NTP
National Toxicology Program

Welcome to the NTP Webinar

Today's Webinar: Lessons Learned in Application of the OHAT Framework for Systematic Review and Evidence Integration to Case Studies

Audio Options

1) **Listening only:** available via this webcast or

2) **Audio with ability to ask questions during Q & A**

United States:	1 (800) 894-5910 or 1 (785) 424-1052	
Belgium:	08-007-2519	or 1(785) 424-1052
Brazil:	0-800-891-6744	or 1(785) 424-1052
Germany:	0-800-186-2030	or 1(785) 424-1052
United Kingdom:	0-808-101-1183	or 1(785) 424-1052
Other international callers:		or 1(785) 424-1052

NTP
National Toxicology Program

Lessons Learned in Application of the OHAT Framework for Systematic Review and Evidence Integration to Case Studies

Office of Health Assessment and Translation
National Institute of Environmental Health Sciences

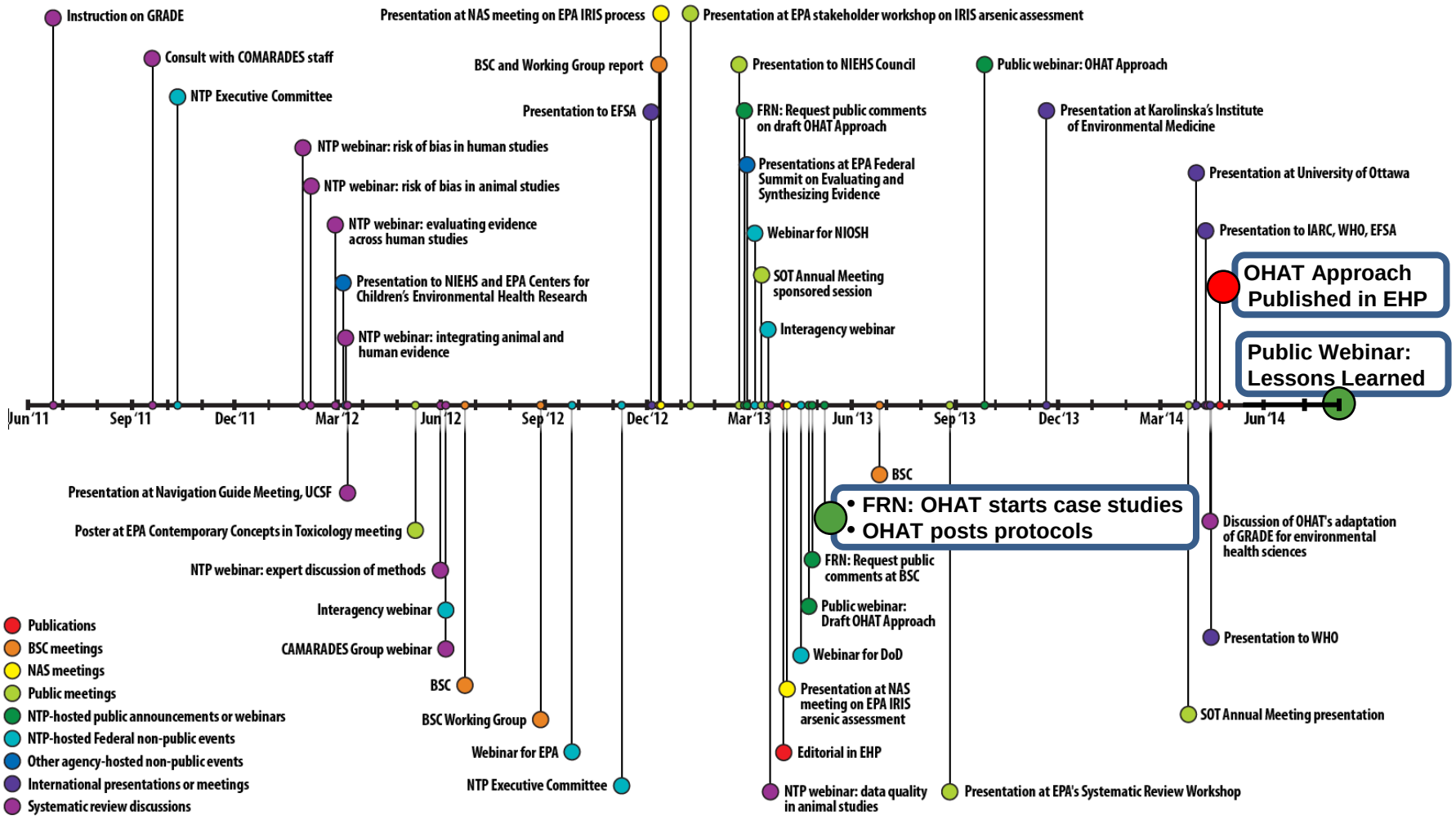
Andrew A. Rooney

July 31, 2014

Format and Logistics

- OHAT staff presentation on a “lesson” or topic
- Question and answer session on that topic
  - After presentation, participants can indicate to the operator if you would like to ask a question
  - Participants will be called upon in the order questions are received and the phone line will be unmuted
  - Participants can ask their question directly
- Lesson topics and timing
  - 5 topics
  - Remaining time (~60 minutes) for additional discussion

Development of OHAT Framework for Systematic Review and Evidence Integration



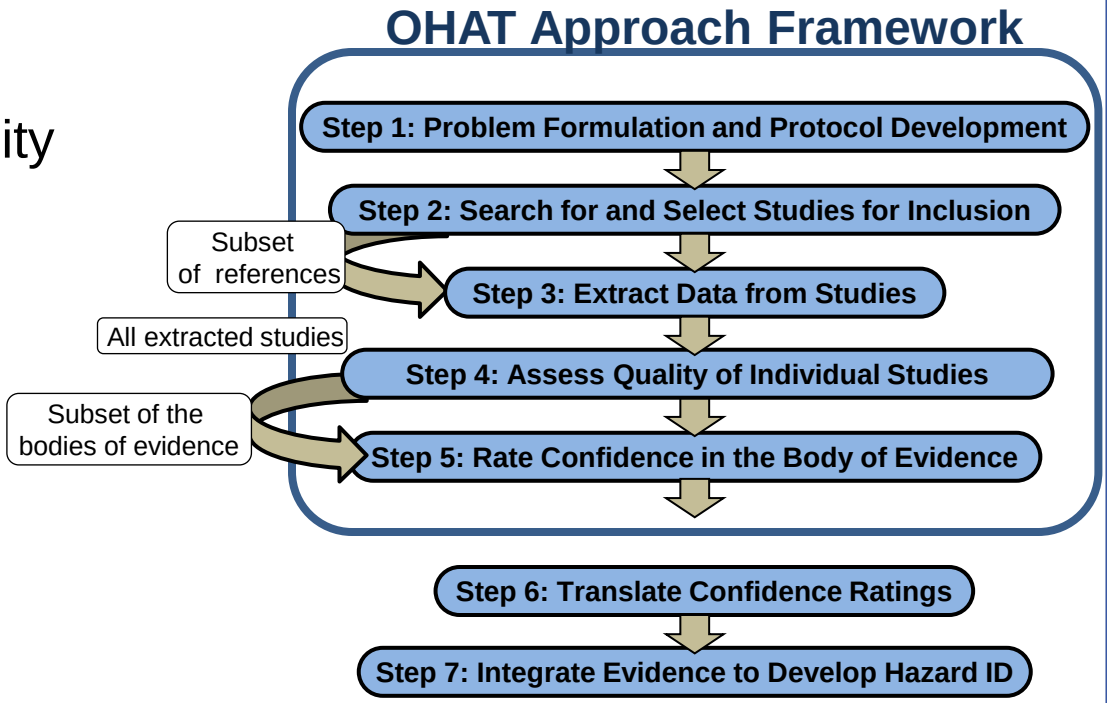
- Case studies to assist with determining if refinement or revision might be needed to Draft OHAT Approach – February 2013

Role of Case Studies in OHAT Implementation of Systematic Review

- Developed to explore methods not to reach conclusions
  - Work flow/oversight, new software tools
  - Feasibility of utilizing “gold standard” systematic review practices
  - Extent to which guidance could be “template” vs “case-specific”

- Two case studies
  - BPA exposure and obesity
  - PFOA/PFOS exposure and immunotoxicity

- Other OHAT projects



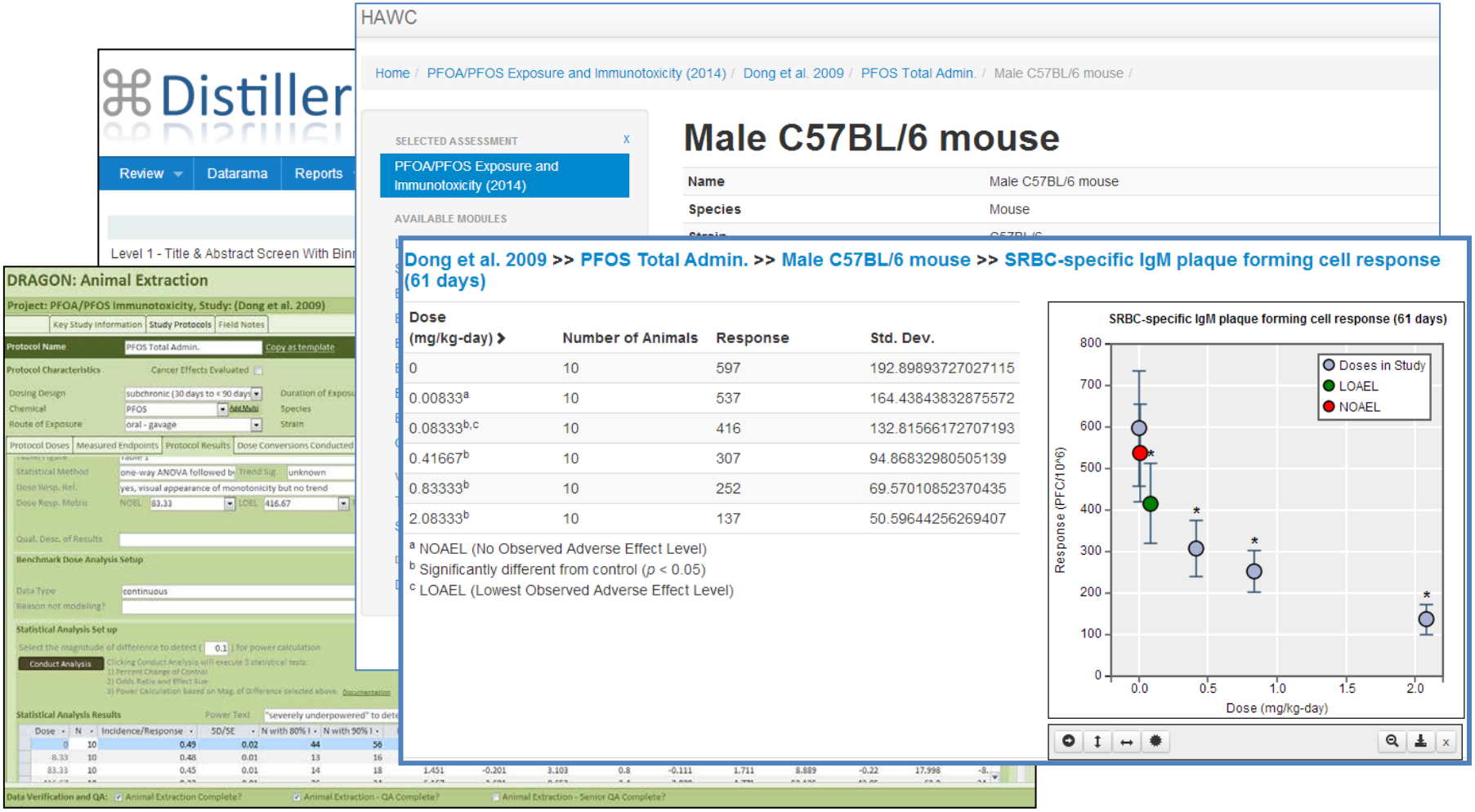
Lessons Learned on Framework, Workflow, etc.

