

NATIONAL
ACADEMIES *Sciences
Engineering
Medicine*

Building Confidence in New Evidence Streams for Human Health Risk Assessment

Weihsueh Chiu, Study Chair

September 22, 2023



Acknowledgements

Committee Members

WEIHSUEH A. CHIU (*Chair*), Texas A&M University
KIM BOEKELHEIDE, Brown University School of Medicine
PATIENCE BROWNE, OECD (until September 2022)
HOLLY DAVIES, Washington State Department of Health
CORIE A. ELLISON, Procter and Gamble
MARIE C. FORTIN, Jazz Pharmaceuticals
NICOLE KLEINSTREUER, NIEHS (until September 2022)
NANCY E. LANE, University of California Davis
HEATHER B. PATISAUL, North Carolina State University
ELIJAH J. PETERSEN, National Institute of Standards and Technology
KRISTI PULLEN FEDINICK, Center for Earth, Energy, and Democracy
MARTYN T. SMITH, University of California Berkeley
ROBYN L. TANGUAY, Oregon State University
CHRISTOPHER VULPE, University of Florida Gainesville
TRACEY J. WOODRUFF, University of California San Francisco
JOSEPH C. WU, Stanford University

Consultants to the Committee

A. JOHN BAILER, Miami University
MALCOLM MACLEOD, University of Edinburgh

Other Consultants

GRACE COONEY, ICF Resources LLC
CARY HAVER, ICF Resources LLC
JENNIFER SAUNDERS
JENNA SPROWLES, ICF Resources LLC

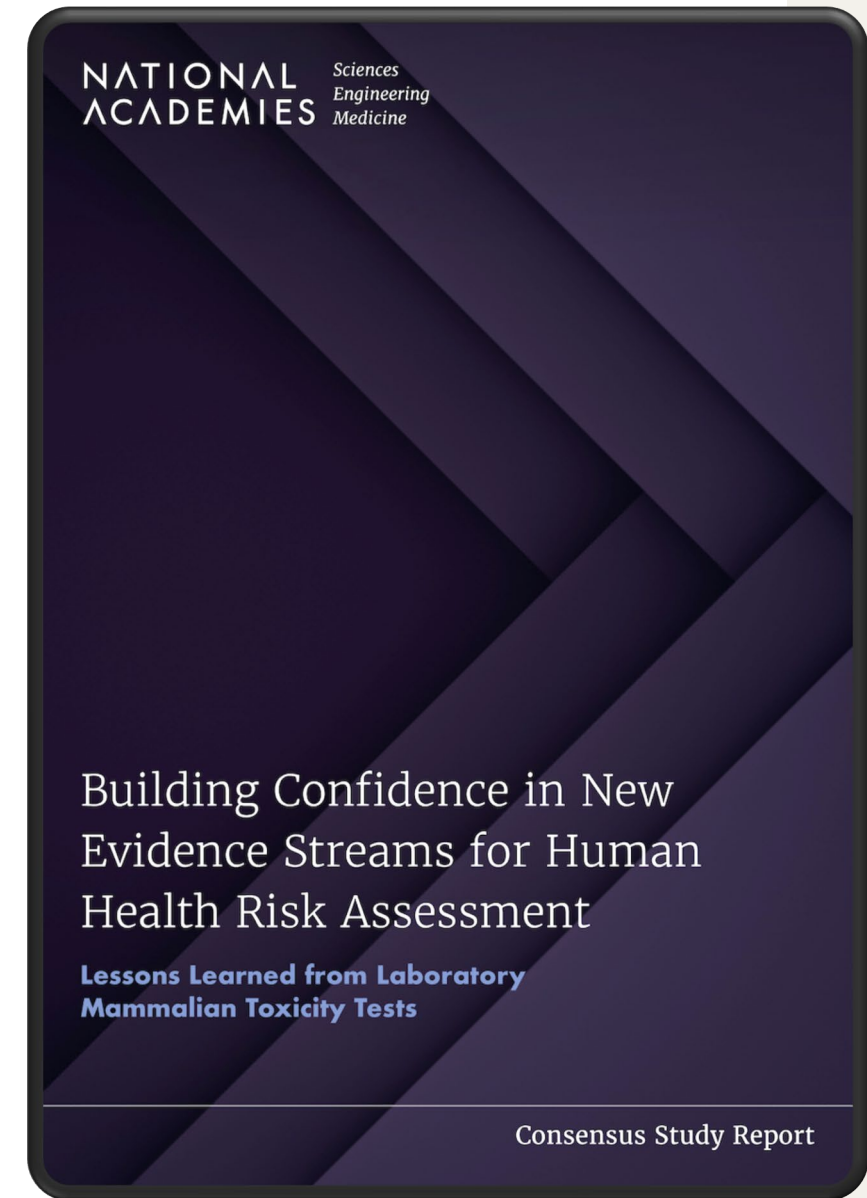
Report Reviewers:

