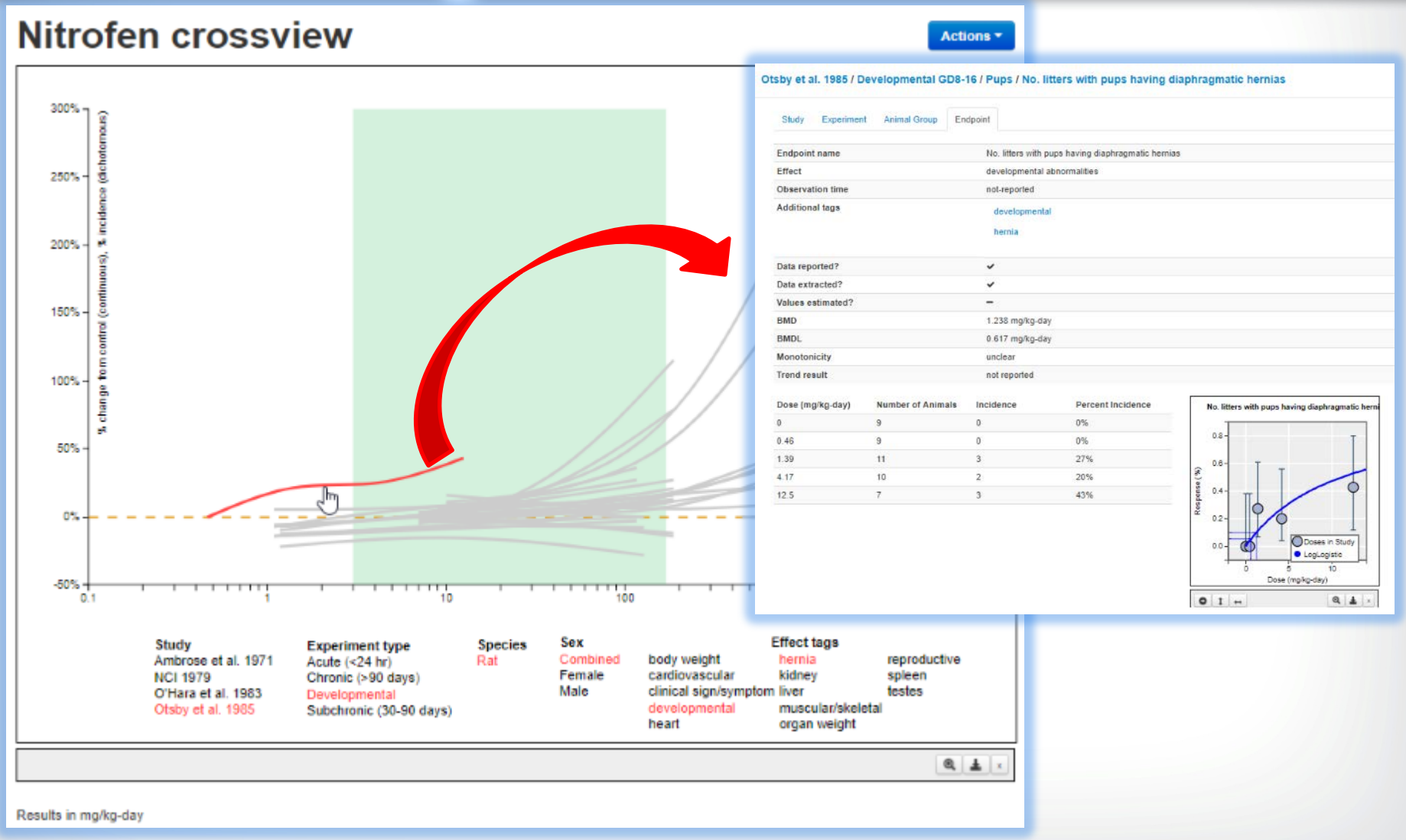


*Note: Report indicates these findings as being treatment related but they are not marked as being statistically significant on table on page 33. On page 29, the report states "Pregnancy and intrauterine development of the fetuses were influenced by all three concentrations of chloroform. Thus, two dams from the 30 ppm group, three dams from the 100 ppm group, and 8 dams from the 300 ppm group exhibited no fetuses, but only empty implantation sites in the uterus. This shows that the embryos died shortly after implantation. In the control group, all dams carried live fetuses til delivery. **Since the number of dams having exclusively empty implantation sites in our previous control studies was only a single dam of a total of 1,275 animals, this condition must be due to the exposure to chloroform.**"