- The framework for systematic review and evidence integration accommodates assessment–specific tailoring
  - Framework published (Rooney et al., 2014)
- Need for “handbook” with instructions for protocol development
  - Incorporates input from public comments, experience with case studies, recommendations from NAS reports, and discussion with groups using systematic review
  - Will be posted on NTP website
- Focus on further development and refinement of software tools
  - Distiller, DRAGON, HAWC, text-mining, etc.
- Principal “lessons learned” focus of today’s webinar

Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA. 2014. Systematic Review and Evidence Integration for Literature-Based Environmental Health Science Assessments. Environ Health Perspect 122(7): 711-718.

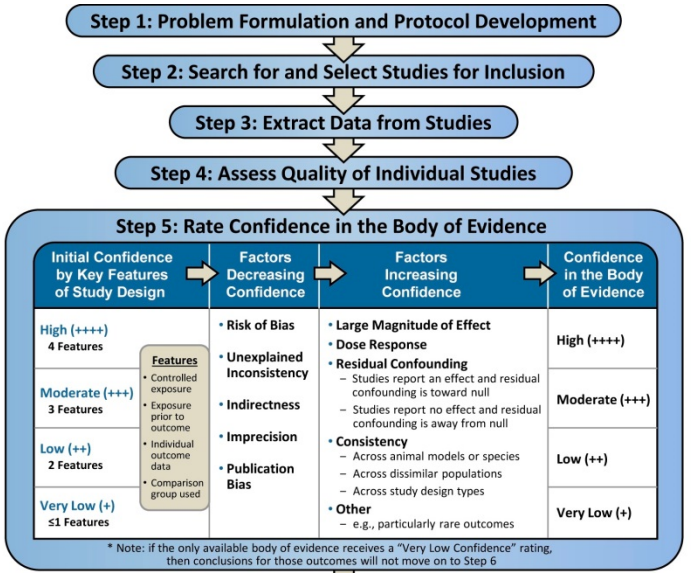
Refinement of Software Tools

- Emerging tools facilitate analysis and display
- Experience leads to efficiencies, but data entry still = **TIME**



Overarching Lessons Learned

- **Scoping and Problem Formulation:** can be time consuming but are critical
- **Searching the Literature:** finding a balance between practical and comprehensive is challenging
- **Piloting:** multiple steps benefit from pilot-testing procedures on a small group of studies and refining the protocol as necessary
- **Assessing Study Quality:** detailed guidance and documentation aid in transparency in applying risk of bias
- **Rating Confidence in the Body of Evidence:** structured summary text as well as graphical aids are helpful for reaching and communicating confidence ratings

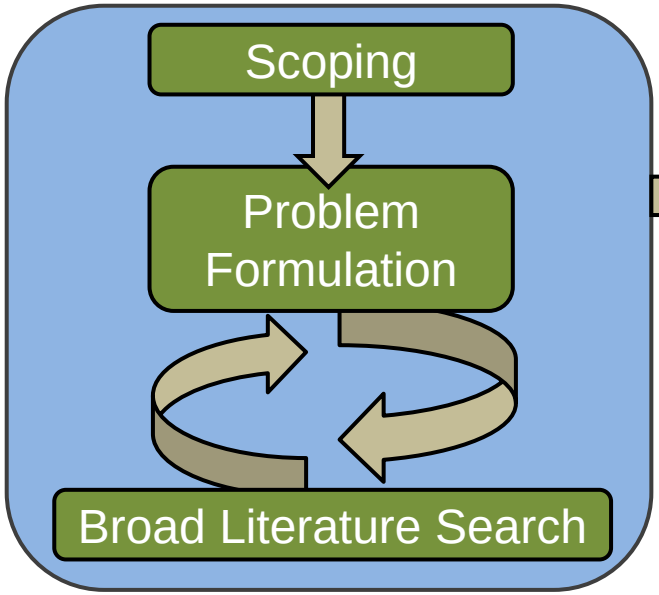


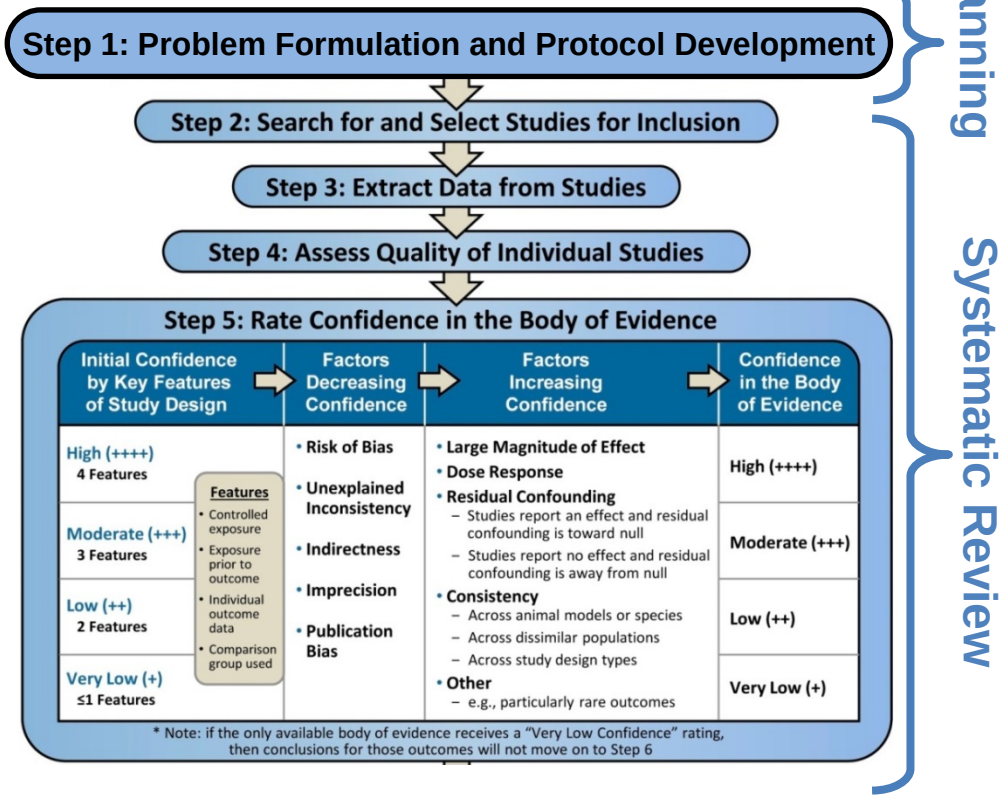
Scoping and Problem Formulation

Lessons Learned on Problem Formulation

- Case studies were focused from outset, so were not major challenges for problem formulation
- Other OHAT projects were greater challenges for problem formulation (e.g., transgenerational inheritance of health effects)
  - Problem formulation is more important when question involves a broad range of health outcomes and large number of search results
  - When possible, we will utilize text-mining and expert-opinion methods to save staff time and resources for systematic review
- Problem formulation of mechanistic evidence is challenging
  - How wide to cast the net?
  - How to approach *in vitro* studies when mechanistic basis for health outcome is unclear?

Scoping and Problem Formulation

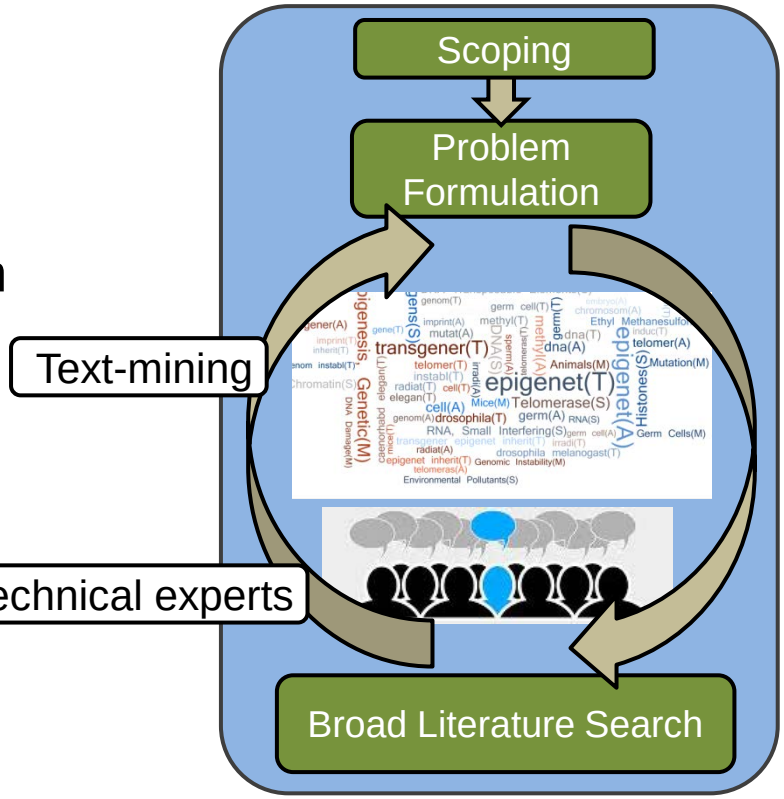




- Recent OHAT experience working with nominations support greater emphasis on scoping and problem formulation
- NRC recommended both as part of the planning phase before starting a systematic review

Scoping and Problem Formulation

- NTP receives nominations
- OHAT Scoping
  - Inform NTP agency points of contact and solicit public input
    - Names of topic experts, unpublished or ongoing research
    - General comments, suggested focus areas
  - Form evaluation design team
- OHAT Problem Formulation
  - Conduct initial broad literature search
  - Identify possible health outcomes
  - Use information to develop the specific research question and determine feasibility



Summary of Lessons Learned:
Problem Formulation

- Thoughtful problem formulation can make the systematic review more efficient
  - Text-mining and expert-opinion methods can assist in refining the question and clarifying the scope
- More work is needed to set the boundaries of the literature search for *in vitro* studies or mechanistic data
  - Much more difficult to define the scope for mechanistic studies than for animal or human studies
- **Questions?**

