NTP
National Toxicology Program

Systematic Review and Evidence Integration for Literature-Based Health Assessments

Andrew Rooney, PhD
National Institute of Environmental Health Sciences

NTP Board of Scientific Counselors Meeting
June 25, 2013

Timeline

- December 2012: NTP Board of Scientific Counselors Meeting
- February 2013: Framework Released for Public Comment
  - Draft OHAT Approach for Systematic Review and Evidence Integration
- April 6, 2013: Case-Study Protocols Released for Public Comment
  - Draft Protocol to Evaluate the Evidence for an Association Between Bisphenol A (BPA) Exposure and Obesity
  - Draft Protocol to Evaluate the Evidence for an Association Between Perfluorooctanoic Acid (PFOA) or Perfluorooctane Sulfonate (PFOS) and Immunotoxicity
- April 23, 2013: Public Q&A at Web-Based Informational Meeting

Draft OHAT Approach for Systematic Review and Evidence Integration for Literature-Based Health Assessments

Step 1: Prepare topic

Step 2: Search for and select studies

Step 3: Extract data from studies

Step 4: Assess individual study quality

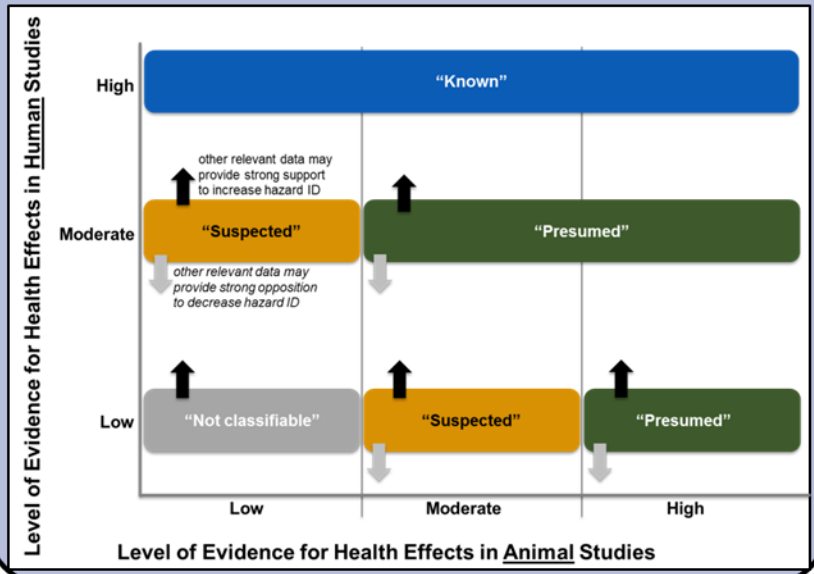
Step 5: Rate confidence in body of evidence

Initial Confidence by Key Features of Study Design	Factors Decreasing Confidence	Factors Increasing Confidence	Confidence in the Body of Evidence
High (++++) 4 Features	❖ Risk of Bias ❖ Unexplained Inconsistency ❖ Indirectness ❖ Imprecision ❖ Publication Bias	❖ Large Magnitude of Effect ❖ Dose Response ❖ All Plausible Confounding • Studies report an effect and residual confounding is toward null • Studies report no effect and residual confounding is away from null ❖ Consistency • Across animal models or species • Across dissimilar populations • Across study design types ❖ Other e.g., particularly rare outcomes	High (++++)
Moderate (+++) 3 Features			Moderate (+++)
Low (++) 2 Features			Low (++)
Very Low (+) ≤1 Features			Very Low (+)

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

Step 6: Translate confidence ratings into level of evidence for health effect

Step 7: Integrate evidence to develop hazard identification conclusions



Presentation Overview

- Major Technical and Scientific Questions Moving Forward
- How Comments Have Informed the Issues
- Outline How NTP is Trying to Reach Resolution
- Illustrate Our Initial Approach with Examples from Case-Studies
- Discussion with the NTP Board of Scientific Counselors

Major Technical and Scientific Questions Moving Forward

- **How Does the Approach Address Study Quality?**
- Excluding Studies or “Tiers” Based on Quality
- Confidence in Body of Evidence – Initial Confidence Rating
- Consideration of Other Relevant Data (e.g., mechanistic)

Many Comments on Study Quality

- Support for Study Quality as Internal Validity or Risk of Bias
- Don’t Restrict Study Quality to Internal Validity
- Suggested Additions

Study Quality in Different Steps of Approach

- Internal Validity or Risk of Bias (STEP 4)
  - Completeness of reporting
  - Confounding
  - Study design and conduct
- External Validity or Directness and Applicability (STEP 5)
  - Route of exposure
  - Timing and duration of exposure
  - Relevance of animal model for human health
- Continued Evaluation
  - Conflict of interest
  - Power