BRIAN BERRIDGE, B2 Pathology Solutions LLC
DEBORAH CORY-SLECHTA, University Of Rochester
REZA J. RASOULPOUR, Corteva Agriscience
ANDREW ROONEY, National Institute Of Environmental Health Sciences
CRAIG ROWLANDS, Underwriters Laboratory Solutions
IVAN RUSYN, Texas A&M University
LAURA VANDENBERG, University Of Massachusetts Amherst
KATRINA WATERS, Pacific Northwest National Laboratory
MAURICE WHELAN, European Commission
LAUREN ZEISE, California Environmental Protection Agency

Review was overseen by:
DAVID DORMAN and **JOEL KAUFMAN (NAM)**

Staff

KATHRYN Z. GUYTON, Project Director
NATALIE ARMSTRONG, Associate Program Officer
LESLIE BEAUCHAMP, Senior Program Assistant
ELIZABETH BOYLE, Senior Program Officer
TAMARA DAWSON, Program Coordinator (until August 2022)
CORRINE LUTZ, Senior Program Officer (until July 2022)
THOMASINA LYLES, Senior Program Assistant



Key Challenges to Progress



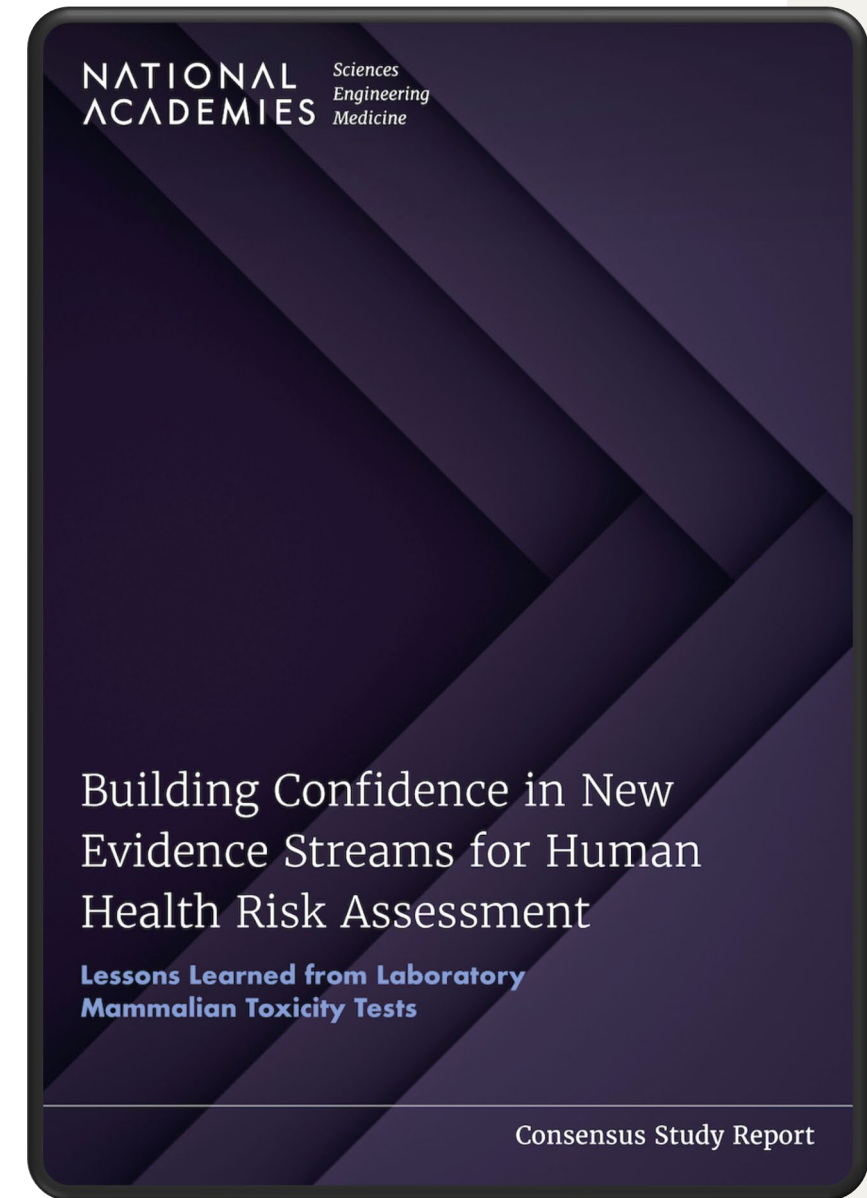
Seminal NASEM reports have provided recommendations for advancing the science of characterizing and assessing exposures and effects of environmental agents.