Searching the Literature

Lessons Learned on Searching the Literature

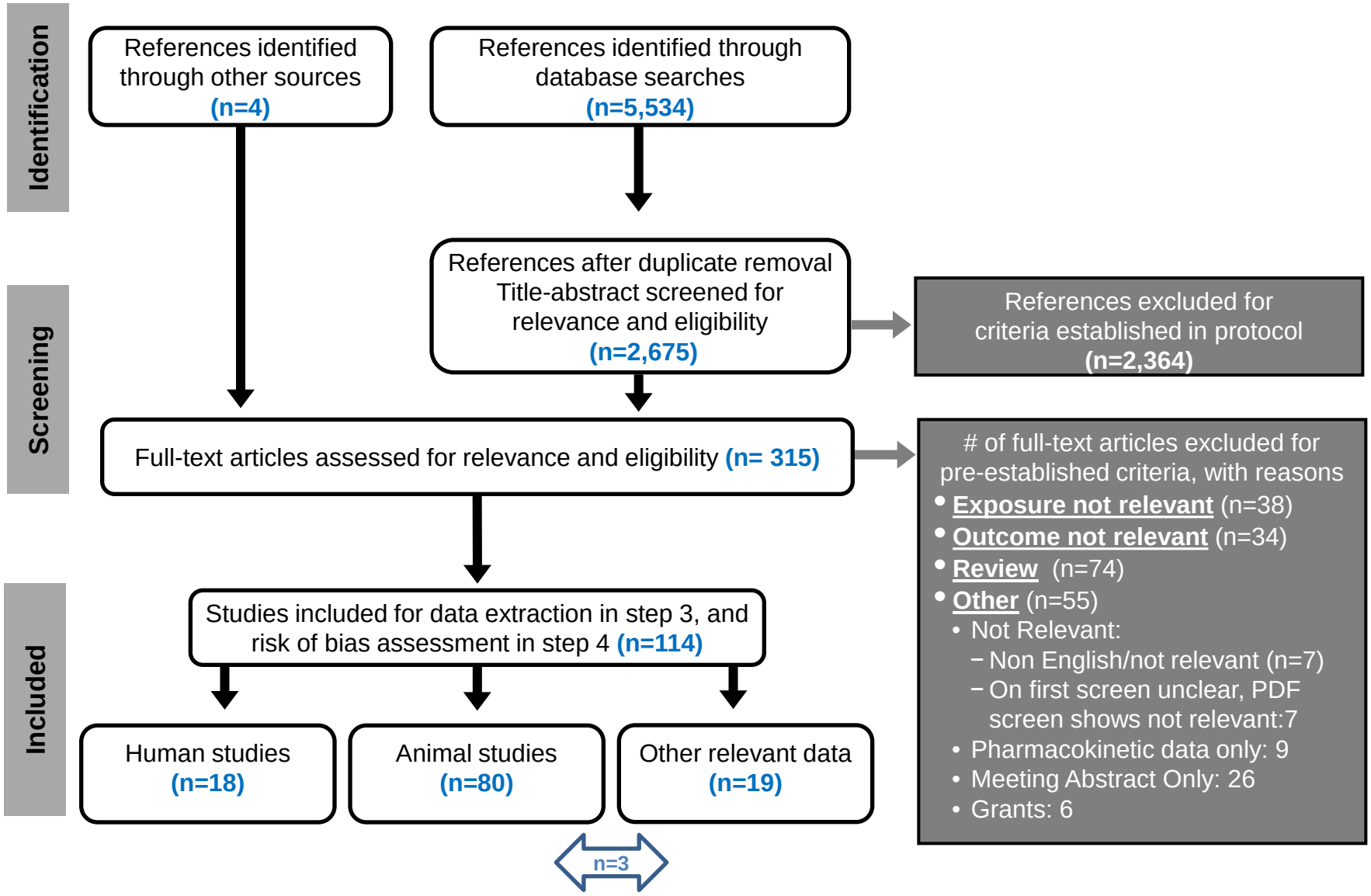
- OHAT's initial literature search approach may be too broad
  - "Gold standard" approach may not be practical for all evaluations
  - Including certain types of studies can be very resource intensive
    - Non-English publications
    - Conference abstracts
- Project-specific decisions need to be made on the scope and extent of the literature search
- OHAT employs additional methods to supplement the literature search
  - Consult technical experts, examine relevant reviews, post literature search results, solicit peer review and public comment

From Literature Search to Screening:
Eligibility Criteria for PFOA/PFOS and Immunotoxicity

- Objective
  - Evaluate the evidence that exposure to PFOA or PFOS is associated with changes in immune-related measures in humans, animals, or *in vitro* model systems.

- Protocol states specific eligibility criteria
  - Types of studies (e.g., no restrictions based on study design)
  - Types of human studies and models (e.g., include wildlife studies)
  - Types of exposures (e.g., based on administered dose)
    - Perfluorooctanoic acid (PFOA)
    - Perfluorooctane sulfonate (PFOS)
  - Types of outcomes (e.g., disease resistance assay)
  - Types of publications (e.g., no restriction on language)

Screening: Example PFOA/PFOS and Immunotoxicity



PDF Retrieval and Full-text Screening Were the Largest Screening Investment for Case Studies

- Title and abstract screening
  - Majority of eligibility decisions can be made on title and abstract
  - Absence of an abstract is a problem

- PDF retrieval and full text screening
  - PDF retrieval required for relevant and for "unclear" references
  - Retrieval of some references is highly time intensive
    - Journals not in NIH holdings
    - Conference or meeting publications
    - Theses, grants
    - Foreign language publications

Summary of Lessons Learned:
Lessons from Searching for and Selecting Studies

- "Gold standard" literature search approach may not be practical for evaluations with larger literature bases
  - Case-by-case determinations on eligibility criteria, e.g., whether or not an evaluation will retrieve non-English studies
- Public release and outreach to supplement literature search
  - OHAT outreach includes soliciting input on list of references
- **Questions?**

Pilot-testing

Lessons Learned on the Value of Pilot-testing Procedures

- Pilot-testing of procedures on a small group of studies improved consistency and reduced the need for discussion, conflict resolution, and error correction

- Pilot-testing and refining the protocol are important at multiple steps
  - Inclusion and exclusion criteria
  - Data extraction guidance
  - Risk of bias assessment

Applying Inclusion and Exclusion Criteria
Example: Piloting and Refining PFOA Exposure Criteria

- Inclusion Criteria
  - **Exposure:**
    - Studies must include exposure to PFOA or PFOS based on:
      - Administered dose or concentration,
      - Biomonitoring data (e.g., urine, blood, or other specimens),
      - Environmental measures (e.g., air, water levels), or
      - Indirect measures such as job title
    - PFOA (perfluorooctanoic acid) synonyms
      - C8
      - Eftop EF-201
      - n-Perfluorooctanoic acid
      - Octanoic acid, 2,2,3,3,4,4,5,5,6,6,7,7,8,8,8-pentadecafluoro-
      - Octanoic acid, pentadecafluoro- [free acid; CAS # 335-67-1]
      - Pentadecafluorooctanoic acid
      - Pentadecafluoro-1-octanoic acid
      - *(abbreviated list of 20+ synonyms for slide)*
- Exclusion Criteria
  - **Exposure:**
    - NA

Initially there was no list of excluded exposures

Pilot-testing the Screening Tool
Addition of Excluded Chemicals Increased Efficiency

Adding “excluded” chemicals was an important refinement to the key-word highlighting

Exposure

- Included chemical
- Unclear, possible chemical
- Excluded chemical

Title-Abstract Screening Question

Does the title or abstract suggest this article contains data related to PFOA or PFOS and Immunotoxicity?

- ○ YES, hand collected
- ○ YES, identified from literature search
- ○ Relevant review or commentary, hand collected
- ○ Relevant review or commentary, from literature search
- ◉ NO, not relevant
- ○ UNCLEAR, need PDF

Reference: in Title-Abstract View

Refid: 14
(Pyridoxalated hemoglobin)-(polyoxyethylene) conjugate solution as blood substitute for normothermic whole body rinse-out T.Agishi, Y. Funakoshi, H.Honda, K.Yamagata, M.Kobayashi and M.Takahashi

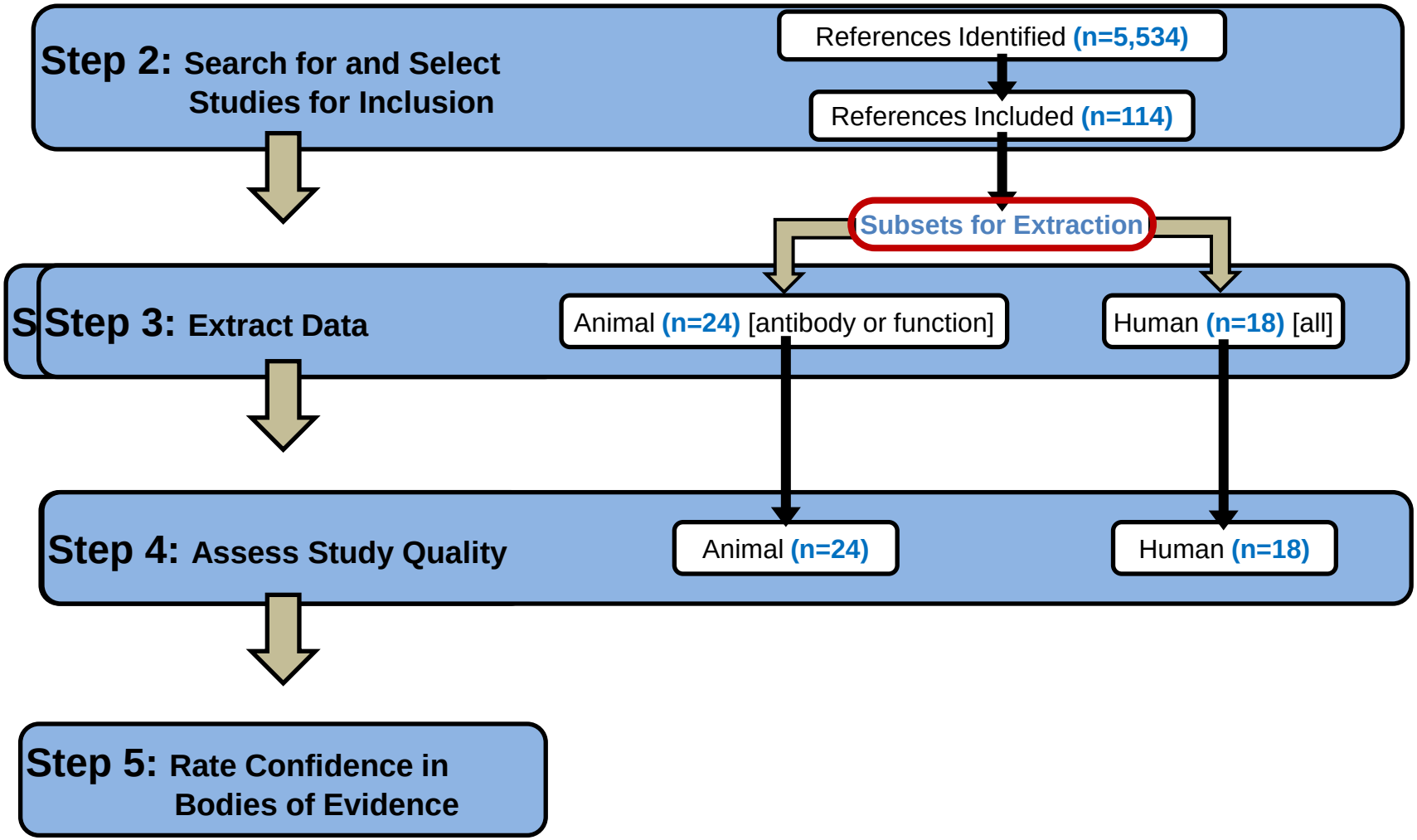
In order to investigate a new possibility for artificial blood with oxygen-carrying capability to be applied to other than mere supplementation, normothermic whole body rinse-out in which artificial blood deriving from **perfluorochemical** emulsion, **Fluosol**-DA 20% (Green Cross Co., Ltd., Osaka, Japan) or stabilized hemoglobin solution, (pyridoxalated hemoglobin)-(polyoxyethylene) conjugate solution (Ajinomoto Co., Ltd., Tokyo, Japan) were used as rinsing fluid for a blood purification experiment. Replacement either with approximately 150 ml/kg of **Fluosol**-DA or stabilized hemoglobin solution showed effective removal of digoxin at a reduction rate of 96.3% or 92.2%, respectively. However, when **Fluosol**-DA was used, a certain amount of **perfluorochemical** should be retrieved by centrifugation to avoid a possible toxic effect on the reticulo-endothelial system. Even though 3 out of 6, and 3 out of 8 dogs, respectively, survived for a long period after the procedure, the experimental dogs were very susceptible to **infection**.

Summary of Lessons Learned:
Lessons on the Value of Pilot-testing Procedures

- Pilot testing procedures and refining the protocol are important at multiple steps
  - Screening
    - Increases efficiency
    - Decreases conflicts between reviewers
    - Reduced workload and cost of retrieving “unclear” references
  - Data Extraction
    - Improves consistency
    - Reduces errors and need for corrections
  - Risk of bias assessment
    - Decreases need for discussion and resolution between assessors
- **Questions?**

Assessing Study Quality

Study Flow and Subsets Used for Case Studies: Example PFOA/PFOS and Immunotoxicity



Study Quality and Utility

- While not an issue for the case study process, we found there is frequently confusion on what is meant by quality
- Multiple aspects of “quality” and “utility” are important for an evaluation
  - **Reporting quality**
    How well was the study reported?
  - **Risk of bias or internal validity**
    How credible are the findings based on design and conduct of the study?
  - **Directness and applicability**
    How well does the study address the topic under review?