Major Technical and Scientific Questions Moving Forward

- How Does the Approach Address Study Quality?
- **Excluding Studies or “Tiers” Based on Quality**
- Confidence in the Body of Evidence – Initial Confidence Rating
- Consideration of Other Relevant Data (e.g., mechanistic)

Excluding Studies or “Tiers” Based on Quality

- Study Quality Impacts Confidence in the Conclusions
  - Should all studies contribute to the conclusions?
  - Can studies have too many problems with internal validity or risk of bias?
  - Would confidence be “diluted”?
- Exclude Studies for Established Reasons in Protocol (STEPs 1&2)

Individual Study Quality (STEP 4)

Studies (on outcome basis)	Example Answers to Risk of Bias Questions												
	Question #1	Question #2	Question #3	Question #4	Question #5	Question #6	Question #7	Question #8	Question #9	Question #10	Question #11	Question #12	Question #13
Bucher et al., 2002	++	++	++	++	++	++	++	++	++	++	++	+	+
Wolfe et al., 2000	++	++	++	+	+	+	+	+	+	–	–	–	––
Thayer et al., 2010	++	++	++	++	+	+	+	+	–	–	–	–	––
Boyles et al., 2011	++	++	+	+	+	+	+	–	–	–	–	––	––
Rooney et al., 2013	++	+	–	–	–	–	–	–	–	–	––	––	––

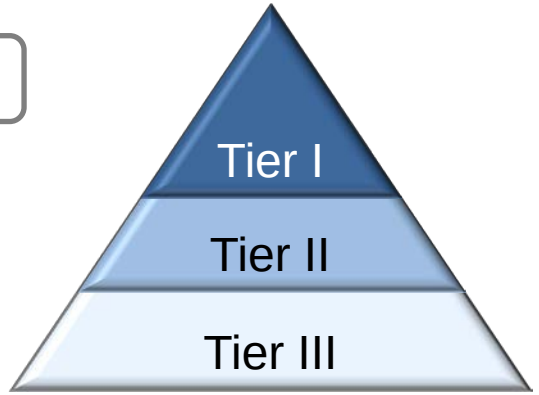
Answers on 4-point scale

++ Definitely **Low** Risk of Bias + Probably **Low** Risk of Bias – Probably **High** Risk of Bias –– Definitely **High** Risk of Bias

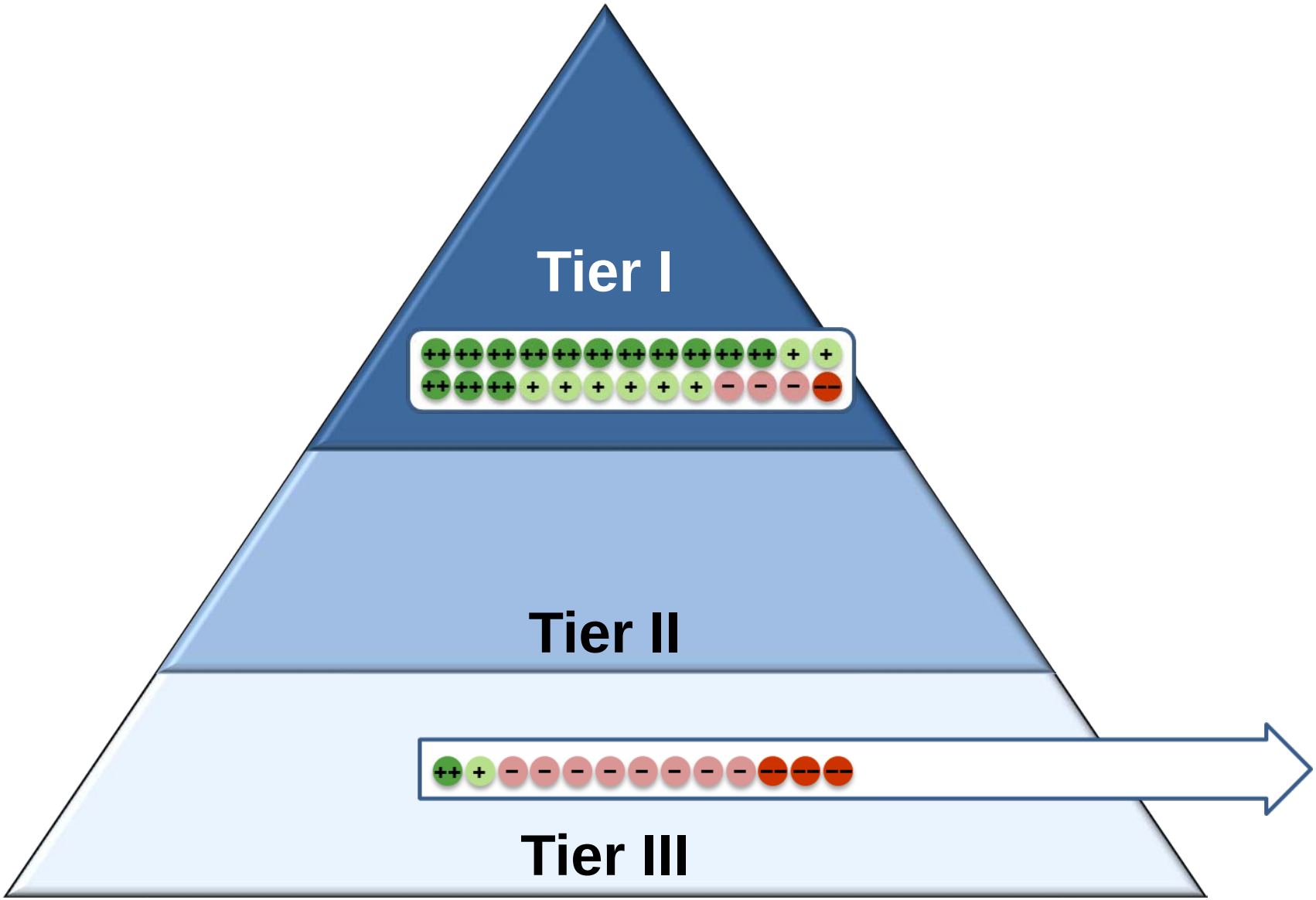
Using Individual Study Quality in Next STEPS