Other NASEM reports have provided recommendations for advancing the science and practice of risk assessment, focusing on systematic review-based methods.



However, there are few examples of the application of NAMs to inform risk assessment decision-making.



Key Human Health Risk Assessment Themes: *Prior Reports*

- Decreasing reliance solely on apical endpoints and “guideline” studies
- Increasing use of *in vitro* and computational approaches
- Increasing role of systematic review-based evidence assessment methods
- Increasing coverage of susceptible and vulnerable populations



The report recommends that EPA continue using prior NASEM reports for advice and recommendations to improve toxicity testing and human health risk assessment.

Definition of “NAM”

EPA (2021) has defined the term “New Approach Methods” (NAMs) to be “*any technology, methodology, approach, or combination that can provide information on chemical hazard and risk assessment to avoid the use of animal testing.*” For the purposes of TSCA, EPA recognizes this new term (i.e., NAMs) as encompassing any “*alternative test methods and strategies to reduce, refine, or replace vertebrate animals.*”

The committee found EPA’s definition too narrow, creating a false dichotomy between data streams, all of which can be informative for human health risk assessment.

The report recommends that EPA broaden the definition of NAM to encompass the full range of strategies and approaches shown in Figure 2-1, all of which can be informative for human health risk assessment.

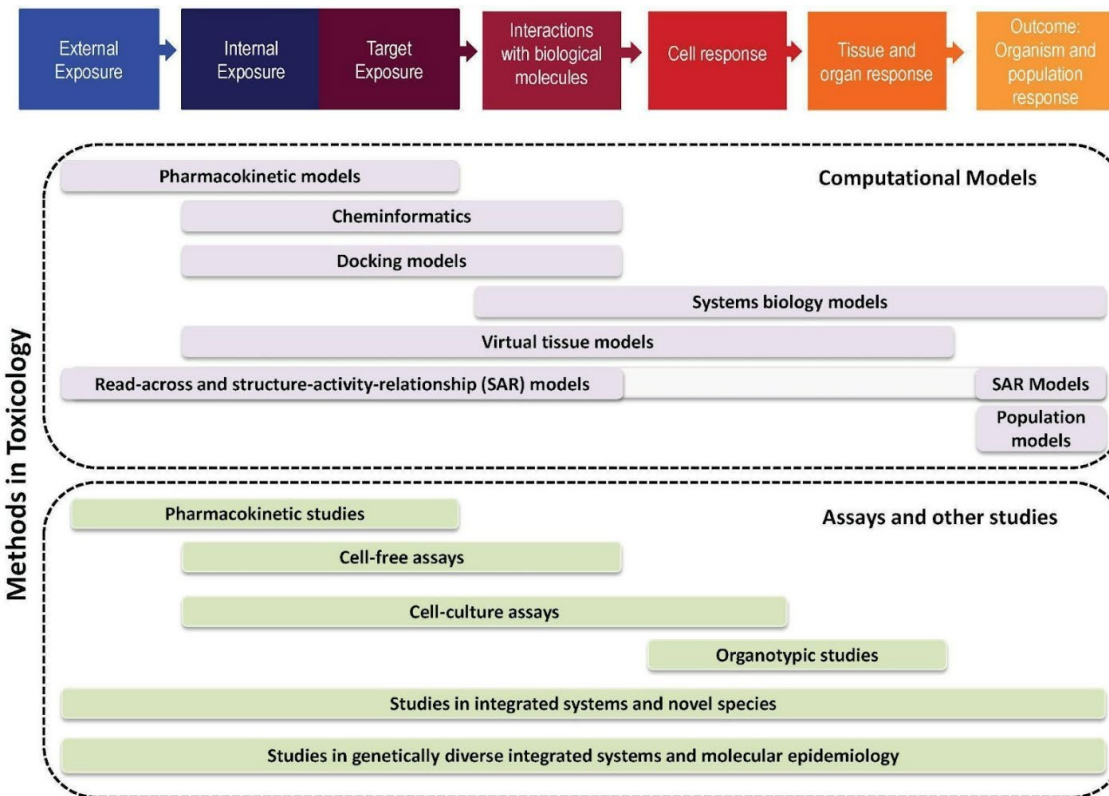


Figure 2-1. A continuum of computational models and biological assays can provide information for human health risk assessment. Reproduced from NASEM (2017) *Using 21st Century Science to Improve Risk-Related Evaluations*, Figure 3-1.

Variability in Toxicity Studies

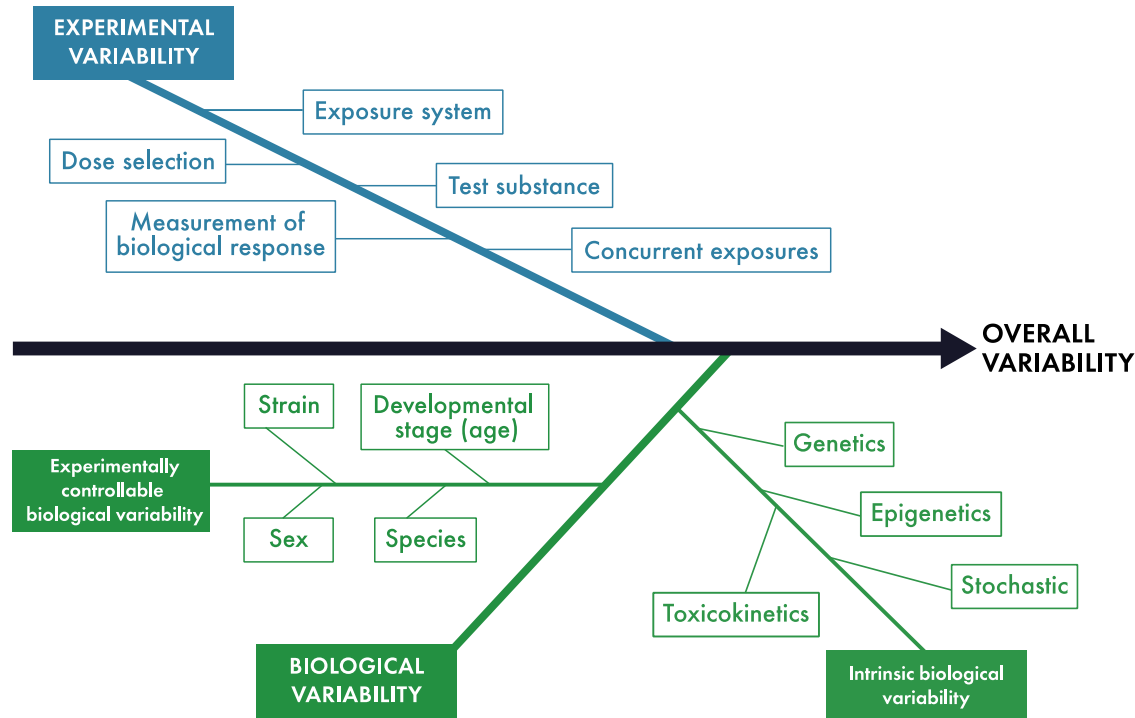


Figure S-3: The multiple sources of variability in laboratory mammalian toxicity tests.