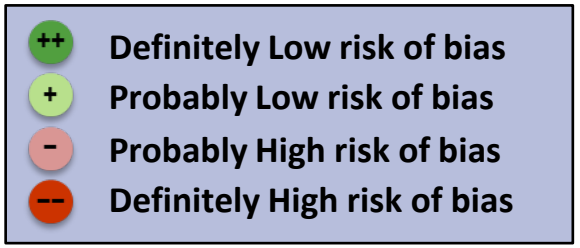
Lessons Learned Assessing Study Quality

- Detailed risk of bias guidance is necessary
  - Specific and consistent guidance could be developed, even for challenging topics
  - Expert input may be required on project specifics (e.g., exposure)

- Pilot-testing and adjusting guidance (up front where possible) minimized discussion needed to reach final risk of bias ratings and supported consistency
  - Pilot-testing the guidance on several studies and adjusting for areas that are unclear
  - Documenting the basis of risk of bias ratings
  - Modifications to the guidance for unforeseen situations

Risk of Bias Assessment

- Evaluation is endpoint/outcome specific
- Answers on 4-point scale from Clarity Group
  - Study design determines which questions are applicable
  - Answers equate to risk of bias rating for each question/criteria



- Risk of bias relies on
  - Detailed guidance for all risk of bias questions
  - Specificity on defining the evidence from a study report that determines the risk of bias rating
  - Risk of bias assessed independently in duplicate
  - Discussion to reach and document basis for risk of bias rating for each question
  - Adjustments to the guidance

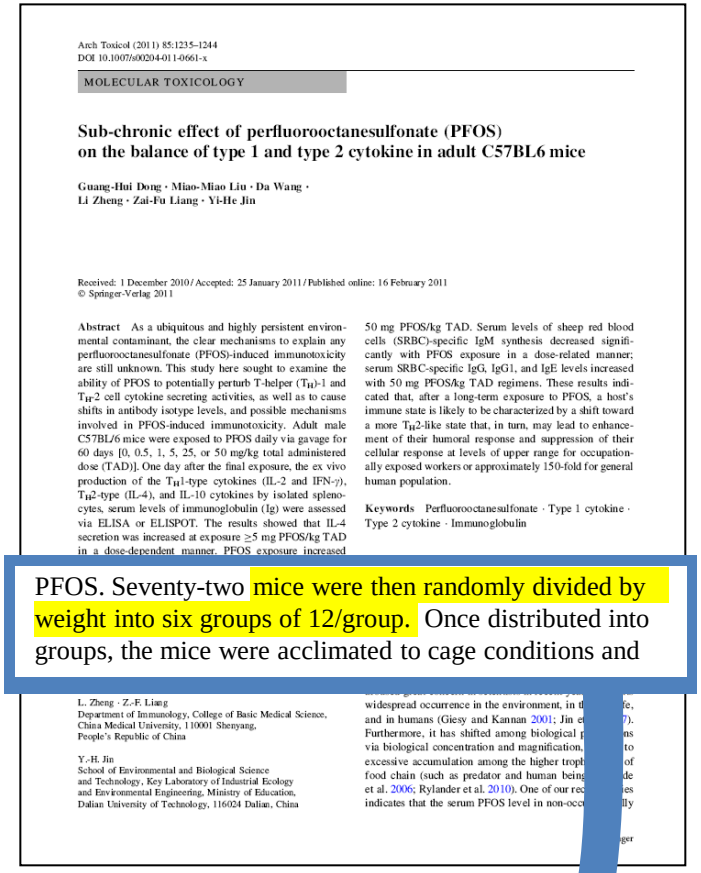
Each Risk of Bias Question Is Answered with Evidence from Study Report or Author Contact

Specific Guidance

Guidance defines all 4 ratings for each question

1. Randomization

- ++ **Definitely Low risk of bias:** There is direct evidence that animals were allocated to any study group including controls using a method with a random component. Restricted randomization (e.g., blocked) will be considered low bias …
- \+ **Probably Low risk of bias:** There is indirect evidence that animals were allocated to any study group including controls using a method with a random component (i.e., authors state that allocation was random …
- – **Probably High risk of bias:** ….
- – – **Definitely High risk of bias:** …



Arch Toxicol (2011) 85:1235–1244
DOI 10.1007/s00204-011-0661-x

MOLECULAR TOXICOLOGY

Sub-chronic effect of perfluorooctanesulfonate (PFOS) on the balance of type 1 and type 2 cytokine in adult C57BL6 mice

Guang-Hui Dong · Miao-Miao Liu · Da Wang · Li Zheng · Zai-Fu Liang · Yi-He Jin

Received: 1 December 2010 / Accepted: 25 January 2011 / Published online: 16 February 2011
© Springer-Verlag 2011

Abstract As a ubiquitous and highly persistent environmental contaminant, the clear mechanisms to explain any perfluorooctanesulfonate (PFOS)-induced immunotoxicity are still unknown. This study here sought to examine the ability of PFOS to potentially perturb T-helper (T_H)-1 and T_H2 cell cytokine secreting activities, as well as to cause shifts in antibody isotype levels, and possible mechanisms involved in PFOS-induced immunotoxicity. Adult male C57BL/6 mice were exposed to PFOS daily via gavage for 60 days [0, 0.5, 1, 5, 25, or 50 mg/kg total administered dose (TAD)]. One day after the final exposure, the ex vivo production of the T_H1-type cytokines (IL-2 and IFN-γ), T_H2-type (IL-4), and IL-10 cytokines by isolated splenocytes, serum levels of immunoglobulin (Ig) were assessed via ELISA or ELISPOT. The results showed that IL-4 secretion was increased at exposure ≥5 mg PFOS/kg TAD in a dose-dependent manner. PFOS exposure increased 50 mg PFOS/kg TAD. Serum levels of sheep red blood cells (SRBC)-specific IgM synthesis decreased significantly with PFOS exposure in a dose-related manner; serum SRBC-specific IgG, IgG1, and IgE levels increased with 50 mg PFOS/kg TAD regimens. These results indicated that, after a long-term exposure to PFOS, a host's immune state is likely to be characterized by a shift toward a more T_H2-like state that, in turn, may lead to enhancement of their humoral response and suppression of their cellular response at levels of upper range for occupationally exposed workers or approximately 150-fold for general human population.

Keywords Perfluorooctanesulfonate · Type 1 cytokine · Type 2 cytokine · Immunoglobulin

PFOS. Seventy-two mice were then randomly divided by weight into six groups of 12/group. Once distributed into groups, the mice were acclimated to cage conditions and

L. Zheng · Z.-F. Liang
Department of Immunology, College of Basic Medical Science, China Medical University, 110001 Shenyang, People's Republic of China

Y.-H. Jin
School of Environmental and Biological Science and Technology, Key Laboratory of Industrial Ecology and Environmental Engineering, Ministry of Education, Dalian University of Technology, 116024 Dalian, China

widespread occurrence in the environment, in t… fe, and in humans (Giesy and Kannan 2001; Jin e… 7). Furthermore, it has shifted among biological p… ns via biological concentration and magnification, … to excessive accumulation among the higher troph… of food chain (such as predator and human being… de et al. 2006; Rylander et al. 2010). One of our rec… es indicates that the serum PFOS level in non-occ… lly

Was administered dose or exposure adequately randomized? | ++

Support for final rating: *"mice were randomly divided by weight"*

Clear Risk of Bias Guidance Can Be Developed… Even for Challenging Issues

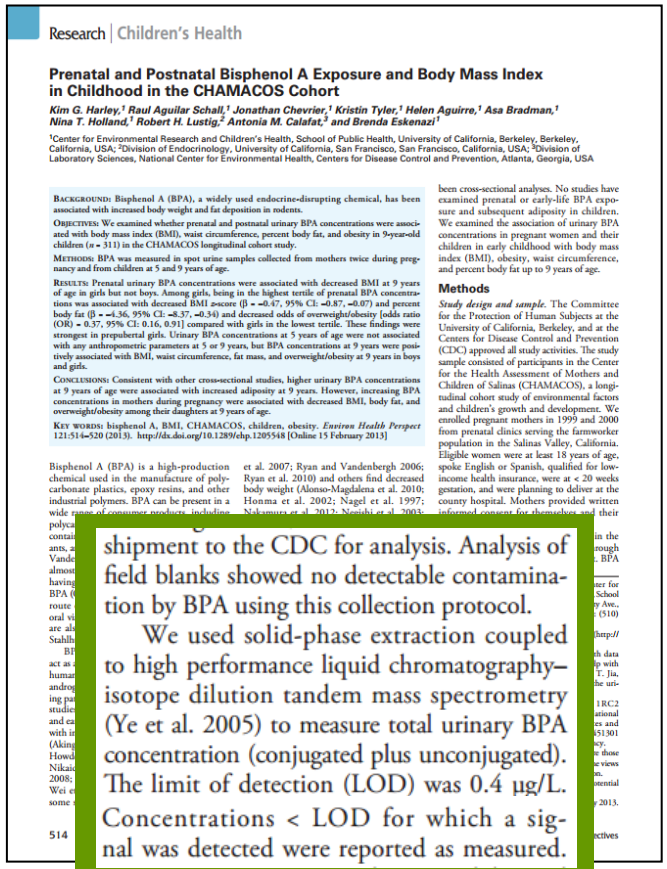
Example: BPA Exposure

Specific Guidance

Guidance defines all 4 ratings for each question

12. Exposure

- ++ **Definitely Low risk of bias:** There is direct evidence that most data points for the aglycone, conjugated and/or total BPA are ***above*** the level of quantitation (LOQ) for the assay; **AND** the study utilized spiked samples to confirm assay performance and the stability of BPA and conjugated BPA in biological samples was appropriately addressed …
- \+ **Probably Low risk of bias:** There is indirect evidence that most data points for the aglycone, conjugated and/or total BPA are ***above*** the level of quantitation (LOQ) …
- – **Probably High risk of bias:** ….
- – – **Definitely High risk of bias:** …



Research | Children's Health

Prenatal and Postnatal Bisphenol A Exposure and Body Mass Index in Childhood in the CHAMACOS Cohort

Kim G. Harley,[1] Raul Aguilar Schall,[1] Jonathan Chevrier,[1] Kristin Tyler,[1] Helen Aguirre,[1] Asa Bradman,[1] Nina T. Holland,[1] Robert H. Lustig,[2] Antonia M. Calafat,[3] and Brenda Eskenazi[1]

[1]Center for Environmental Research and Children's Health, School of Public Health, University of California, Berkeley, Berkeley, California, USA; [2]Division of Endocrinology, University of California, San Francisco, San Francisco, California, USA; [3]Division of Laboratory Sciences, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Background: Bisphenol A (BPA), a widely used endocrine-disrupting chemical, has been associated with increased body weight and fat deposition in rodents.

Objectives: We examined whether prenatal and postnatal urinary BPA concentrations were associated with body mass index (BMI), waist circumference, percent body fat, and obesity in 9-year-old children ($n = 311$) in the CHAMACOS longitudinal cohort study.

Methods: BPA was measured in spot urine samples collected from mothers twice during pregnancy and from children at 5 and 9 years of age.

Results: Prenatal urinary BPA concentrations were associated with decreased BMI at 9 years of age in girls but not boys. Among girls, being in the highest tertile of prenatal BPA concentrations was associated with decreased BMI z-score (β = –0.67, 95% CI: –0.87, –0.07) and percent body fat (β = –4.36, 95% CI: –8.37, –0.34) and decreased odds of overweight/obesity [odds ratio (OR) = 0.37, 95% CI: 0.16, 0.91] compared with girls in the lowest tertile. These findings were strongest in prepubertal girls. Urinary BPA concentrations at 5 years of age were not associated with any anthropometric parameters at 5 or 9 years, but BPA concentrations at 9 years were positively associated with BMI, waist circumference, fat mass, and overweight/obesity at 9 years in boys and girls.

Conclusions: Consistent with other cross-sectional studies, higher urinary BPA concentrations at 9 years of age were associated with increased adiposity at 9 years. However, increasing BPA concentrations in mothers during pregnancy were associated with decreased BMI, body fat, and overweight/obesity among their daughters at 9 years of age.