- "Tiers" from Individual Study Quality Assessed in STEP 4
  - Restrict confidence rating conclusions to top tier studies
  - How do we assess the impact of removing low-quality studies on confidence conclusions developed in STEP 5?

High quality = Fewer challenges to internal validity



Sensitivity Analysis



Major Technical and Scientific Questions Moving Forward

- How Does the Approach Address Study Quality?
- Excluding Studies or Tiers Based on Quality
- **Confidence in Body of Evidence – Initial Confidence Rating**
- Consideration of Other Relevant Data (e.g., mechanistic)

Draft OHAT Approach for Systematic Review and Evidence Integration for Literature-Based Health Assessments

You Are HERE

Step 1: Prepare topic

Step 2: Search for and select studies

Step 3: Extract data from studies

Step 4: Assess individual study quality

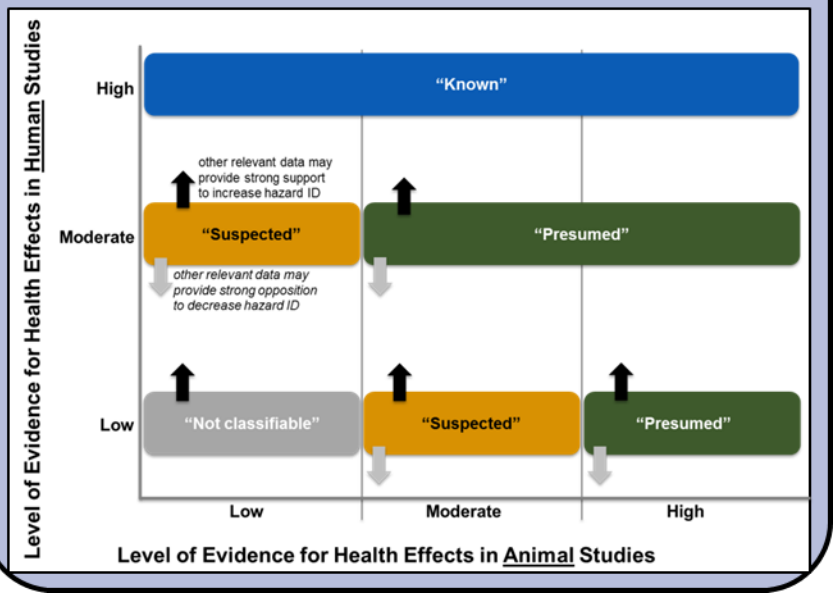
Step 5: Rate confidence in body of evidence

Initial Confidence by Key Features of Study Design	Factors Decreasing Confidence	Factors Increasing Confidence	Confidence in the Body of Evidence
High (++++) 4 Features	❖ Risk of Bias ❖ Unexplained Inconsistency ❖ Indirectness ❖ Imprecision ❖ Publication Bias	❖ Large Magnitude of Effect ❖ Dose Response ❖ All Plausible Confounding • Studies report an effect and residual confounding is toward null • Studies report no effect and residual confounding is away from null ❖ Consistency • Across animal models or species • Across dissimilar populations • Across study design types ❖ Other e.g., particularly rare outcomes	High (++++)
Moderate (+++) 3 Features			Moderate (+++)
Low (++) 2 Features			Low (++)
Very Low (+) ≤1 Features			Very Low (+)