HAWC: Dose-Response Displays

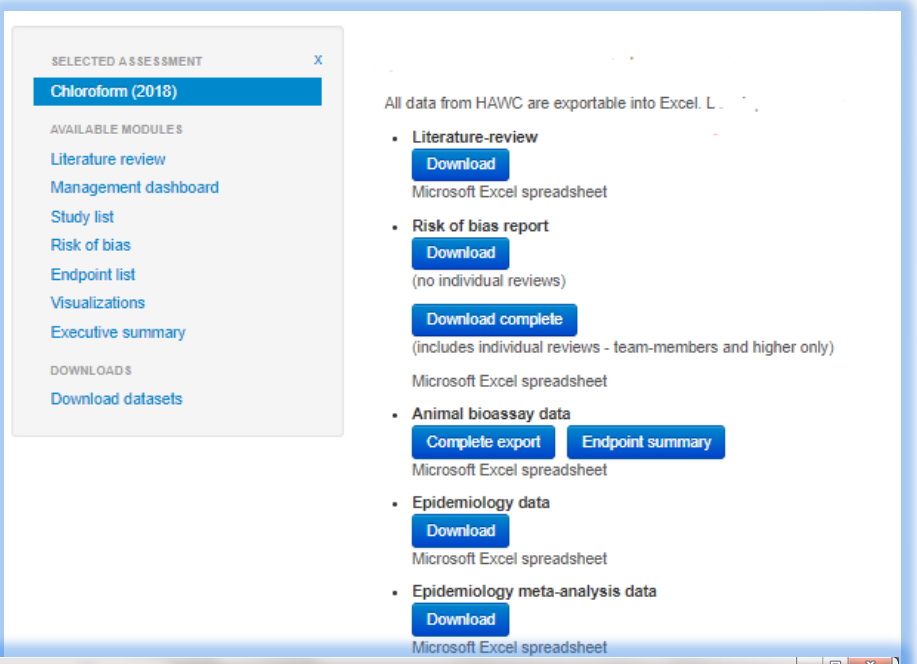


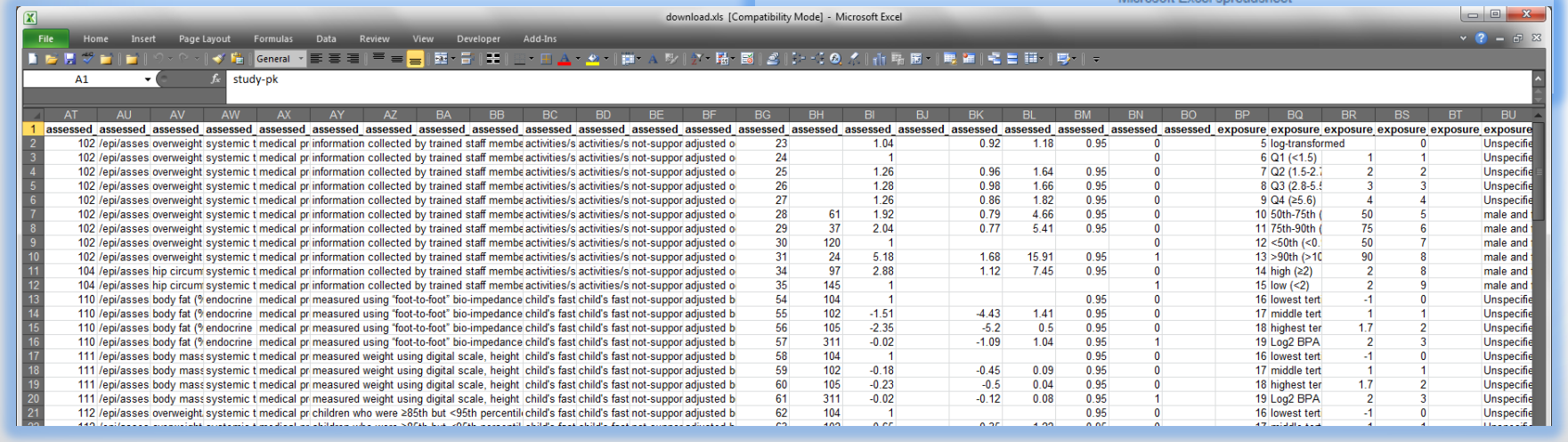
HAWC: Dose-Response Displays



HAWC: Download Reports

- **Entire database for an assessment can be downloaded in Microsoft Excel exports**





HAWC Benchmark Dose Modeling

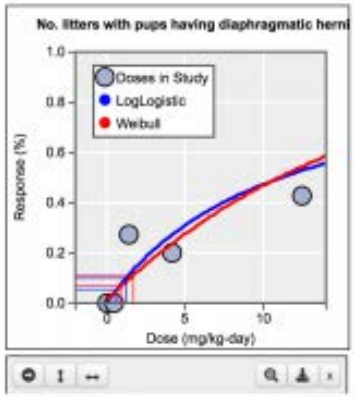
Benchmark dose modeling

BMD setup | Results | Model recommendation and selection

BMDS output summary

Model	Global p-value	AIC	BMD (10%)	BMDL (10%)	Residual of interest	Output
Logistic	0.1974	42.2346	4.60161	2.81096	0.222	View
LogLogistic	0.4283	37.9348	1.23831	0.617243	1.709	View
Probit	0.1974	42.1248	4.26256	2.62289	0.185	View
LogProbit	0.1432	42.643	4.1219	1.71245	0.103	View
Multistage	0.1648	40.7611	1.70173	0.930926	2.039	View
Gamma	0.1648	40.7611	1.70173	0.930926	2.039	View
Weibull	0.1647	40.7611	1.70175	0.930926	2.04	View

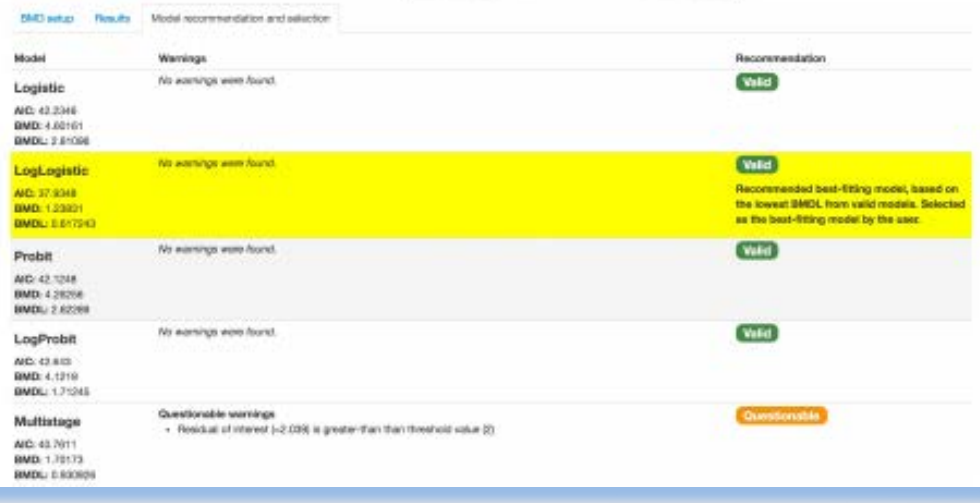
Selected model (if any) highlighted in yellow



The USEPA BMDS software is integrated into HAWC for fitting dose-response models to animal bioassay data.

Model recommendations also performed using BMDS technical guidance and automation approaches from Wignall et al. 2014.

Selections must be done manually.



Advantages

- Structured extraction to promote consistency and completeness
- Free, open source and customizable
- Enhance opportunities for database interpretability
- Integration with automated data-extraction tools
- Web-based to promote team collaboration
- Ability to export data files promotes further analysis of findings and quantification (in assessments or for methods development)
- Creates possibilities for web-based, interactive reports

IRIS has Addressed the Major NAS 2014 Recommendations

NAS 2014 Topics	IRIS Process Improvements
Looking Forward	• Specialized software tools for efficiency and more user friendly and transparent formats for evidence display have been adopted • Strategic planning on use of text and data-mining tools and automation • Specialized tools facilitate transparent documentation, consistency across assessments, and database interoperability

See Demonstrations:

- SWIFT Review and SWIFT Active
- Health Assessment Workspace Collaborative
- Heath Effects Research Online

EPA
United States
Environmental Protection
Agency

SESSION 4: STUDY SELECTION FOR DEVELOPING TOXICITY VALUES, AND ADVANCING RESEARCH ON QUANTITATIVE ANALYSES FOR EVIDENCE INTEGRATION AND DOSE-RESPONSE ANALYSES

David Bussard*, Jason Lambert*, Ted Berner, Allen Davis, Jeff Gift, Karen Hogan, Leonid Kopylev, Ravi Subramaniam

[*Speaking]

Office of Research and Development
NCEA, IRIS

NAS 2014: Three High Priority (Box 8-1) Recommendations on Quantification

- **TOXICITY VALUES**: "EPA should develop criteria for determining when evidence is sufficient to derive toxicity values."
  - Overall hazard conclusions inform decision whether to develop toxicity values.
  - Better documenting considerations on which studies are carried forward to dose-response.