- Any biologically based assay system will always have intrinsic and irreducible biological variation.
- Practically, intrinsic biological variability and experimental variability can be difficult to distinguish.
- **Variability is not fundamentally a negative attribute.** Minimizing variability may limit understanding of the distribution of toxic response, and therefore the generalizability of a study's results.

The report recommends that

- **EPA generally refrain from identifying a threshold of acceptable variability across all NAMs based on laboratory mammalian studies.**
- **EPA should prioritize increasing external validity through broader coverage of biological variability.**

How to Bridge Differing Contexts for Evaluating Scientific Confidence?

Systematic review-based approaches are considered current best practices for establishing scientific confidence for hazard and dose-response.

- Designed to **evaluate existing evidence** (human epidemiology or laboratory mammalian toxicity studies), based on a large literature and NASEM report guidance.

In contrast, *scientific confidence frameworks for NAMs* have focused on design of assays and strategies.

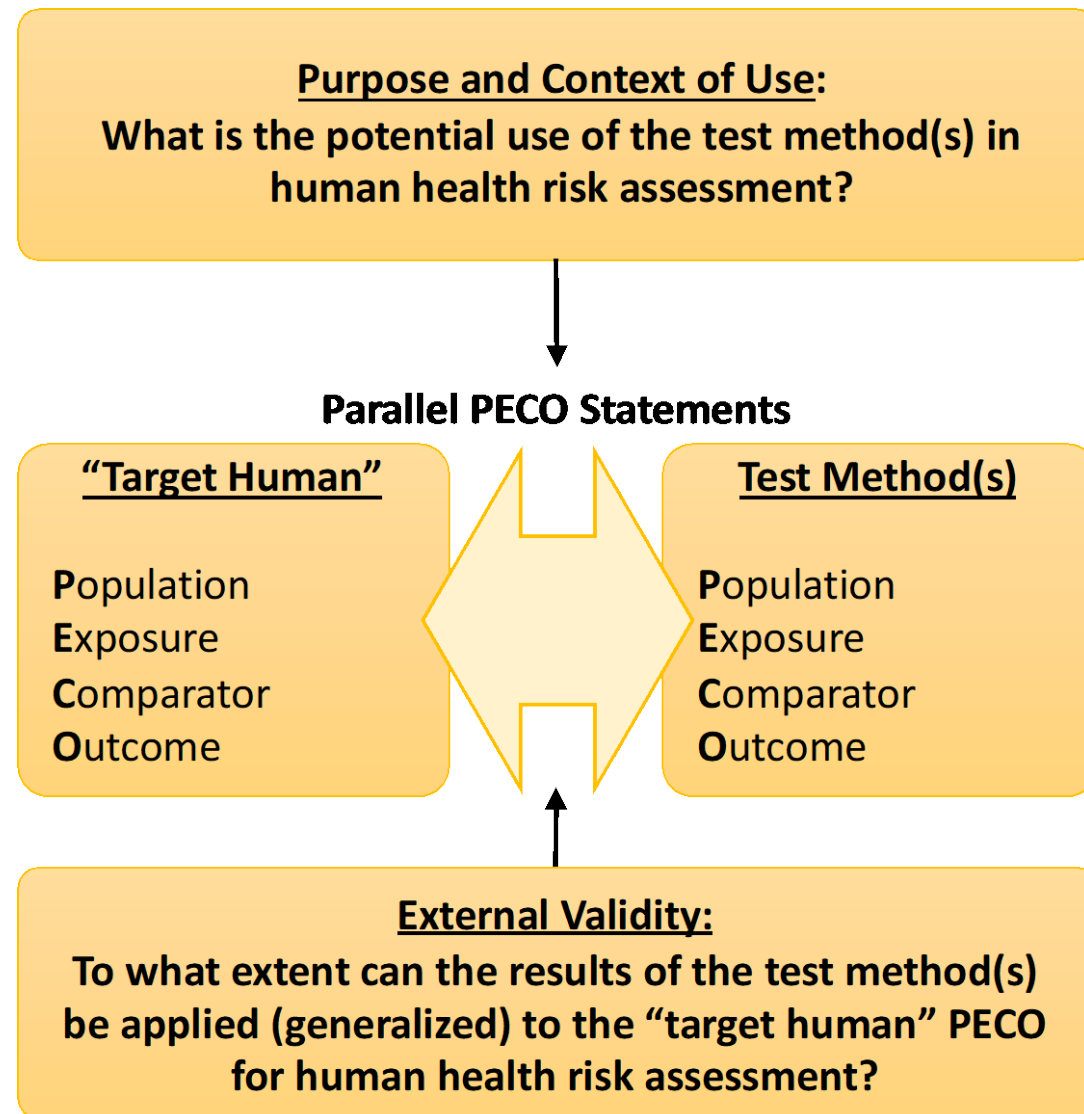
- Goal is to determine if a NAM **will generate acceptable data for use** in hazard identification and dose-response, based on large literature on concepts related to assay validation.