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

Step 6: Translate confidence ratings into level of evidence for health effect

Step 7: Integrate evidence to develop hazard identification conclusions



Confidence in the Body of Evidence (Step 5) Initial Confidence Rating

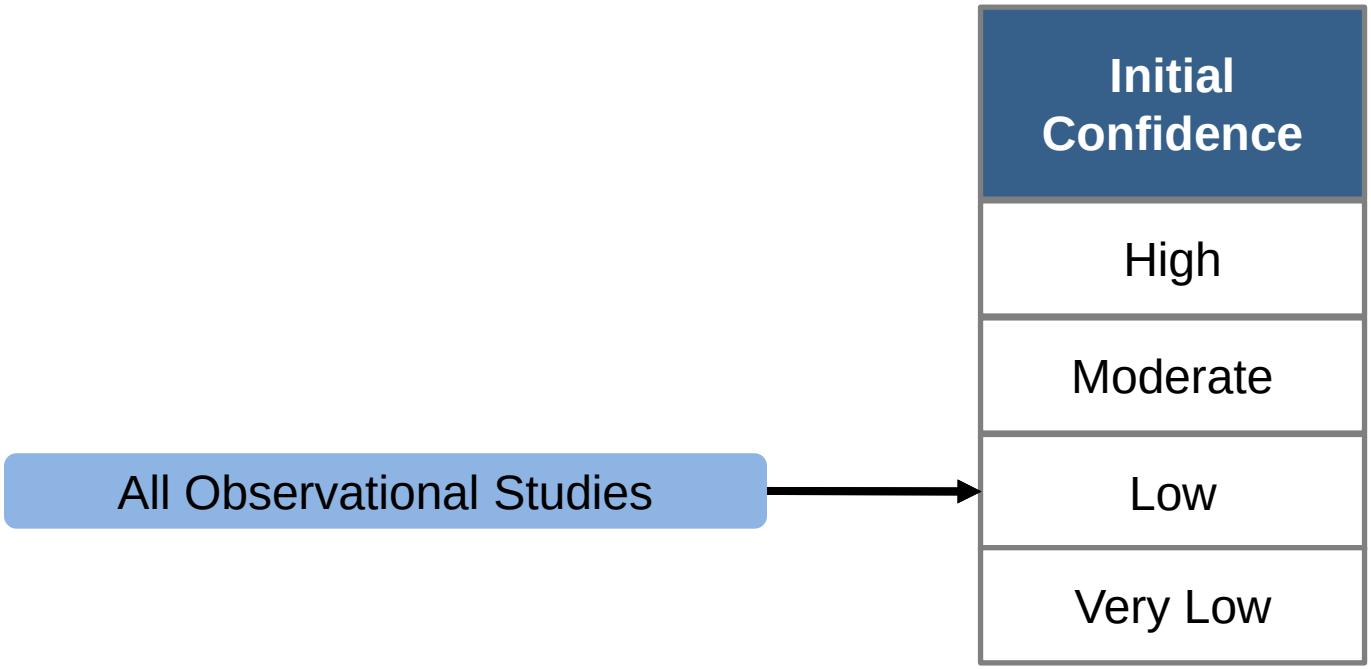
- Based on Established Method (GRADE)
  - Clear presentation of elements considered for downgrading or upgrading confidence in a body of evidence
  - Framework for documenting scientific judgment decisions
  - Elements cover Bradford Hill causality considerations

- Initial Confidence
  - Where do you start?