- **POINTS OF DEPARTURE (PODs)**: "EPA should clearly present two dose-response estimates: a central estimate (such as a maximum likelihood estimate or a posterior mean) and a lower-bound estimate for a POD from which a toxicity value is derived."
  - Central estimates (MLEs) of BMDs provided in IRIS assessments along with BMDLs.
  - Will start to use WHO/IPCS approach to characterize distributions in final values.
  - Model averaging to characterize model uncertainty.

- **QUANTITATIVE CAPABILITIES**: "EPA should expand its ability to perform quantitative modeling of evidence integration; in particular, it should develop the capacity to do Bayesian modeling of chemical hazards. ...The Committee emphasizes that... IRIS assessments should not be delayed while this capacity is being developed."
  - Meta-analysis of human and animal studies increasing: hazard decisions and dose-response.
  - Bayesian methods are being explored to help characterize uncertainty.
  - New approach methods and assays are increasingly being evaluated quantitatively.

Evidence Integration Conclusions Inform when to Develop Toxicity Values

Evidence integration conclusion	Quantitative toxicity value provided?
Strongest conclusion for a human health effect (for cancer, a descriptor of *Known*)	Yes.
Moderately strong conclusion for a human health effect (for cancer, a descriptor of *Likely*)	Yes.
Weakest conclusion for a human health effect (for cancer, a descriptor of *Suggestive*)	Determined by situation (e.g., may provide values when useful for decision purpose and the evidence includes a well-conducted study)
Inadequate information	No, although bounding estimate from a study that does not show positive results can be derived where useful for decision purpose.
Strong support for no human health effect	No.

Decision-Making for Advancing Studies to Develop Toxicity Values

IRIS has further clarified the considerations that inform the selection of studies to estimate human dose-response relationships (next slide).

- IRIS continues to find that this decision process is not reducible to a formula.
- Expert judgment is essential for judging the relative merits of individual studies and which studies support more integrative quantitative analyses (e.g., meta-analysis).
- IRIS must often utilize studies with a range of attributes and levels of reporting. For example, the available studies on many mission-critical chemicals do not provide data on an individual subject basis.
- For full transparency, IRIS continues to emphasize documentation of the factors it weighed in emphasizing certain studies, or combinations of studies, over others.

More Explicitly Defining the Attributes IRIS Uses to Evaluate Studies for Derivation of Toxicity Values

In addition to qualitative study evaluation judgments (i.e., *medium* or *high* confidence studies are preferred), studies are assessed across several study attributes

Example Primary Considerations for Selection of Studies for Derivation of Toxicity Values			
Study attribute		**Human studies**	**Animal studies**
Test species		**Human data are generally preferred** to eliminate interspecies extrapolation uncertainties (e.g., in toxicodynamics and specific health outcomes).	**Animals that respond most like humans are preferred**. Outcomes associated with species known to show differences in sensitivity can provide support with suitable qualification.
Human relevance of the exposure paradigm	**Exposure route**	Studies involving **typical human environmental exposure routes are preferred** (e.g., oral, inhalation). A validated toxicokinetic model can be used to extrapolate across exposure routes.	
	Exposure duration	For chronic toxicity values, **chronic or subchronic studies are preferred**. **Exceptions exist** (e.g., when a population or lifestage is more sensitive during a particular time window)	
	Exposure levels	**Exposures near the range of typical environmental human exposures are preferred. Studies with a broad exposure range and multiple exposure levels are preferred** to the extent that they can provide information about the shape of the exposure-response relationship* and facilitate extrapolation to more relevant (generally lower) exposures.	
Susceptibility		Studies that yield **risk estimates in the most susceptible groups are preferred**. Inclusion of design features in the analysis (e.g., matching procedures, blocking; covariates or other procedures for statistical adjustment) **that adequately address the relevant sources of potential critical confounding for a given outcome are preferred**.	

*U.S. EPA Benchmark Dose Technical Guidance (2012)

IRIS Assessments Are Providing Central MLE Estimates of BMDs Along with BMDLs

Table 2-1. Summary of derivation of PODs

Endpoint and reference	Species/ sex	Model[a]	BMR	BMD mg/kg-d	BMDL mg/kg-d	POD mg/kg-d	POD_{ADJ}[b] mg/kg-d	POD_{HED}[c] mg/kg-d
Developmental								
Neurobehavioral changes: Open field crossed squares at PND 69 Chen et al. (2012)	Male and Female Sprague-Dawley rats	Exponential 4	1 SD	0.23	0.11			
Neurobehavioral changes: Elevated plus maze open arm entries at PND 70 Chen et al. (2012)	Female Sprague-Dawley rats	Exponential 4	1 SD	0.21	0.092	0.092[d]	0.092	0.092
Neurobehavioral changes: Morris water maze hidden platform trial escape latency at PNDs 71–74 Chen et al. (2012)	Male and Female Sprague-Dawley rats	Hill CV Hill CV Hill CV Hill NCV	1 SD (9 sec)	PND71: 0.49 PND72: 0.33 PND73: 0.27 PND74: 0.23	0.16 0.16 0.12 0.13			
Cardiovascular effects at PND 53 Jules et al. (2012)	Long-Evans rats	LOAEL (0.6 mg/kg-d) (15% ↑ in systolic blood pressure; 33% ↑ in diastolic blood pressure)				0.6	0.6	0.15