The committee aimed to integrate and bridge these different contexts, to enable a seamless handoff between them.

Bridging Different Contexts via “Parallel” PECO Statements

PECO¹: A cornerstone of evidence-based practice to frame and answer a human health hazard-related question. Laboratory mammalian toxicity tests are intended as surrogates for a “target human” PECO for the same tissue or system. *However, PECO statements are not currently routinely used for in silico, in vitro, and nonmammalian toxicity tests.*

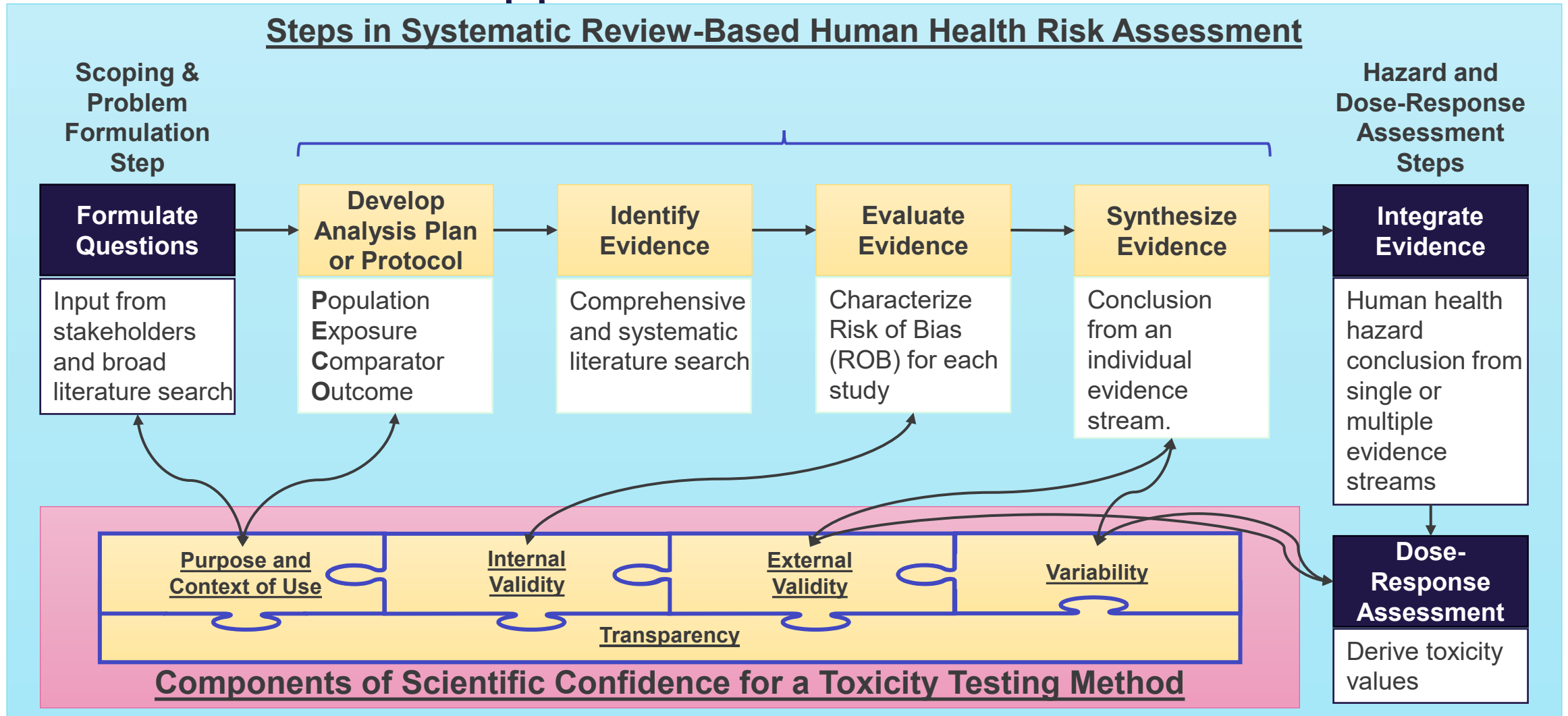
The report recommends that EPA address this gap by defining a “target human” PECO for each NAM, thereby providing information as to how it would inform human health hazard identification or dose-response.