Key words: bisphenol A, BMI, CHAMACOS, children, obesity. *Environ Health Perspect* 121:514–520 (2013). http://dx.doi.org/10.1289/ehp.1205548 [Online 15 February 2013]

Bisphenol A (BPA) is a high-production chemical used in the manufacture of polycarbonate plastics, epoxy resins, and other industrial polymers. BPA can be present in a wide

et al. 2007; Ryan and Vandenbergh 2006; Ryan et al. 2010) and others find decreased body weight (Alonso-Magdalena et al. 2010; Honma et al. 2002; Nagel et al. 1997;

been cross-sectional analyses. No studies have examined prenatal or early-life BPA exposure and subsequent adiposity in children. We examined the association of urinary BPA concentrations in pregnant women and their children in early childhood with body mass index (BMI), obesity, waist circumference, and percent body fat up to 9 years of age.

Methods

Study design and sample. The Committee for the Protection of Human Subjects at the University of California, Berkeley, and at the Centers for Disease Control and Prevention (CDC) approved all study activities. The study sample consisted of participants in the Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS), a longitudinal cohort study of environmental factors and children's growth and development. We enrolled pregnant mothers in 1999 and 2000 from prenatal clinics serving the farmworker population in the Salinas Valley, California. Eligible women were at least 18 years of age, spoke English or Spanish, qualified for low-income health insurance, were at < 20 weeks gestation, and were planning to deliver at the county hospital. Mothers provided written

shipment to the CDC for analysis. Analysis of field blanks showed no detectable contamination by BPA using this collection protocol.

We used solid-phase extraction coupled to high performance liquid chromatography–isotope dilution tandem mass spectrometry (Ye et al. 2005) to measure total urinary BPA concentration (conjugated plus unconjugated). The limit of detection (LOD) was 0.4 µg/L. Concentrations < LOD for which a signal was detected were reported as measured.

514

Can we be confident in the exposure characterization?	++

Support for final rating: *analysis conducted at CDC NHANES using validated method, measured levels above LOD*

Risk of Bias

Example: PFOA/PFOS Animal Studies (subset assessed)

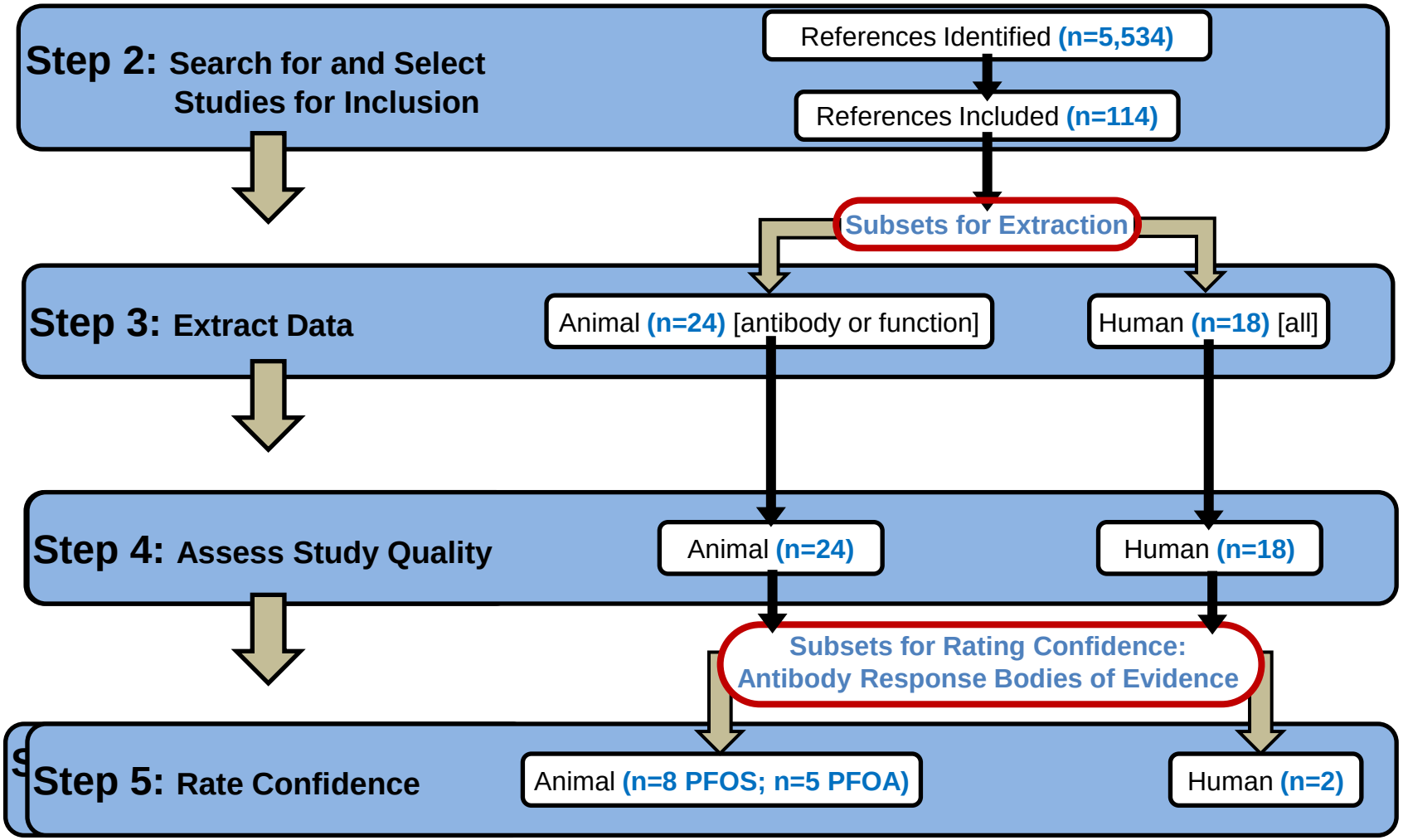
Risk of Bias Question	Dewitt 2008	Dewitt 2009	Dong 2009	Dong 2011	Fair 2011	Fairley 2007	Guruge 2009	Hu 2010	Hu 2012	Iwai 2006	Keil 2008	Lefebvre 2008	Loveless 2008	Peden-Adams 2008	Peden-Adams 2009	Qazi 2009	Qazi 2010	Son 2009	Yang 2000	Yang 2001	Yang 2002a	Yang 2002b	Zhang 2013	Zheng 2009	Zheng 2011
Randomization	+	−	++	++	−	++	+	+	++	−	−	−	+	+	+	−	−	+	−	−	+	+	++	++	++
Allocation Concealment	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
Confounding (design/analysis)	++	+	++	++	++	+	++	++	++	++	+	++	++	+	−	−	−	−	++	−	+	+	++	++	++
Unintended Exposure	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Identical Experimental Conditions	++	++	+	+	++	++	++	++	++	+	++	+	++	++	++	++	++	++	+	+	+	+	++	++	++
Adhere to Protocol	+	+	+	+	−	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Blinding of Researchers During Study	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Missing Outcome Data	−	+	++	++	−−	−	+	−	−	+	−−	−	−	+	++	+	++	+	−	−	+	−	++	++	++
Assessment of Confounding Variables	+	+	++	++	++	−	+	+	++	++	+	+	+	++	++	−	+	+	+	−	−	+	+	++	++
Exposure Characterization	++	−	+	+	−	−	−	+	−	−	−	+	+	+	+	+	+	−	−	−	−	−	+	+	+
Outcome Assessment	+	+	+	+	+	+	++	+	+	−	++	+	+	+	+	+	+	+	+	+	+	+	+	++	+
Blinding of Outcome Assessors	+	+	+	+	++	+	+	+	+	+	+	+	−−	+	++	+	+	+	+	+	+	+	+	+	+
Outcome Reporting	+	+	+	++	−−	+	+	+	+	−	+	+	−−	+	+	+	++	−	++	+	+	+	+	++	+

Summary of Lessons Learned:
Lessons Learned Assessing Study Quality

- Specific and consistent risk of bias (RoB) rules can be developed, even for challenging topics
- Detailed RoB guidance is necessary in the protocol
  - Expert input will be required
- Piloting and adjusting guidance minimized discussion needed to reach RoB ratings and supported consistency
  - Pilot-testing guidance and adjusting for areas that are unclear
  - Documenting the basis of risk of bias ratings
  - Further modifications to the guidance for unforeseen circumstances
- Other risk of bias tools are philosophically similar
  - We expect the specific tool will evolve over time
- **Questions?**

Rating Confidence in the Body of Evidence

Study Flow and Subsets Used for Case Studies: Example PFOA/PFOS and Immunotoxicity



Body of Evidence: Definitions

- **What is a confidence rating?**
  - A conclusion on a body of evidence that is developed by considering its strengths and weaknesses
  - Ratings reflect confidence that study findings reflect the true association between exposure and effect
- **What comprises a "body of evidence"?**
  - Studies with data on the same or related outcomes as defined in the protocol
- **What do we mean by "initial confidence"?**
  - The starting point for a study or group of studies prior to examining strengths and weaknesses

Lessons Learned Rating Confidence in the Body of Evidence

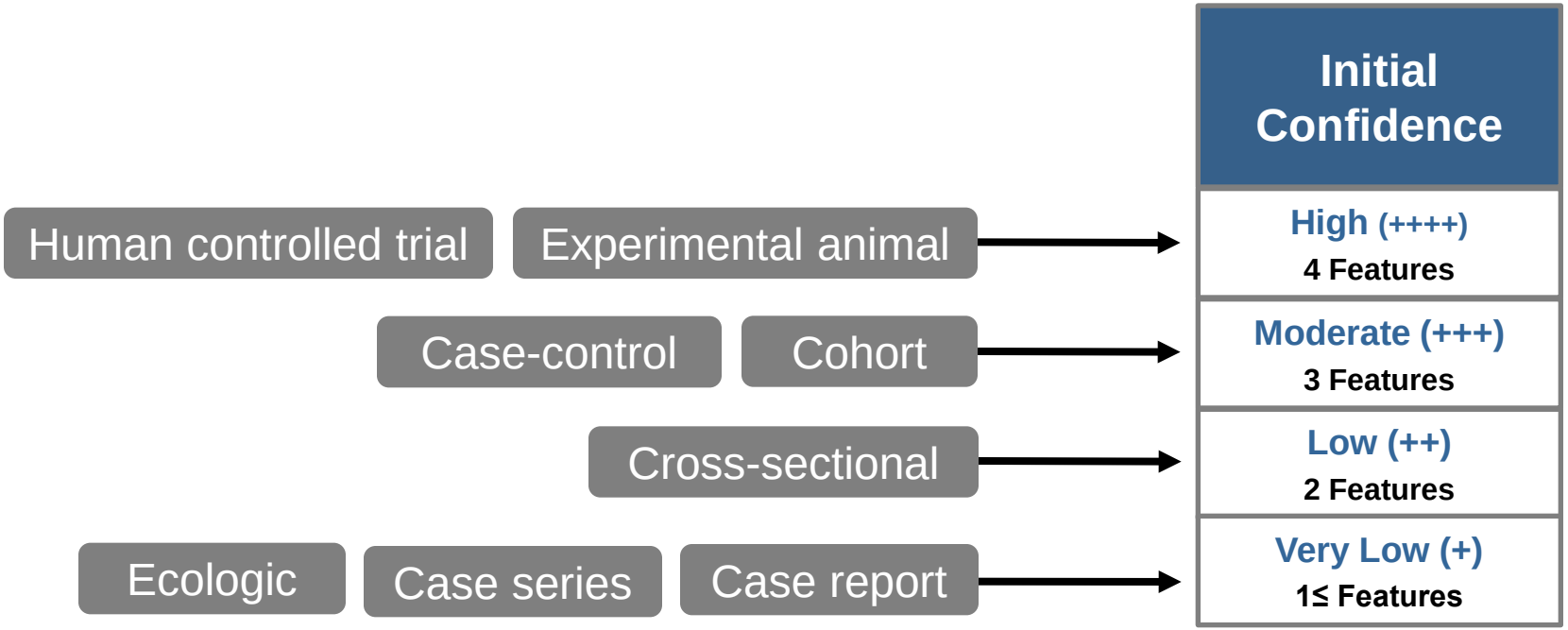
- Initial confidence rating by key study design features transparently grouped studies
- Summary text as well as graphical aids were helpful for reaching and communicating confidence ratings for bodies of evidence
- Publication bias was challenging to ascertain