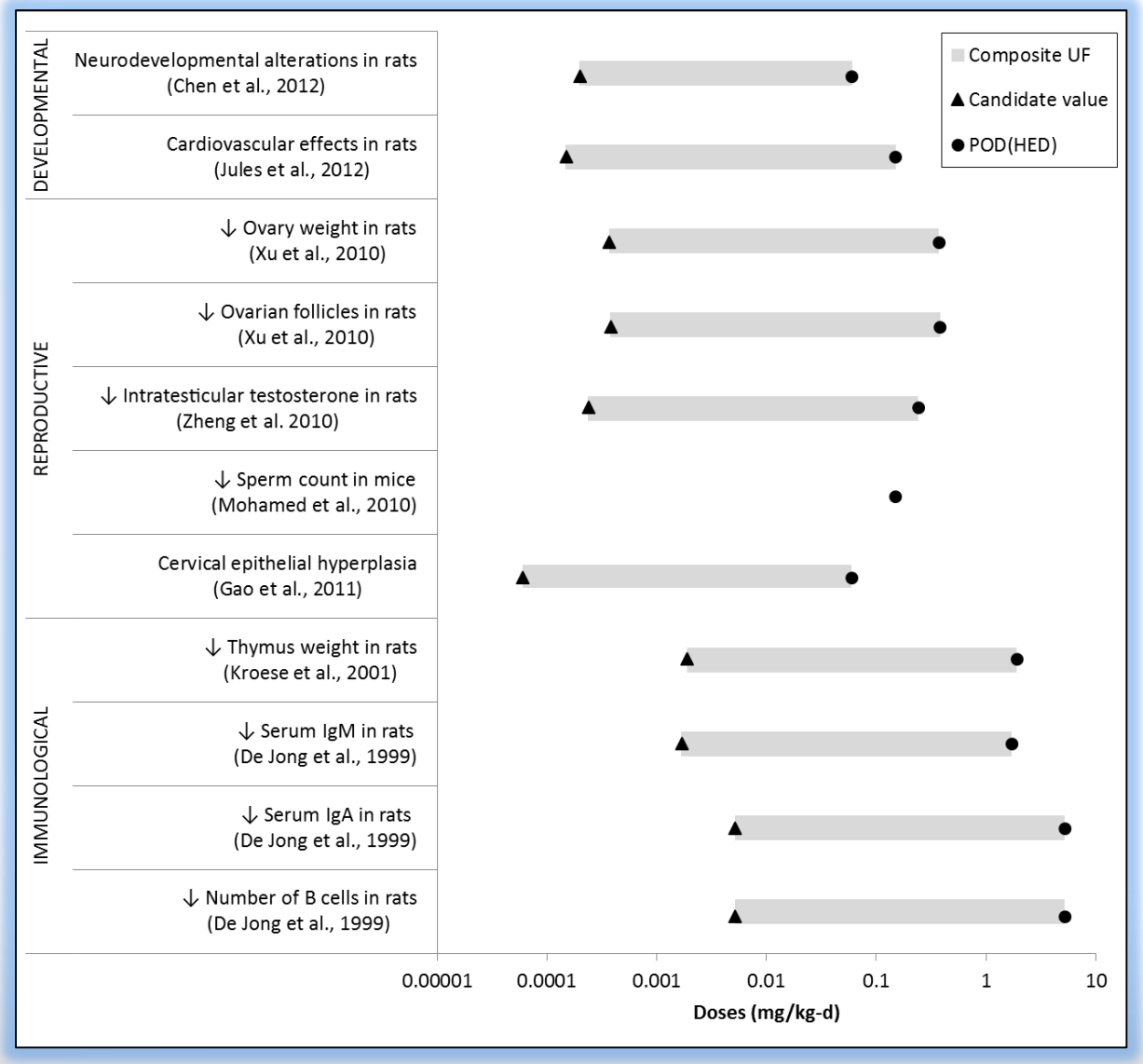
Recent animal study example to the left: **Benzo[a]pyrene (EPA, 2017)**

p. 2-8 Toxicological Review of Benzo[a]pyrene https://cfpub.epa.gov/ncea/iris/iris_documents/documents/toxreviews/0136tr.pdf

Recent epidemiology example: **Ethylene oxide (EPA, 2016)**

p. 4-109 Toxicological Review of the Inhalation Carcinogenicity of Ethylene Oxide. (EPA, 2016) https://cfpub.epa.gov/ncea/iris/iris_documents/documents/toxreviews/1025tr.pdf

IRIS is also Presenting Arrays of Candidate Toxicity Values



Benzo[a]pyrene (EPA, 2017)

Improvements in Characterizing Uncertainty

1) Model Averaging: characterizing model uncertainty

- Currently evaluating several methods
- Approach for dichotomous data expected to undergo peer review in 2018

$$\Pr(BMD \mid D) = \sum_{i=1}^{9} \pi_i \Pr(BMD \mid M_i, D)$$

Posterior Distribution of the BMD

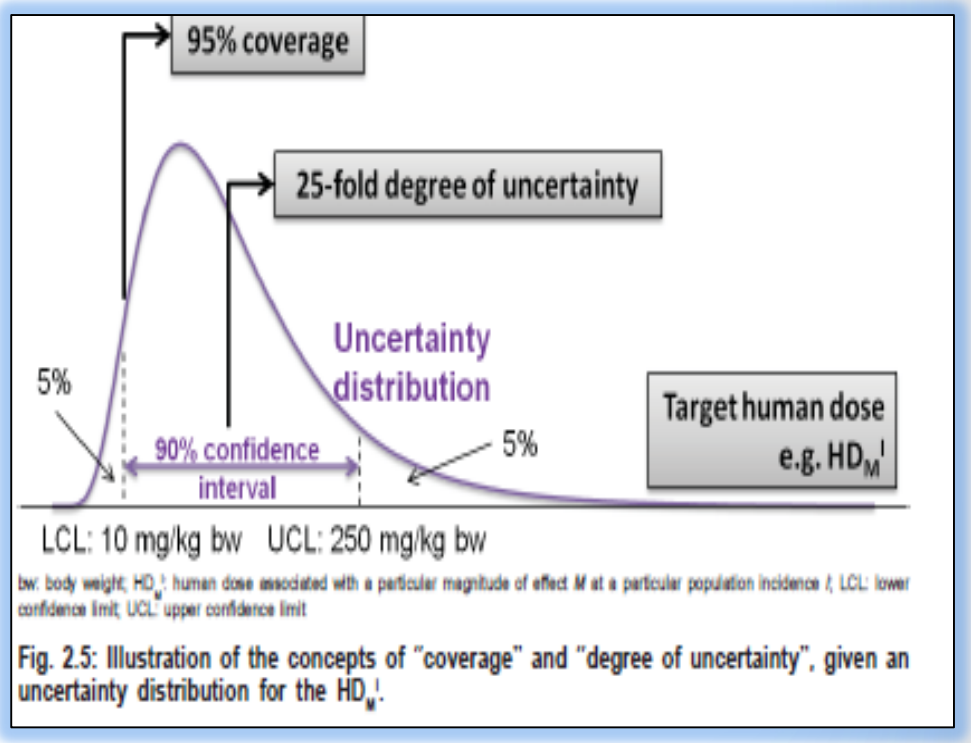
$$\alpha = \int_{-\infty}^{BMD_\alpha} \Pr(BMD \mid D) dBMD$$