Initial Confidence in Body of Evidence

- Initial Confidence Based on Study Design

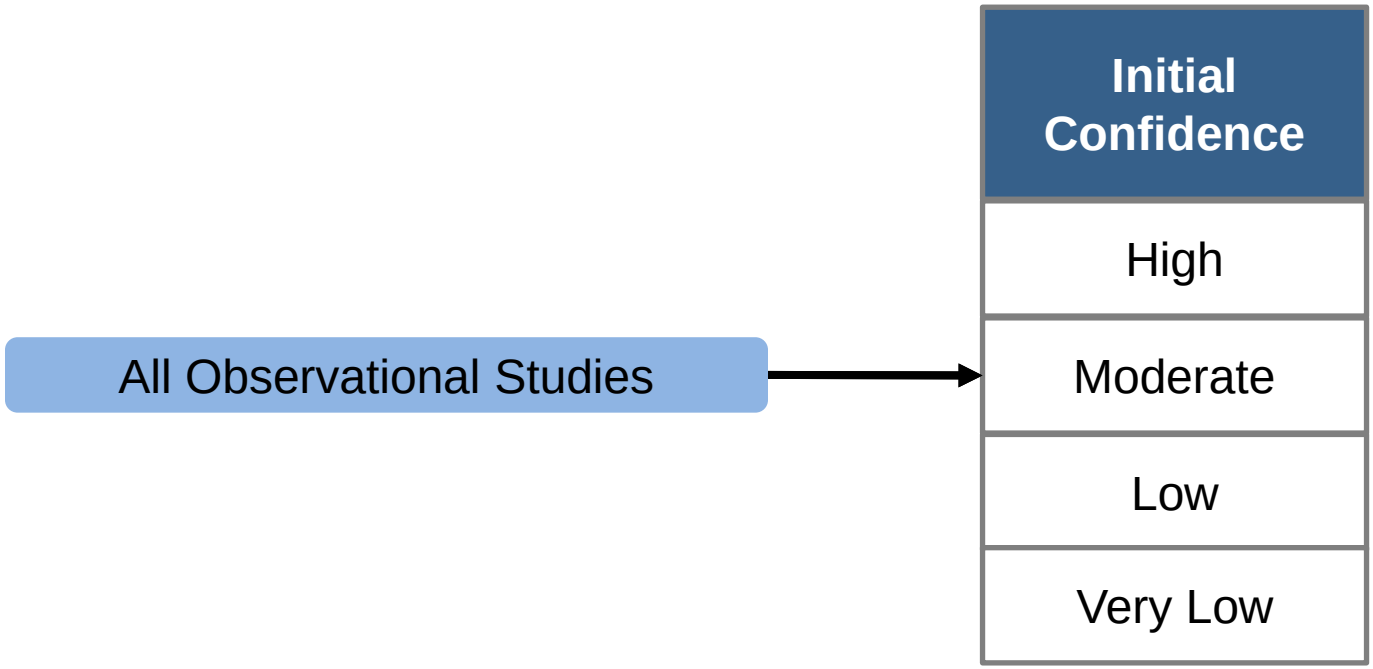
- Options for Observational Studies
  - Start all observational studies as “low” (GRADE)



Initial Confidence in Body of Evidence

- Initial Confidence Based on Study Design

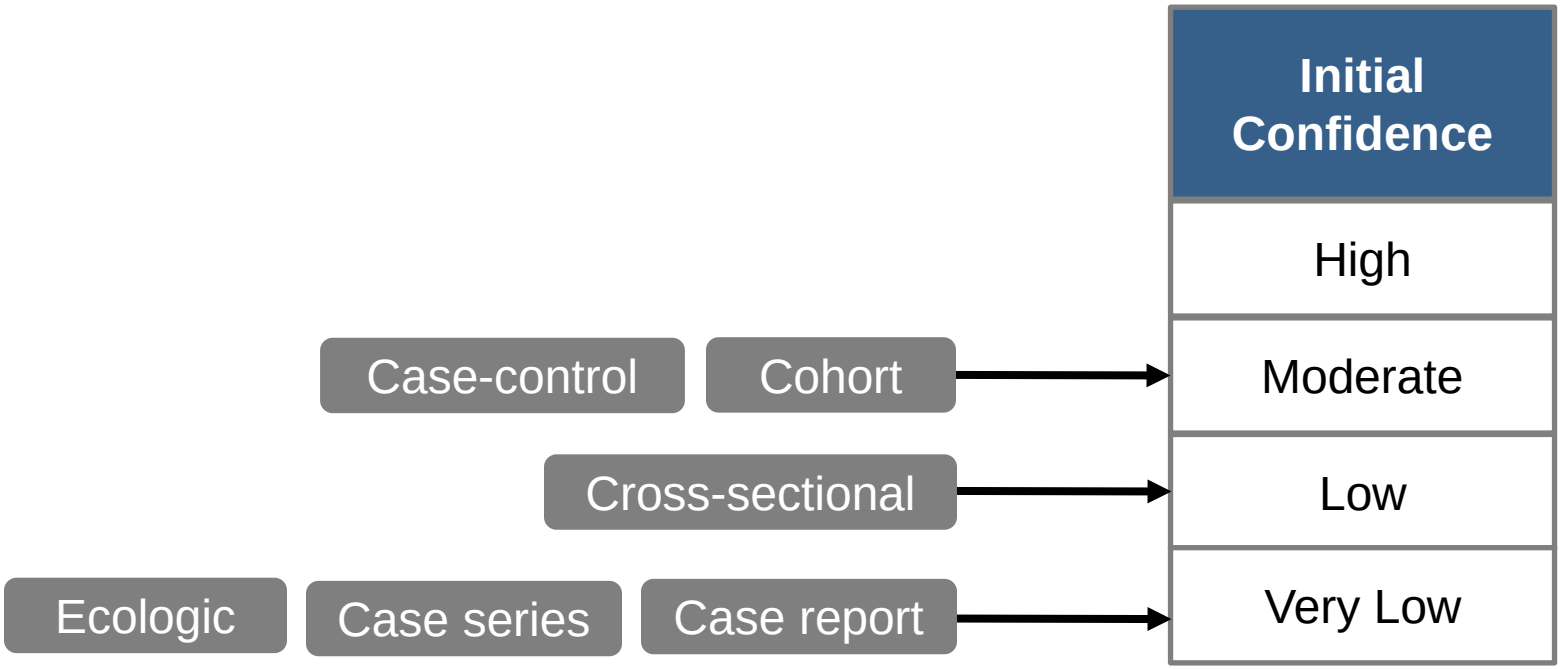
- Options for Observational Studies
  - Start all observational studies as “moderate” (Navigation Guide)