Initial Confidence in a Body of Evidence

- Studies are stratified based on design features
- Stratification reflects importance of observational studies for environmental health assessments

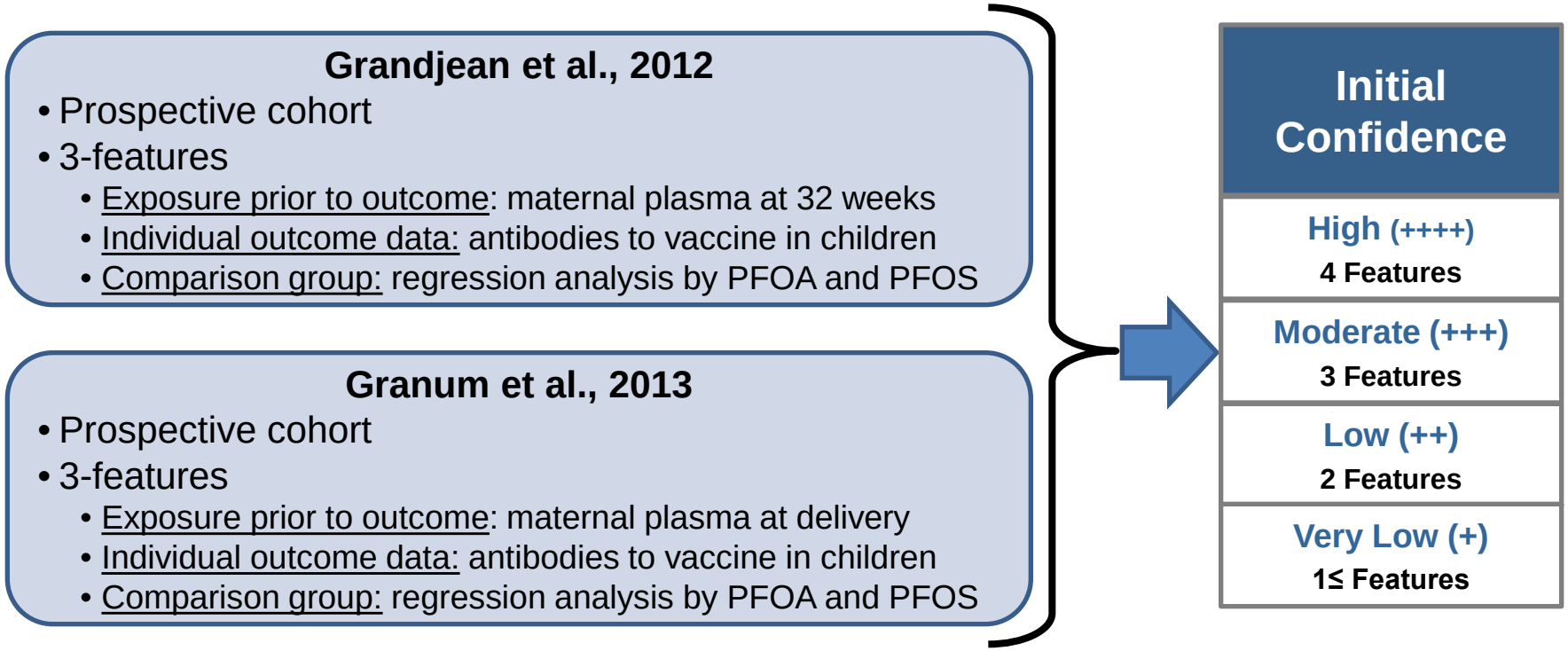
Features

- Controlled exposure
- Exposure prior to outcome
- Individual outcome data
- Comparison group used



Initial Confidence by Study Design Features

- **Example:** PFOA or PFOS exposure and immunotoxicity
  - Specific outcome = antibody response (e.g., antibodies to vaccines)



Initial Confidence by Study Design Features

- **Example:** PFOS exposure and immunotoxicity
  - Specific outcome = antibody response (e.g., to SRBC)

Dong et al., 2009
- PFOS Exposure-male mice
- 4-features
- Outcome at 60 days
  - ↓ IgM antibodies to SRBC-PFC

Dong et al., 2011
- PFOS Exposure-male mice
- 4-features
- Outcome at 60 days
  - ↓ IgM antibodies to SRBC-ELISA

Keil et al., 2008
- PFOS Exposure-developmentally to m/f mice
- 4-features
- Outcome at 28 days
  - ↓ IgM antibodies to SRBC-PFC

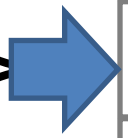
Peden-Adams et al., 2008
- PFOS Exposure-m/f mice
- 4-features
- Outcome at 28 days
  - ↓ IgM antibodies to SRBC-PFC
  - ↓ IgM antib.to TNP-LPS-ELISA

Lefebvre et al., 2008
- PFOS Exposure-m/f rats
- 4-features
- Outcome at 28 days
  - No effect or increase IgG antibodies to KLH by PFC

Qazi et al., 2010
- PFOS Exposure-male mice
- 4-features
- Outcome at 28 days
  - No effect on IgM antibodies to SRBC or TNP-LPS-PFC&ELISA

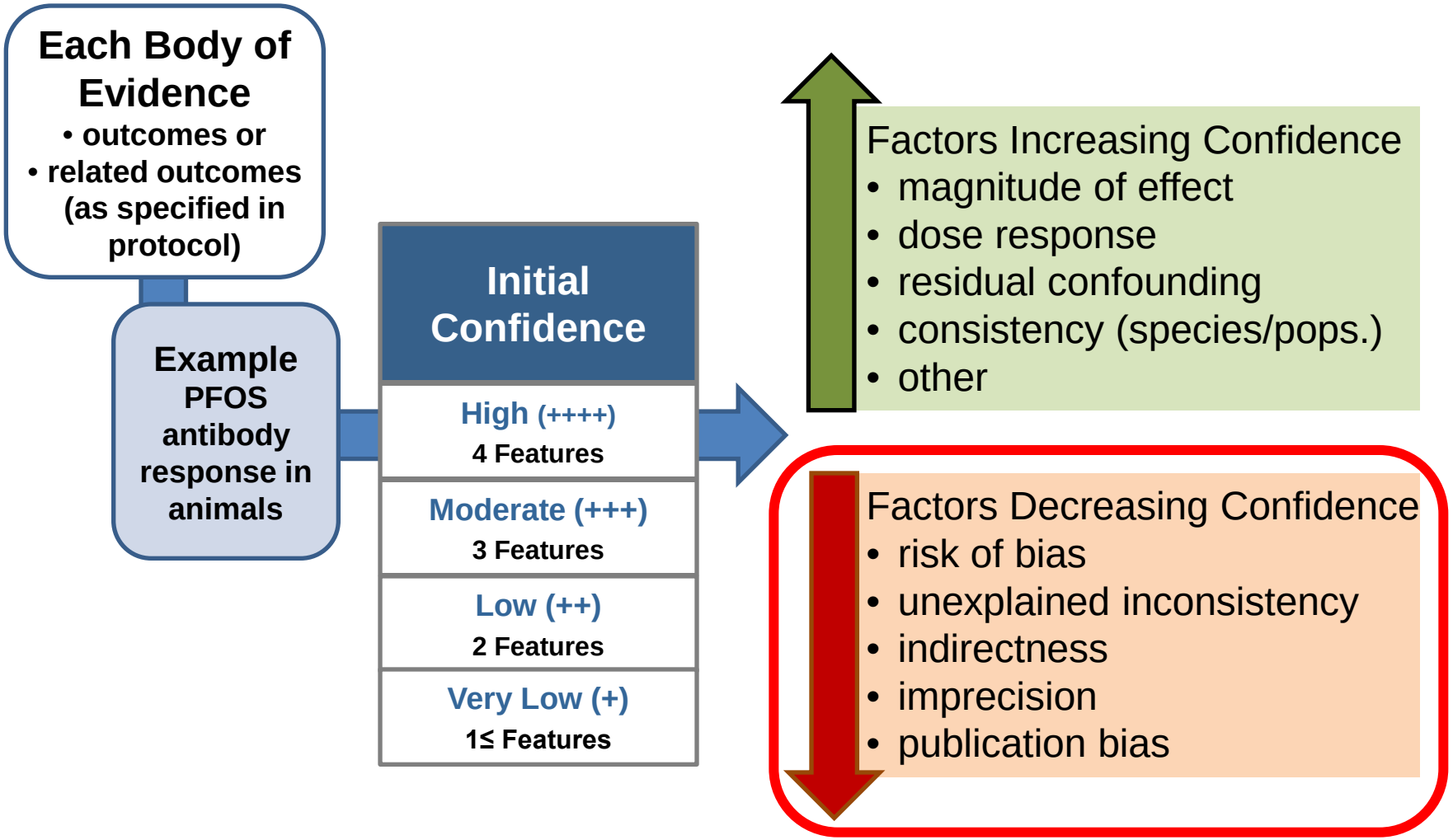
Zheng et al., 2009
- PFOS Exposure-male mice
- 4-features
- Outcome at 7 days
  - ↓ IgM antibodies to SRBC-PFC

Zheng et al., 2011
- PFOS Exposure-male mice
- 4-features
- Outcome at 7 days
  - ↓ IgM antibodies to SRBC-ELISA



Initial Confidence
High (++++) 4 Features
Moderate (+++) 3 Features
Low (++) 2 Features
Very Low (+) 1≤ Features

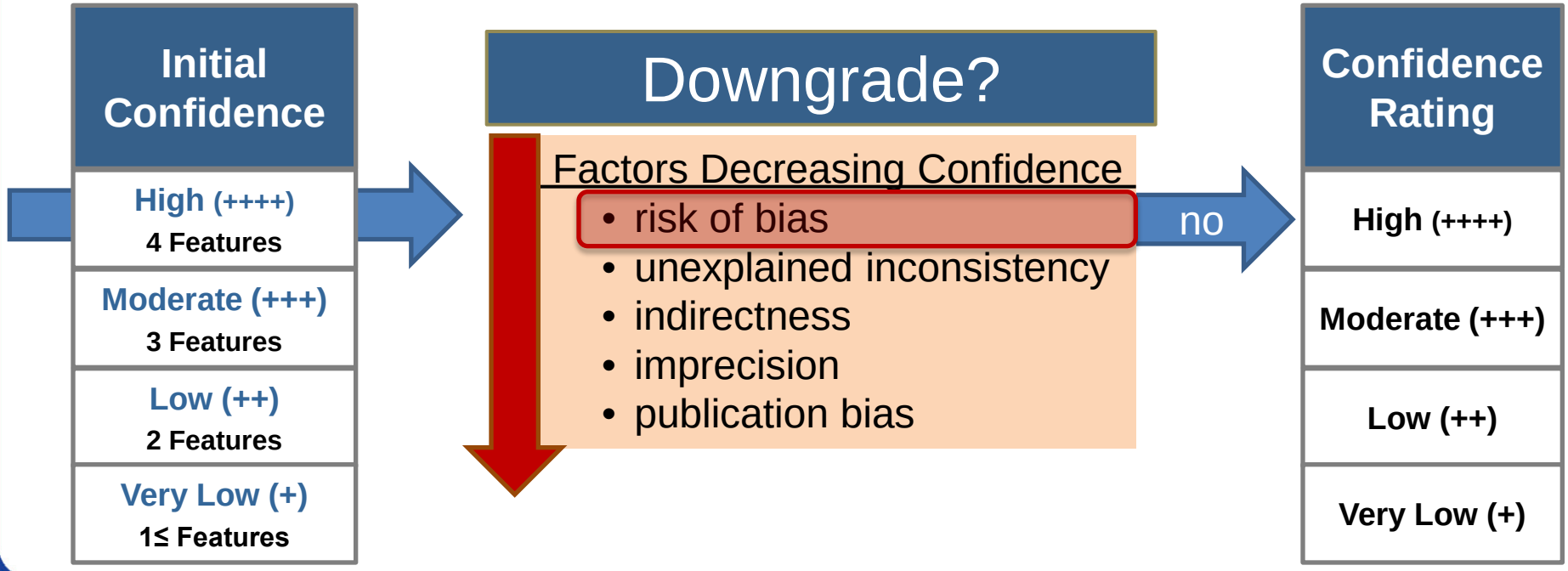
After Setting Initial Confidence, Ratings Are Developed by Considering Factors that Increase and Decrease Confidence