Calculation of the BMDL

Improvements in Characterizing Uncertainty

2) Distributions and Central Estimates: characterizing uncertainty in the human toxicity value

- WHO/IPCS guidance (IPCS, 2014)
- Risk-specific doses in terms of ranges, for explicitly described:
  - Effect magnitudes
  - Confidence levels
  - Human population incidence rates.
- A probabilistic approach to adjustments from animal to human; a framework for refining toxicity values.



bw: body weight; HD_M^I: human dose associated with a particular magnitude of effect *M* at a particular population incidence *I*; LCL: lower confidence limit; UCL: upper confidence limit

Fig. 2.5: Illustration of the concepts of "coverage" and "degree of uncertainty", given an uncertainty distribution for the HD_M^I.

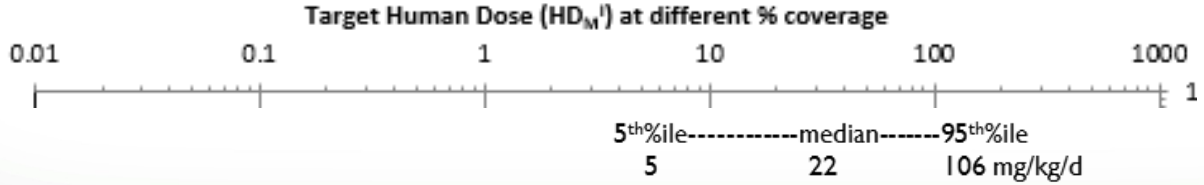
Improvements in Characterizing Uncertainty

WHO/IPCS Approach:

IRIS intends to provide such calculations along with traditional Reference Values:

- Confidence intervals on risk-specific doses
- Central estimates
- Estimates of incidence as a function of dose
- Use of appropriate probability math for uncertainty adjustments (instead of UFs) to allow for a more probabilistic and scientific value for use in risk assessment

By characterizing ranges of risk-specific doses, this provides more than a "conservative" estimate (it provides useful context by estimating the full distribution)



Use of Quantitative Modeling to Inform Evidence Integration

Meta-Analysis:

Increasingly Being Used to Interpret Sets of Results across Similar Populations

- Formal tools continue to be used to combine similar human epidemiology studies to improve decisions about hazard and about slope of dose-response.
- These approaches have also been used to better understand animal data that differ between studies of similar species and endpoints.
- As software tools and best practices become more common and easier to apply to environmental health studies, IRIS intends to consider their use more routinely.

Other examples: Libby Amphibole Asbestos (2014) and Trimethylbenzene analysis (Davis and Kraft, 2017) – see poster session; Arsenic assessment (in process)

Use of Quantitative Modeling to Inform Evidence Integration

Bayesian Approaches:

More Frequent Use Across Different Applications, and Research is Ongoing

- **Characterizing Uncertainty**
  - Bayesian approaches were used to characterize uncertainty in PBPK modeling and evaluate inter-related model inputs (Perchlorate peer review, 2018).
  - Bayesian Analysis is compatible with the WHO/IPCS Approach for characterizing uncertainty
- **Model Averaging**
  - Bayesian approaches are being applied to individual BMD models, and then model averaging is used to characterize uncertainty
- **Meta-Analysis**
  - Bayesian meta-analysis is currently being used to evaluate arsenic epidemiology studies
- **Bayesian Networks** (**exploratory research is currently underway**)
  - Possess the potential to integrate across evidence streams and bridge data gaps, borrowing strength from diverse data.
  - Software and mathematics are currently available.

Future work to better meet Agency needs for "benefits analysis"

Economics benefits analysis would ideally estimate incidence resulting from different decision options.

- We have provided human dose response functions from some analyses based on epidemiology data. (Evaluation of the Inhalation Carcinogenicity of Ethylene Oxide, EPA, 2016).

IRIS is also evaluating analogous predictions from animal data that could inform benefits analysis, including modifications of the IPCS approach.

Advancing Application of New Approach Methods (NAM) and Data in HHRA

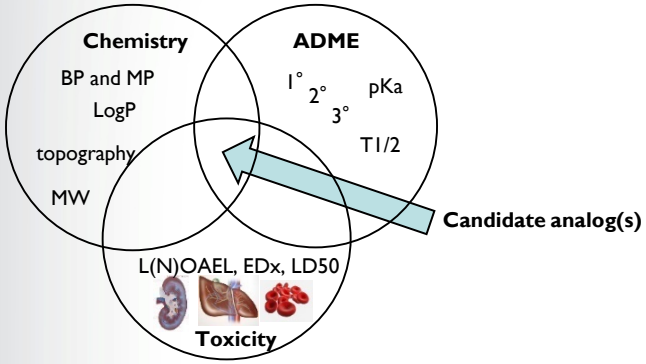
- Over the past decade, several reports, books, resource documents, etc. have been published regarding the use of New Approach Methods (NAM) across the human health risk assessment paradigm (i.e., shifting the paradigm)
- Numerous labs, centers, workgroups, and initiatives across federal, private, and academic institutions have been formed to advance NAM

- EPA/ORD/NCEA, in conjunction with partners (e.g., NCCT, NTP) has been actively engaged in the conceptualization and evaluation of NAM across a broad landscape of HHRA applications

NAM Toolbox to Date

- **Data-mining**: ToxRefDB-comprehensive collection and collation of extant hazard and exposure data –(Martin et al. 2009. Env Health Perspect 117: 392-399)
- **Chemoinformatics**: structure-activity/read-across; QSAR –(Wang et al. 2012. Regul Toxicol Pharmacol 63: 10-19; Craig et al. 2014. J Appl Toxicol 34: 787-794)
- **High-Throughput (HT) Exposure modeling**: ExpoCast –(Egeghy et al. 2016. Env Health Perspect. 124(6):697-702)
- **HT Toxicokinetics**: in vitro to in vivo (IVIVE) modeled dosimetry –(Wambaugh et al. 2015. Tox Sci 147: 55-67)
- **Bioactivity**: short-term animal; cell-free and/or cell-based HT assay data –(Judson et al. 2011. Chem Res Toxicol 24: 451-462; Dean et al. 2017. Tox Sci 157(1):85-99)
- **Adverse Outcome Pathway (AOP)**: expert-driven identification of signal transduction pathways along the exposure to outcome continuum. –(Edwards et al. 2016. J Pharmacol Exp Ther. 356(1):170-181)