Examples of “Parallel” PECO

Target Human PECO	Toxicity Testing Method	Test Method PECO
<p>P: Human population</p> <p>E: Chronic oral exposure to chemical</p> <p>C: No/lower exposure</p> <p>O: Any cancer</p>	<p>Two-year cancer rodent bioassay for chemical X in drinking water</p>	<p>P: Rodents</p> <p>E: Chemical in drinking water for 2 years</p> <p>C: Drinking water without X</p> <p>O: Any cancer</p>
<p>P: Human population</p> <p>E: Internal (serum) exposure to Chemical X via any route</p> <p>C: No/lower internal exposure</p> <p>O: Long QTc, positive or negative chronotropy, asystole</p>	<p>High throughput screening for chemical X using iPSC-derived cardiomyocytes</p>	<p>P: iPSC-derived cardiomyocyte from single or multiple donors</p> <p>E: Chemical dissolved in media with DMSO</p> <p>C: Negative controls: DMSO in media; Positive controls: known positive drugs for each outcome (e.g., sotalol, isoproterenol, propranolol)</p> <p>O: Delayed action potential, increased or decreased spontaneous beat rate, asystole</p>
<p>P: Human population</p> <p>E: Chemical X via any route</p> <p>C: No/lower exposure</p> <p>O: Adverse developmental outcomes</p>	<p>Zebrafish-derived early life stage chemical screening</p>	<p>P: Diverse strains of early life stages zebrafish</p> <p>E: Chemical X dissolved in media</p> <p>C: Negative controls: DMSO in media; Positive controls: known positive developmentally active controls</p> <p>O: Lethality, developmental delay, altered morphology, altered motor responses</p>
<p>P: Human population</p> <p>E: Chemical X via dermal exposure</p> <p>C: No/lower exposure</p> <p>O: Skin allergy/sensitization</p>	<p>Murine Local Lymph Node Assay (OECD TG 429/442A/442B)</p>	<p>P: Mice (adult female CBA/JNCrlj strain)</p> <p>E: Chemical dermally applied in vehicle (e.g. acetone: olive oil)</p> <p>C: Negative: Vehicle (acetone: olive oil) Positive: 25% hexyl cinnamic aldehyde in acetone: olive oil.</p> <p>O: Proliferation of lymphocytes in the lymph nodes draining the site of substance application.</p>

Building Scientific Confidence in NAMs for Human Health Risk Assessment Applications