Initial Confidence in Body of Evidence

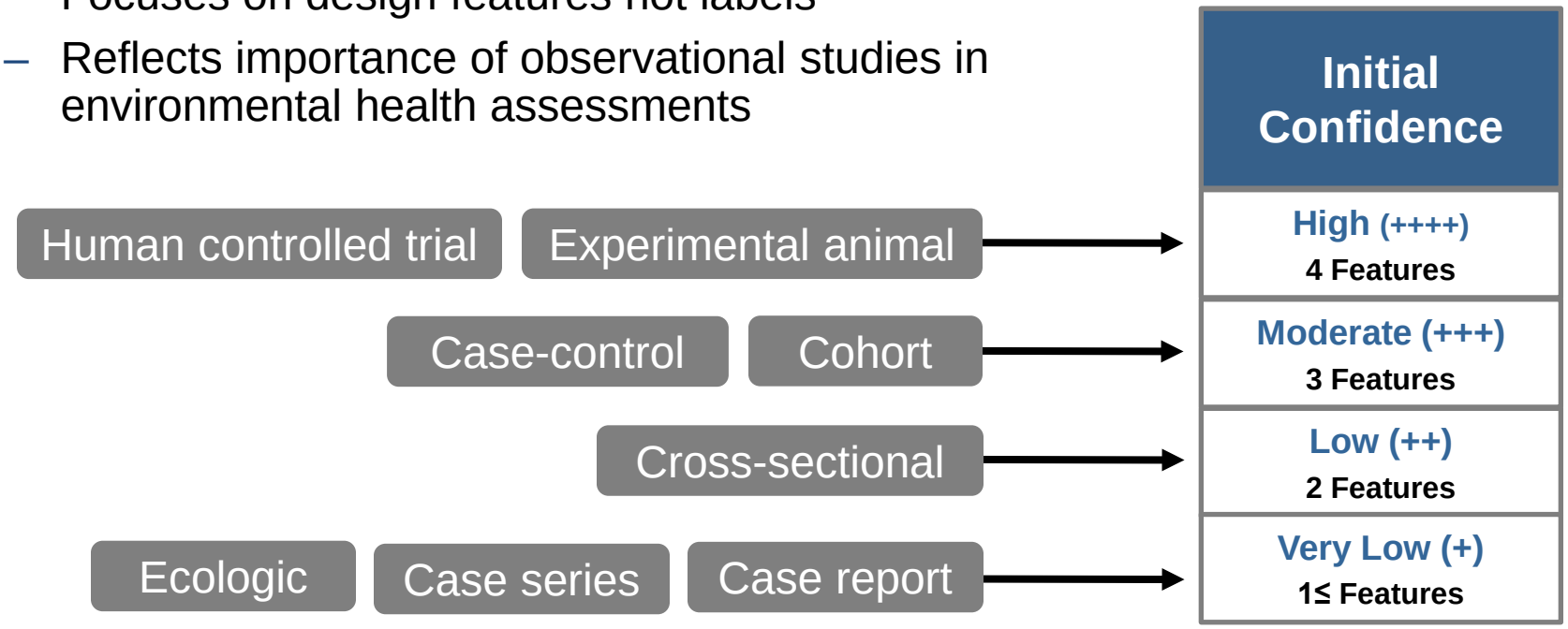
- Initial Confidence Based on Study Design

- Options for Observational Studies
  - Stratify based on study design labels (Initial OHAT method)