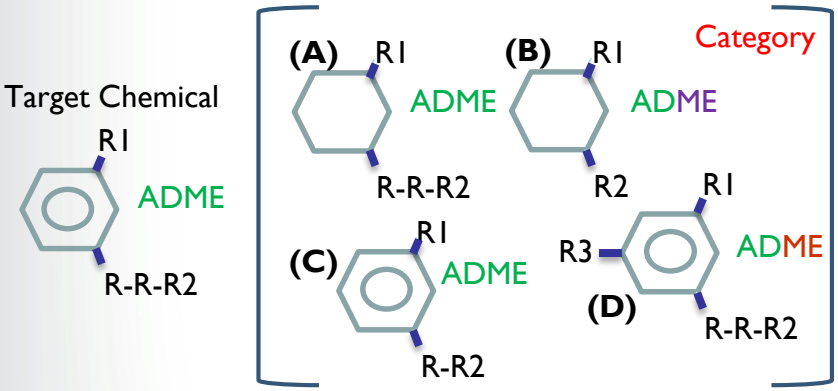
Chemoinformatics



Expert-driven Read-Across

Data-poor chemicals
- Inferred/interpolated hazard
- Surrogate based POD and subsequent derivation of RfVs

IRIS-type chemicals
- Data-gap filling
- Augment WOE
- Potential for reducing uncertainties



ADME = Absorption, Distribution, Metabolism, Elimination

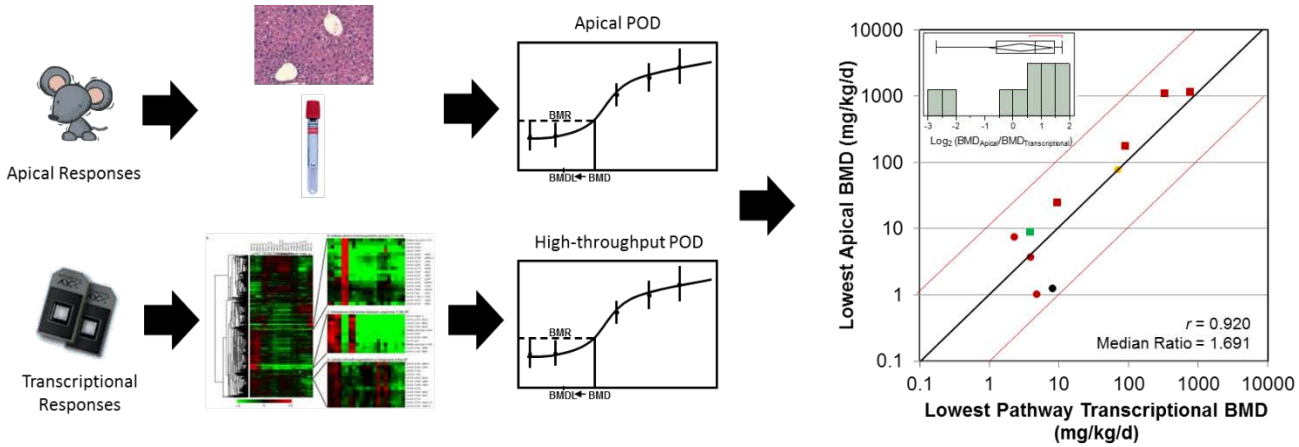
Category approach

Data-poor chemicals
- Data-gap filling
- Extrapolated hazard
- Less applicable for quantitative assessment currently

IRIS-type chemicals
- Data-gap filling
- Augment WOE
- Foundational member of category (i.e., anchor chem)

- Similarity in structure and physicochemical properties between a chem of concern and a population of analogs
- Robustness of approach dependent on density of analogs populating a category
- Highly reliant on WOE supporting toxicity endpoints across category
- Presumes common Adverse Outcome Pathway or Mode of Action across category members

Bioactivity (e.g., transcriptomics)



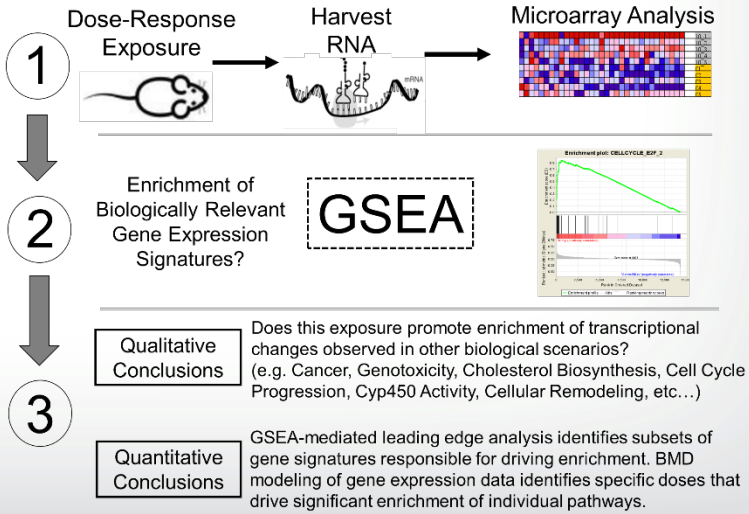
- Close relationship between genotype/phenotype across two different routes of exposure, rodent species, and multiple target tissues
- In vitro?? Will need to optimize metabolism protocols; integrate IVIVE

<u>Data-poor chemicals</u>

- Evidence base for hazard
- Empirical dose-response based on pathway perturbations
- Reduce need for longer-term animal studies