Downgrade Considerations: RoB

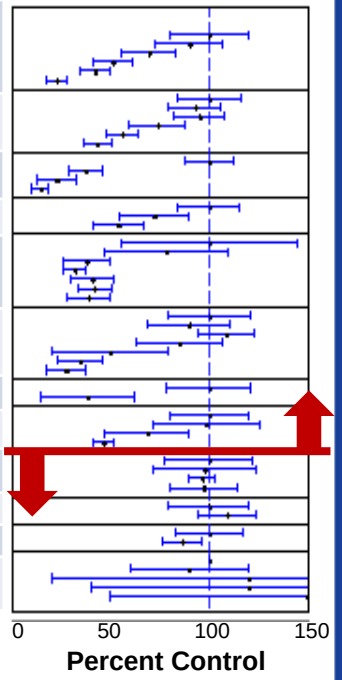
- Low risk of bias for almost all questions
- Key questions
  - –Randomization – mixed
  - –Outcome Assessment – no issues
- Probably high risk of bias for allocation concealment

Risk of Bias	Dong 2009	Dong 2011	Zheng 2009	Zheng 2011	Peden-Adams 2008	Keil 2008	Qazi 2010	Lefebvre 2008
Randomization	++	++	++	++	+	–	–	–
Allocation Concealment	–	–	–	–	–	–	–	–
Confounding (design/analysis)	++	++	++	++	+	+	–	++
Unintended Exposure	+	+	+	+	+	+	+	+
Identical Experimental Conditions	+	+	++	++	++	++	++	+
Adhere to Protocol	+	+	+	+	+	+	+	+
Blinding of Researchers During Study	+	+	+	+	+	+	+	+
Missing Outcome Data	++	++	++	++	+	––	++	–
Assessment of Confounding Variables	++	++	++	++	++	+	+	+
Exposure Characterization	+	+	+	+	+	–	+	+
Outcome Assessment	+	+	++	+	+	++	+	+
Blinding of Outcome Assessors	+	+	+	+	+	+	+	+
Outcome Reporting	+	++	++	+	+	+	++	+



Downgrade Considerations: Inconsistency

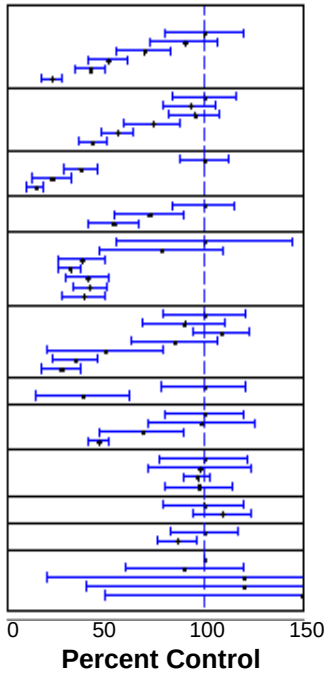
Reference	Species, Strain (sex)	Duration	Route	Max Dose/day	Risk of Bias: Randomization	Antibody Endpoint	Response
Dong 2009	Mice, c57BL/6 (male)	60 days	gavage	2.0833mg/kg	++	IgM to SRBC by PFC	
Dong 2011	Mice, c57BL/6 (male)	60 days	gavage	0.833mg/kg	++	IgM to SRBC by ELISA	
Zheng 2009	Mice, c57BL/6 (male)	7 days	gavage	40mg/kg	++	IgM to SRBC by PFC	
Zheng 2011	Mice, c57BL/6 (male)	7 days	gavage	20mg/kg	++	IgM to SRBC by ELISA	
Peden-Adams 2008	Mice, B6C3F1 (male)	28 days	gavage	0.18mg/kg	+	IgM to SRBC by PFC	
	Mice, B6C3F1 (female)	28 days	gavage	0.18mg/kg	+	IgM to SRBC by PFC	
	Mice, B6C3F1 (female)	21 days	gavage	0.18mg/kg	+	IgM to TNP-LPS by ELISA	
Keil 2008	Mice, B6C3F1 (male)	gestation	gavage	5mg/kg	-	IgM to SRBC by PFC	
	Mice, B6C3F1 (female)	gestation	gavage	5mg/kg	-	IgM to SRBC by PFC	
Qazi 2010	Mice, B6C3F1 (male)	28 days	diet	0.25mg/kg	-	IgM to SRBC PFC/ELISA	
	Mice, B6C3F1 (male)	28 days	diet	0.25mg/kg	-	IgM to TNP-LPS by ELISA	
Lefebvre 2008	Rats, SD (m/f)	28 days	diet	7.58 mg/kg	-	IgG to KLH by ELISA	



- **Factors to consider for consistency of PFOS-antibody response**
  - Lefebvre 2008: Species, route, randomization, antibody response (IgG), antigen (KLH)
  - Qazi 2010: route, randomization,
  - Keil 2008: gestational exposure, randomization

Downgrade Considerations: Inconsistency

Reference	Species, Strain (sex)	Duration	Route	Max Dose/day	Risk of Bias: Randomization	Antibody Endpoint	Response
Dong 2009	Mice, c57BL/6 (male)	60 days	gavage	2.0833mg/kg	++	IgM to SRBC by PFC	
Dong 2011	Mice, c57BL/6 (male)	60 days	gavage	0.833mg/kg	++	IgM to SRBC by ELISA	
Zheng 2009	Mice, c57BL/6 (male)	7 days	gavage	40mg/kg	++	IgM to SRBC by PFC	
Zheng 2011	Mice, c57BL/6 (male)	7 days	gavage	20mg/kg	++	IgM to SRBC by ELISA	
Peden-Adams 2008	Mice, B6C3F1 (male)	28 days	gavage	0.18mg/kg	+	IgM to SRBC by PFC	
	Mice, B6C3F1 (female)	28 days	gavage	0.18mg/kg	+	IgM to SRBC by PFC	
	Mice, B6C3F1 (female)	21 days	gavage	0.18mg/kg	+	IgM to TNP-LPS by ELISA	
Keil 2008	Mice, B6C3F1 (male)	gestation	gavage	5mg/kg	-	IgM to SRBC by PFC	
	Mice, B6C3F1 (female)	gestation	gavage	5mg/kg	-	IgM to SRBC by PFC	
Qazi 2010	Mice, B6C3F1 (male)	28 days	diet	0.25mg/kg	-	IgM to SRBC PFC/ELISA	
	Mice, B6C3F1 (male)	28 days	diet	0.25mg/kg	-	IgM to TNP-LPS by ELISA	
Lefebvre 2008	Rats, SD (m/f)	28 days	diet	7.58 mg/kg	-	IgG to KLH by ELISA	



Downgrade?

Initial Confidence

- High (++++)
- Moderate (+++)
- Low (++)
- Very Low (+)

Factors Decreasing Confidence

No → risk of bias

- unexplained inconsistency
- indirectness
- imprecision
- publication bias

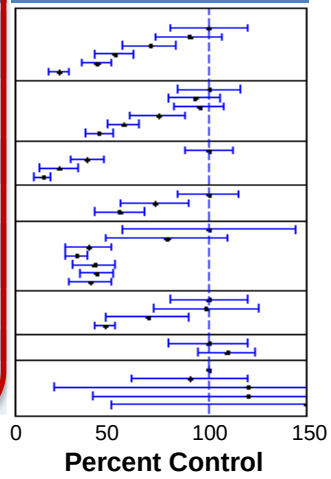
no

Confidence Rating

- High (++++)
- Moderate (+++)
- Low (++)
- Very Low (+)

Downgrade Considerations: Indirectness

Reference	Species, Strain (sex)	Duration	Route	Dose	Antibody Endpoint	Response
Dong 2009	Mice, c57BL/6 (male)	60 days	gavage	0, 0.0083 to 2.0833mg/kg-d	IgM to SRBC by PFC	
Dong 2011	Mice, c57BL/6 (male)	60 days	gavage	0, 0.0083 to 0.833mg/kg-d	IgM to SRBC by ELISA	
Zheng 2009	Mice, c57BL/6 (male)	7 days	gavage	0, 5, 20, 40mg/kg-d	IgM to SRBC by PFC	
Zheng 2011	Mice, c57BL/6 (male)	7 days	gavage	0, 5, 20mg/kg-d	IgM to SRBC by ELISA	
Peden-Adams 2008	Mice, B6C3F1 (m/f)	28 days	gavage	0, 0.00018 to 0.18mg/kg-d	IgM to SRBC by PFC	
Keil 2008	Mice, B6C3F1 (m/f)	gestation	maternal gavage	0, 0.1, 1, 5mg/kg-d	IgM to SRBC by PFC	
Qazi 2010	Mice, B6C3F1 (male)	28 days	diet	0, 0.25mg/kg-day	IgM to SRBC by PFC*	
Lefebvre 2008	Rats, SD (m/f)	28 days	diet	0, 0.14, to 7.58 mg/kg-d	IgG to KLH by ELISA	



Downgrade?

Initial Confidence

- High (++++)
- Moderate (+++)
- Low (++)
- Very Low (+)

Factors Decreasing Confidence

No → risk of bias

No → unexplained inconsistency

- indirectness
- imprecision
- publication bias

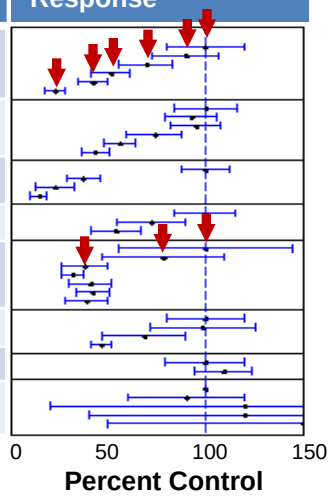
no

Confidence Rating

- High (++++)
- Moderate (+++)
- Low (++)
- Very Low (+)

Downgrade Considerations: Imprecision

Reference	Species, Strain (sex)	Duration	Route	Dose	Antibody Endpoint	Response
Dong 2009	Mice, c57BL/6 (male)	60 days	gavage	0, 0.0083 to 2.0833mg/kg-d	IgM to SRBC by PFC	
Dong 2011	Mice, c57BL/6 (male)	60 days	gavage	0, 0.0083 to 0.833mg/kg-d	IgM to SRBC by ELISA	
Zheng 2009	Mice, c57BL/6 (male)	7 days	gavage	0, 5, 20, 40mg/kg-d	IgM to SRBC by PFC	
Zheng 2011	Mice, c57BL/6 (male)	7 days	gavage	0, 5, 20mg/kg-d	IgM to SRBC by ELISA	
Peden-Adams 2008	Mice, B6C3F1 (m/f)	28 days	gavage	0, 0.00018 to 0.18mg/kg-d	IgM to SRBC by PFC	
Keil 2008	Mice, B6C3F1 (m/f)	gestation	maternal gavage	0, 0.1, 1, 5mg/kg-d	IgM to SRBC by PFC	
Qazi 2010	Mice, B6C3F1 (male)	28 days	diet	0, 0.25mg/kg-day	IgM to SRBC by PFC*	
Lefebvre 2008	Rats, SD (m/f)	28 days	diet	0, 0.14, to 7.58 mg/kg-d	IgG to KLH by ELISA	



Downgrade?