Initial Confidence in Body of Evidence

- Initial Confidence Based on Key Study Design Features (current)
  - Controlled exposure
  - Exposure prior to outcome
  - Individual outcome data
  - Comparison group used
- This Method Stratifies Initial Confidence:
  - Focuses on design features not labels
  - Reflects importance of observational studies in environmental health assessments



Case-Study Example: Initial Confidence

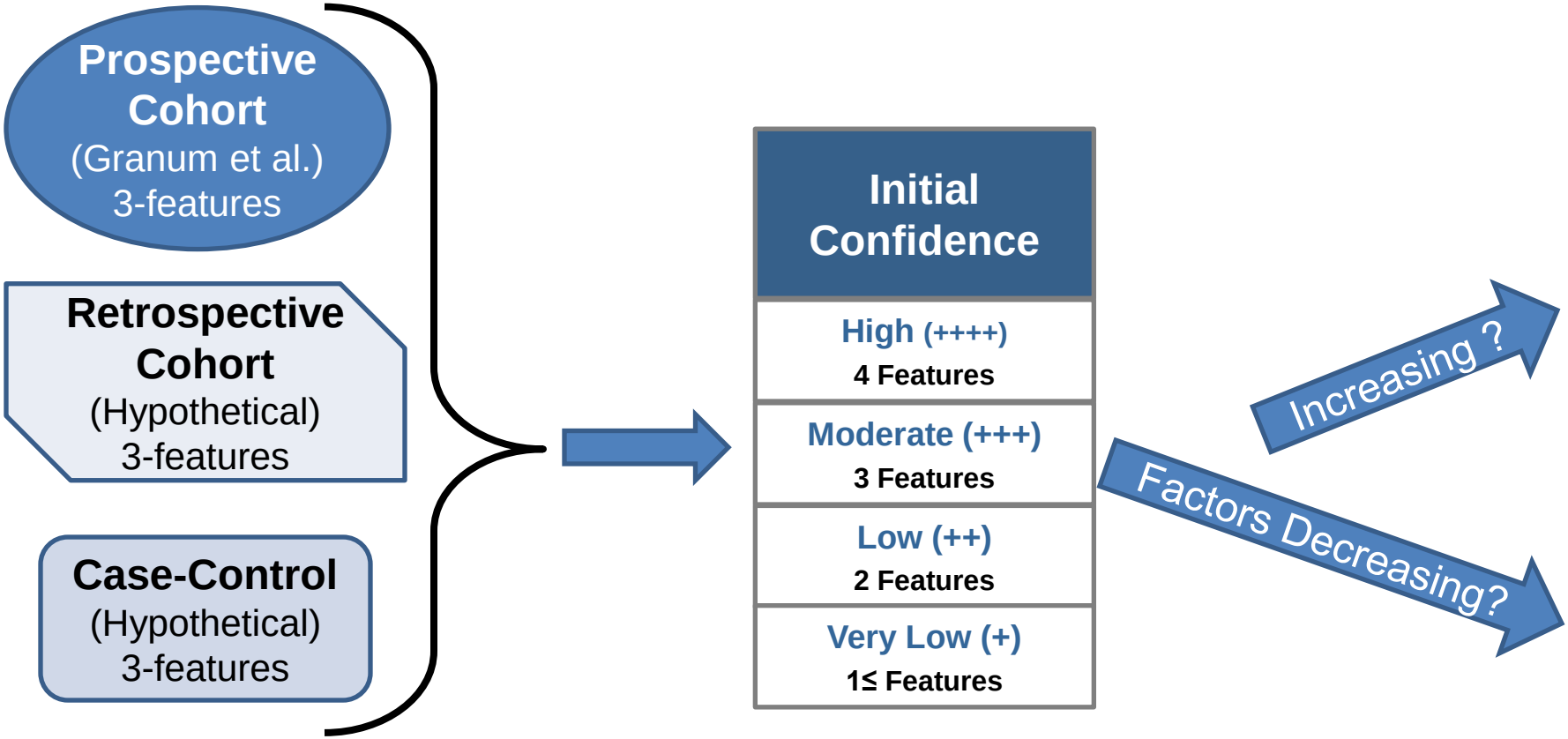
Study Design Feature	Granum *et al.*, 2013 (prospective birth-cohort) (sub-cohort of Norwegian Mother and Child Cohort Study)	
Controlled exposure	No	
Exposure prior to outcome	Yes	Maternal blood levels at delivery
		Child blood levels at 3 years of age
Individual outcome data	Yes	Measured in 3-year-old children
Comparison group used	Yes	Multivariate regression of exposure (PFOA or PFOS) and health outcomes