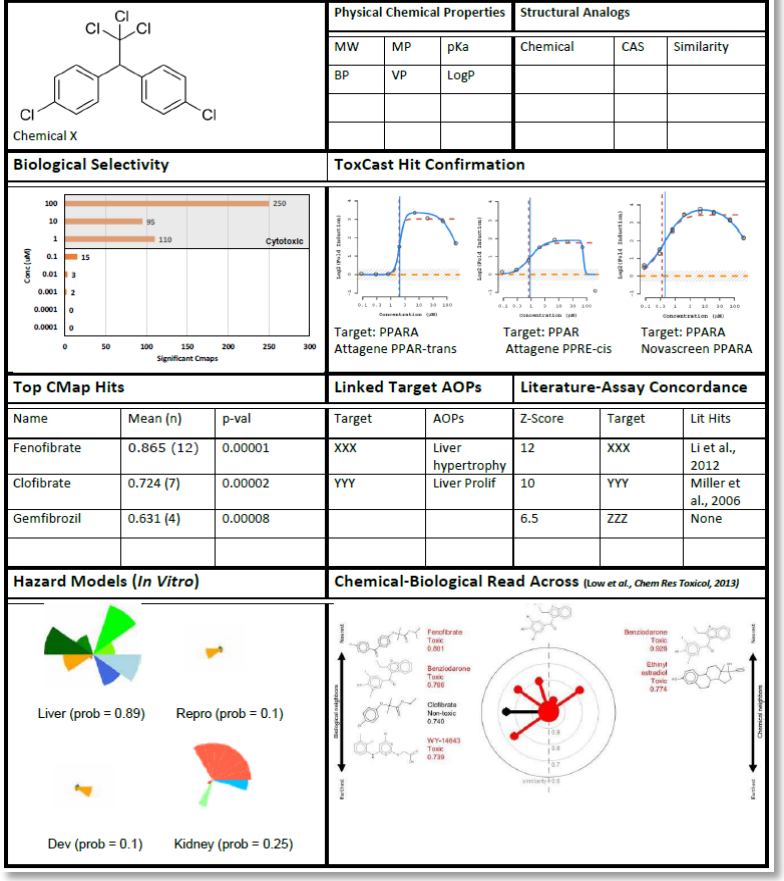
<u>IRIS-type chemicals</u>

- Augment WOE (e.g., MOA/AOP)
- Opportunity to alert off-target effects
- Potential for reducing uncertainties



Integrated Application to Risk Assessment

RapidTox Dashboard



Chemical X

Physical Chemical Properties

MW	MP	pKa
BP	VP	LogP

Structural Analogs

Chemical	CAS	Similarity

Biological Selectivity

ToxCast Hit Confirmation

Target: PPARA Attagene PPAR-trans

Target: PPAR Attagene PPRE-cis

Target: PPARA Novascreen PPARA

Top CMap Hits

Name	Mean (n)	p-val
Fenofibrate	0.865 (12)	0.00001
Clofibrate	0.724 (7)	0.00002
Gemfibrozil	0.631 (4)	0.00008

Linked Target AOPs

Target	AOPs	Z-Score
XXX	Liver hypertrophy	12
YYY	Liver Prolif	10
		6.5

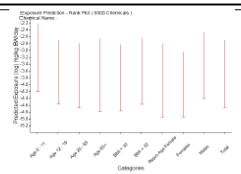
Literature-Assay Concordance

Target	Lit Hits
XXX	Li et al., 2012
YYY	Miller et al., 2006
ZZZ	None

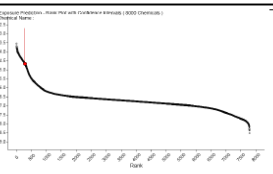
Hazard Models (*In Vitro*)

Chemical-Biological Read Across (Low et al., Chem Res Toxicol, 2013)

Exposure Prediction



Exposure Rank



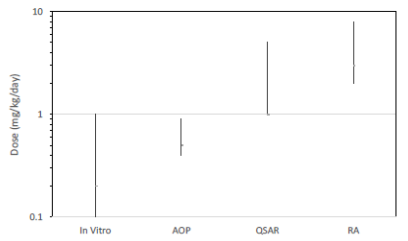
Literature Summary

- Chemical X induces peroxisome proliferation in primary rat hepatocytes (Johns et al., 2003)
- Rats treated with Chemical X showed increased hypertrophy in the liver (Applehans et al, 1998)

Point of Departure Estimate

Method	POD (mg/kg/d)
In Vitro Assay	0.1
In Vitro Assay (AOP-derived)	0.5
QSAR	0.4
Read Across	1.0

Forrest Plot



Assessment Summary

	Value	Confidence	UFs	RfD
Chemical Selectivity:	Moderate	Moderate		
Likely Hazards:	Liver toxicity	High		
Likely AOP/MOA:	PPARA receptor activation causing hepatocyte prolif	High		
Point-of-Departure Estimate	0.5 mg/kg/d		X-X-X-X	10 mg/kg/d
Margin-of-Exposure Estimate				25,000

Comments

- **Associated narrative can be modular based on fit-for-purpose**
- **Systematic WOE always, but can be graded based on decision context**
- **Characterization of qualitative and quantitative uncertainties**

IRIS has Addressed the Major NAS 2014 Recommendations

NAS 2014 Topics	IRIS Process Improvements
Evidence Integration for Hazard Identification (Chapter 6) and Derivation of Toxicity Values (Chapter 7)	• Developing and applying quantitative tools in support of evidence synthesis and integration, including meta-analytical approaches • Expanded development and use of more advanced quantitative methods in software tools, such as BMDS • Developed more explicit criteria for deriving toxicity values, including the intent to derive quantitative toxicity values when IRIS reaches one of the stronger evidence integration conclusions, as well more specific criteria for the evaluation of individual studies • Providing MLE estimates of BMDs, along with BMDLs • Applying and exploring quantitative approaches to better characterize uncertainty, including probabilistic and Bayesian approaches
Future Directions (Chapter 8 "Lessons Learned" and "Looking Forward")	• Quantitative assessment methods will be updated in a continuing, strategic fashion, including capacity building (e.g., training; evolving best practices) for current approaches including meta-analysis, probabilistic analyses, and Bayesian methods