Initial Confidence

- High (++++)
- Moderate (+++)
- Low (++)
- Very Low (+)

Factors Decreasing Confidence

No → risk of bias
No → unexplained inconsistency
No → indirectness
- imprecision
- publication bias

no

Confidence Rating

- High (++++)
- Moderate (+++)
- Low (++)
- Very Low (+)

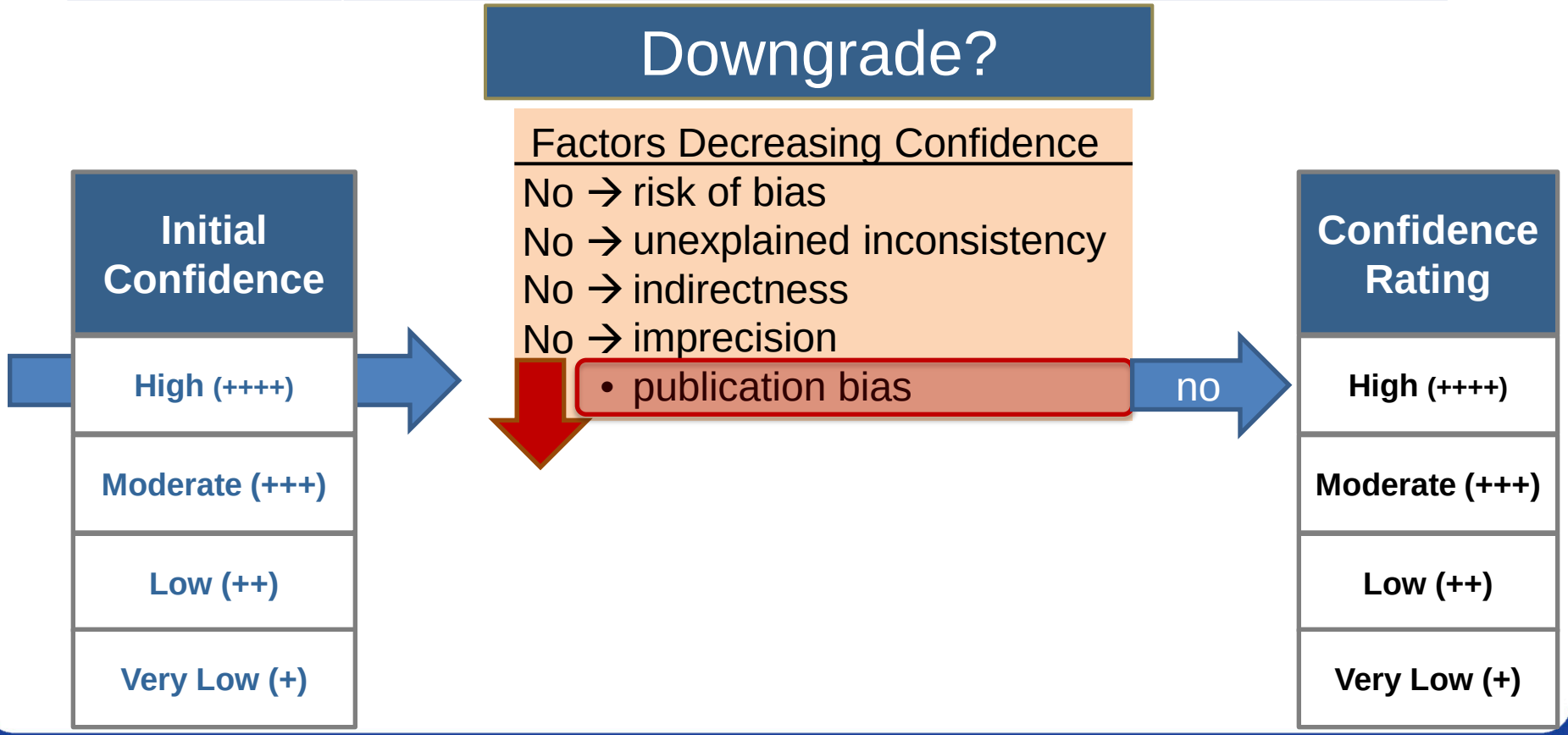
Downgrade Considerations: Publication Bias

- Evidence of unpublished studies or “lag bias”?
  - 25 meeting abstracts or reports immune/ PFOS/ PFOA in last 10 years
  - 6 on antibody response
    - All resulted in publications
    - Only 2 on PFOS

Meeting Abstract	Chemical	Published or missing
Keil et al., 2005	**PFOS**	published
Luebke et al., 2006	PFOA	published
Dewitt et al., 2009	PFOA	published
Dewitt et al., 2009	PFOA	published
Peden-Adams, 2009	**PFOS**	published
Loveless, 2009	PFOA	published

Downgrade Considerations: Publication Bias

Reference	Funding Source
Dong 2009	National Natural Science Foundation of China and China Medical University
Dong 2011	National Natural Science Foundation of China and Liaoning Province Sci.& Tech, and Ed. Foundation
Zheng 2009	National Natural Science Foundation of China and China Medical University
Zheng 2011	National Natural Science Foundation of China and Liaoning Province Sci.& Tech, and Ed. Foundation
Peden-Adams 2008	Medical College of South Carolina and Nevada EPSCOR undergraduate fellowship
Keil 2008	National Institute for Occupational Safety
Qazi 2010	3M Company
Lefebvre 2008	Health Canada



An Evidence Profile Transparently Outlines the Basis for Decreasing the Confidence Rating

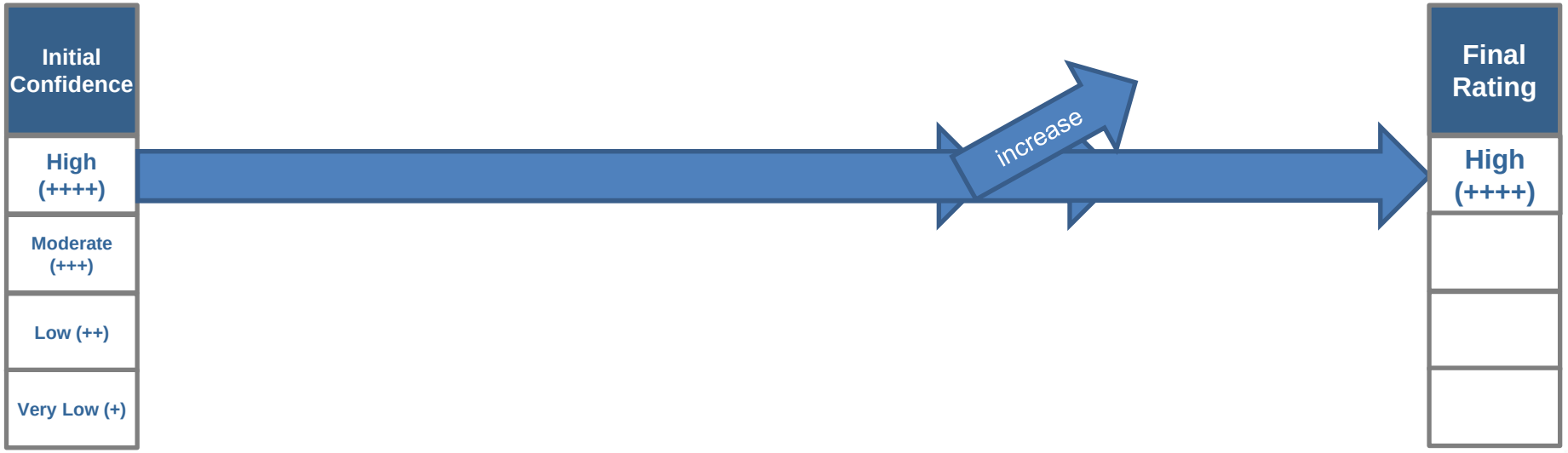
- Reaching and communicating judgments
  - Graphical representations
  - Summary text

Judgments are reached and documented for each factor

Body of Evidence	Risk of Bias	Unexplained Inconsistency	Indirectness	Imprecision	Publication Bias
Animal	not serious	not serious	not serious	not serious	undetected
(8 rodent studies) Initial Rating = High	• General low • Key questions − Randomize – mixed low and probably high − Outcome-low for all studies • Probably high for allocation concealment	• Consistent suppression • Potential inconsistent response, most apparent inconsistency differed by: − Species (Rat vs mouse) − Outcome (IgG vs IgM) − Antigen (SRBC vs KLH) − Risk of bias (randomize)	• SRBC IgM response by PFC or ELISA are among best measures of antibody response	• General small, confidence interval (CI) • Non-overlapping CIs between control and exposed	• No evidence of lag bias • Funding − Government − University − Industry • Studies report no conflict of interest

Confidence Rating: Evidence Profile

Body of Evidence	Risk of Bias	Unexplained Inconsistency	Indirectness	Imprecision	Publication Bias	Magnitude	Dose Response	Residual Confounding	Consistency Across Species/ Model	Final Rating
Animal	not serious	not serious	not serious	not serious	undetected	not large	**yes (increase)**	no	no	**HIGH**
(8 Studies) Initial Rating • High	• General low • Key question – Randomize= mixed low and probably high – Outcome=low • Probably high for allocation concealment	• Consistent suppression • Potential inconsistent response, most apparent inconsistency differed by: – Species (rat vs mouse), – Outcome (IgG vs IgM), – Antigen (SRBC vs KLH) – Risk of Bias (randomize)	• SRBC IgM response by PFC or ELISA are among best measures of antibody response	• General small, confidence interval (CI) • Non-overlapping CIs between control and exposed	• No evidence of lag bias • Funding – Government – Universities – Industry	• Not sufficiently large to overcome potential bias	• Dose-response observed in multiple studies	• No evidence of confounding that would bias toward null	• All positive results from mice	



Lessons Learned Rating the Body of Evidence

- Initial confidence rating by key study design features transparently grouped studies
- Summary text as well as graphical aids are helpful for reaching and communicating confidence ratings for bodies of evidence
  - Graphical tools are key to developing and communicating ratings
  - The evidence profile should contain brief explanation of ratings
- Publication bias is difficult to ascertain
  - Track meeting abstracts to look for lag bias
  - Examine sources of funding to look for potential bias
- **Questions?**

Systematic Review is Feasible in Environmental Health

- OHAT framework accommodates changes to address the specifics of each assessment
- “Handbook” with instructions for developing protocols will be posted on NTP website
- Contribution and ongoing need for development and refinement of software tools
- Focus on problem formulation promotes early public outreach
- Greater transparency provided by systematic review is worth the time-investment of learning methods
  - Case studies helped identify efficiencies and value-added steps Example:
    - Data entry allows graphical display
    - Graphical displays facilitate the process of rating confidence

Acknowledgements

- **Office of Health Assessment and Translation**
  - Abee Boyles
  - Kembra Howdeshell
  - Katie Pelch
  - Andrew Rooney, Deputy Director
  - Kyla Taylor
  - Kristina Thayer, Director
  - Vickie Walker
- **Office of Liaison, Policy and Review**
  - Mary Wolfe, Director
  - Lori White
- **Office of Library and Information Services**
  - Stephanie Holmgren
- **Approach Technical Advisors and Experts**
  - **Lisa Bero,** Director, San Francisco Branch, United States Cochrane Center at UC San Francisco
  - **Gordon Guyatt,** Co-chair, GRADE Working Group, McMaster U
  - **Malcolm Macleod,** CAMARADES Centre, University of Edinburgh
  - **Karen Robinson,** Co-Director, Evidence-Based Practice Center, The Johns Hopkins Bloomberg School of Public Health
  - **Holger Schünemann**, Co-chair, GRADE Working Group, McMaster U.
  - **Tracey Woodruff,** Director, Program on Reproductive Health and the Environment, UCSF
- **NTP Board of Scientific Counselors**
- **NTP BSC Working Group**
  - **Lynn Goldman, Chair,** Dean, School of Public Health and Health Services, George Washington U.
  - **Reeder Sams, Vice-chair,** Acting Deputy Director, NCEA/RTP Division, USEPA
  - **Lisa Bero,** Director, San Francisco Branch, United States Cochrane Center at UC San Francisco
  - **Edward Carney,** Senior Science Leader, Mammalian Toxicology, Dow Chemical Company
  - **David Dorman,** Professor, North Carolina State University
  - **Elaine Faustman**, Director, Institute for Risk Analysis and Risk Communication, U. Washington
  - **Dale Hattis**, Research Professor, George Perkins Marsh Institute, Clark University
  - **Malcolm Macleod,** CAMARADES Centre, University of Edinburgh
  - **Tracey Woodruff,** Director, Program on Reproductive Health and the Environment, UCSF
  - **Lauren Zeise,** Chief, Reproductive and Cancer Hazard Assessment Branch, OEHHA, California EPA
- **Protocol Technical Advisors**
- **Public Comment**

Questions or Comments?