Granum et al. (2013) *J Immunotoxicology*

- **Initial Confidence of a Single Study**
  - GRADE: Low
  - Navigation Guide: Moderate
  - Initial OHAT method (“label”): Moderate
  - Current OHAT method (“design feature”): Moderate

Initial Confidence by Study Design Features

- Starting Point for Evaluating Confidence in the Body of Evidence



Major Technical and Scientific Questions Moving Forward

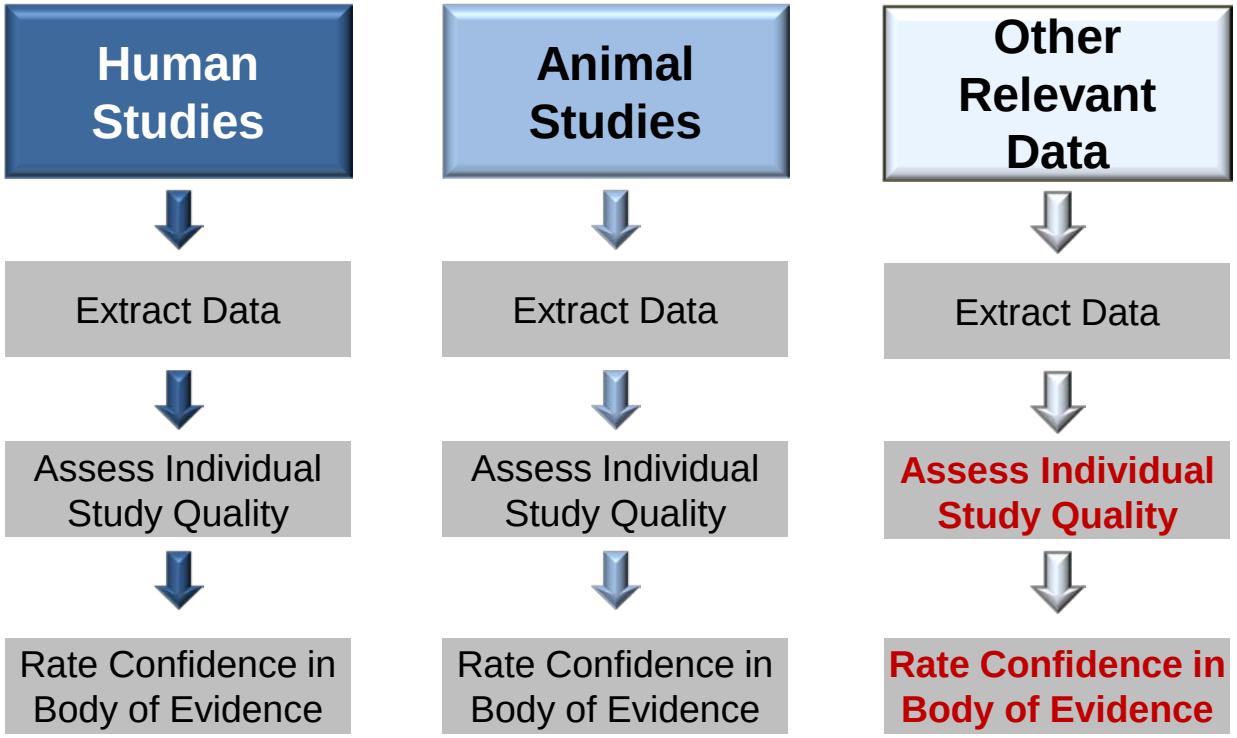
- How Does the Approach Address Study Quality?
- Tiering or Excluding Studies Based on Quality
- Confidence in the Body of Evidence – Initial Confidence Rating
- **Consideration of Other Relevant Data (e.g., mechanistic)**

Consideration of Other Relevant Data

- Three Evidence Streams
  - Human studies
  - Animal studies (non-human)
  - Other relevant data (*in vitro*, mechanistic, etc.)

- Need:
  - To develop a parallel approach for considering other relevant data
  - To prepare for the future – datasets lacking human and animal studies

Challenges to Parallel Approach



- **Near-term Research**:
  Explore development of a study quality (internal validity) tool for *in vitro* studies
- **Biological Plausibility:**
  Considering factors that parallel those used to evaluate confidence in other evidence streams

How NTP is Trying to Reach Resolution on Major Technical Questions Moving Forward

- Study Quality
  - Internal validity (Step 4)
  - External validity (Directness in Step 5)
- Tiering to Consider the Impact of High Risk of Bias Studies
- Initial Confidence Rating on Study Design Features
- Parallel Approach for Other Relevant Data

Acknowledgements

- **Office of Health Assessment and Translation**
  - Abee Boyles
  - Kembra Howdeshell
  - Andrew Rooney, Deputy Director
  - Michael Shelby
  - 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**