EPA
United States
Environmental Protection
Agency

COLLABORATION, TRAINING, AND FINAL THOUGHTS

Tina Bahadori* and Kris Thayer

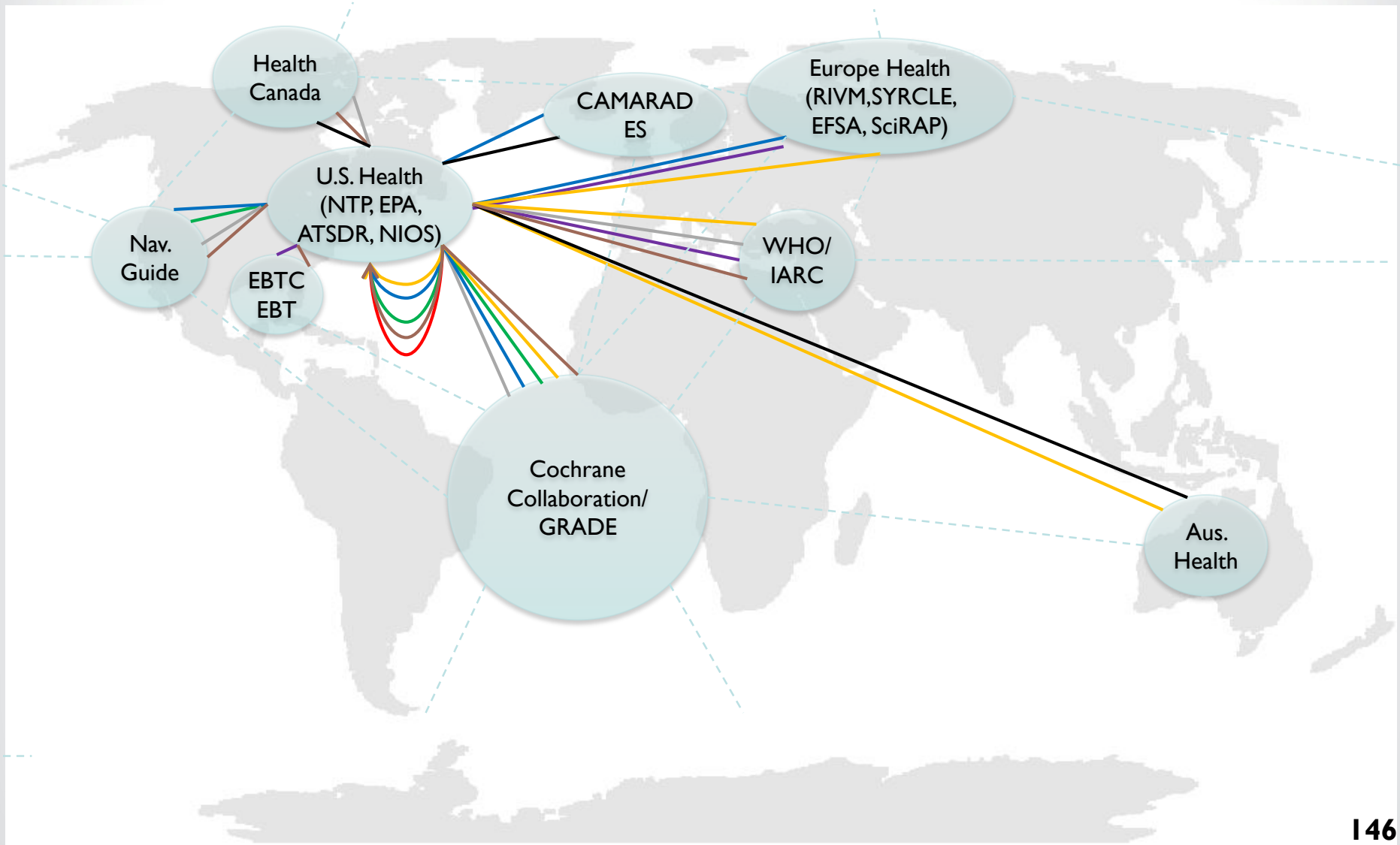
[*Speaking]

Office of Research and Development
NCEA, IRIS

Training and Collaboration

- Held multiple training sessions for IRIS Program staff in 2017, ranging from demos, seminars, to retreats. More to come in 2018…
- Developed support teams to provide teaching and assistance for systematic review tasks and use of new software ("train the trainer" model)
- Active engagement in the EPA Systematic Review Communities of Practice
- Engagement with external stakeholders, other Agency offices, state and other Agencies on systematic review methods and software training
  - e.g., MOUs with NTP, NIOSH, ATSDR, WHO
  - Interagency funding agreement with NIEHS/NTP for text-mining and software tool development and evaluation
- Establishing several academic MOUs to promote hands on training on use of systematic review in chemical assessments

International Collaborations



IRIS has Addressed the Major NAS 2014 Recommendations

NAS 2014 Topics	IRIS Process Improvements
General Process Issues (Chapter 2)	• Quality management pipeline implemented • Program and project management processes implemented • Frequent opportunities for stakeholder engagement • Draft IRIS Handbook of program SOPs is being reviewed within EPA • Re-occurring staff training and template IAPs and protocols promote consistency and quality control
Problem Formulation and Protocol Development (Chapter 3)	• IAPs allow early comment on problem formulation • More frequent Agency engagement facilitates scope refinement • Assessment protocols describe methods and allow for iteration
Evidence Identification (Chapter 4)	• Consultation with information technologists and subject experts • Adopts current systematic review best practices, including use of specialized tools • Transparent documentation (e.g., literature flow diagrams)

IRIS has Addressed the Major NAS 2014 Recommendations

NAS 2014 Topics	IRIS Process Improvements
Evidence Evaluation (Chapter 5)	• Individual studies are evaluated for reporting quality, risk of bias, and sensitivity • Decisions and supporting rationale are clearly documented • Study evaluations impact subsequent assessment decisions
Evidence Integration for Hazard Identification (Chapter 6)	• Structured frameworks provide transparency in expert judgments across human, animal, and mechanistic studies (based on Hill) • Standardized templates documenting key evidence integration decisions have been developed (evidence profile tables) • Developing and applying quantitative tools in support of evidence synthesis and integration, including meta-analytical approaches • Expanded development and use of more advanced quantitative methods in software tools, such as BMDS

IRIS has Addressed the Major NAS 2014 Recommendations

NAS 2014 Topics	IRIS Process Improvements
Derivation of Toxicity Values (Chapter 7)	• Developed more explicit criteria for deriving toxicity values, including the intent to derive quantitative toxicity values when IRIS reaches one of the stronger evidence integration conclusions, as well more specific criteria for the evaluation of individual studies • Providing MLE estimates of BMDs, along with BMDLs • Applying and exploring quantitative approaches to better characterize uncertainty, including probabilistic and Bayesian approaches

IRIS has Addressed the Major NAS 2014 Recommendations

NAS 2014 Topics	IRIS Process Improvements
Future Directions (Chapter 8 "Lessons Learned" and "Looking Forward")	• Processes being implemented include flexibility to incorporate evolving methods in systematic review and risk assessment • Increased collaboration with federal partners and international experts prevents duplication of effort and maintains cutting edge approaches • Current research efforts and training serve to ensure that methods and staff are able to adapt to changing scientific contexts and sources of evidence, including new and emerging data types • Specialized software tools for efficiency and more user friendly and transparent formats for evidence display have been adopted • Strategic planning on use of text and data-mining tools and automation • Specialized tools facilitate transparent documentation, consistency across assessments, and database interoperability • Quantitative assessment methods will be updated in a continuing, strategic fashion, including capacity building (e.g., training; evolving best practices) for current approaches including meta-analysis, probabilistic analyses, and Bayesian methods