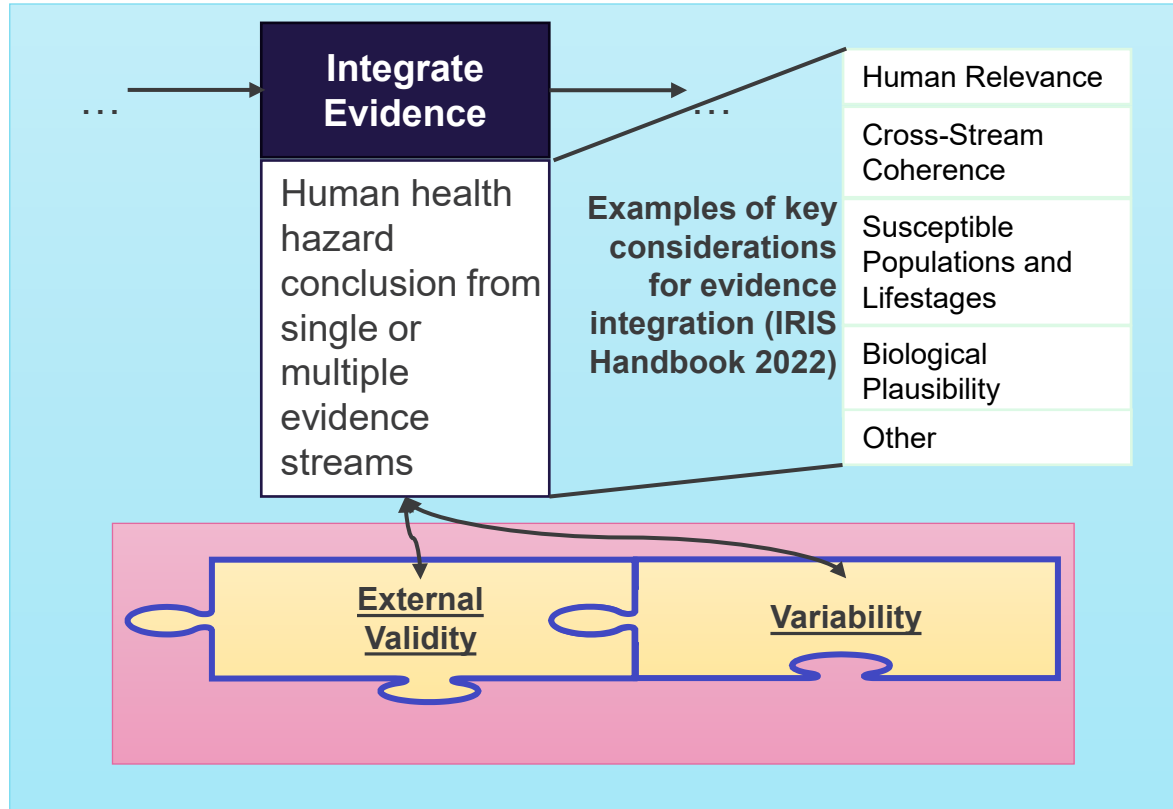


Example Approach for Structured Evaluation of External Validity: *Predicting Human Arrhythmic Cardiotoxicity Using iPSC-Derived Cardiomyocytes*

External validity domain	Qualitative considerations	Quantitative considerations
Biological considerations: Population - How strong is the biological basis for the test method as a biologically relevant model for the human population?	Moderate: Human iPSC-cardiomyocytes (hiPSC-CM) spontaneously beat and have morphology and gene expression similar to human left ventricular cardiomyocytes. However, they express a more fetal phenotype and are not paced.	Moderate: If hiPSC-CM from only a single donor is used, so do not address factors such as sensitive or vulnerable subpopulations.
Biological considerations: Outcome - How strong is the biological basis for the test method outcome as a model for human outcomes measured?	High: In vivo transporters and receptors involved in cardiomyocyte function (e.g., hERG channel, beta1 and beta2- adrenergic receptors) are expressed in hiPSC-CM. Beating parameters measured in vitro are similar to those measured in vivo using an electrocardiogram.	Moderate: Spontaneous beating is at a slower rate than in vivo paced beating, so outcome measurements may need to be corrected for these differences.
Exposure considerations: How accurately does exposure in the test method model human exposures?	Moderate: Chemical must be direct acting due to lack of metabolic capacity in hiPSC-CM, and soluble in DMSO.	Moderate: Quantitative uncertainties regarding differences in protein binding between media and serum, binding to testing materials (e.g., plastic), and partitioning. No consideration of background exposures.
Concordance: How accurately does the test method predict human outcomes to exposure?	Moderate: Accurate bioactivity predictions for a large number of positive and negative reference drugs for each outcome, especially QTc prolongation, but not environmental chemicals. Patient-specific iPSC-CM accurately predict susceptibility to doxorubicin-induced cardiotoxicity. However, these were not the result of systematic reviews.	Moderate: For QTc prolongation, the in vitro free concentration EC01 predicts within 3-fold the in vivo free blood concentration EC01 reported in clinical trials for 10 positive reference drugs. However, these were not the result of a systematic review and only include drugs (Blanchette et al. 2019).

Table 5-5

Building Scientific Confidence in NAMs: *Evidence Integration*

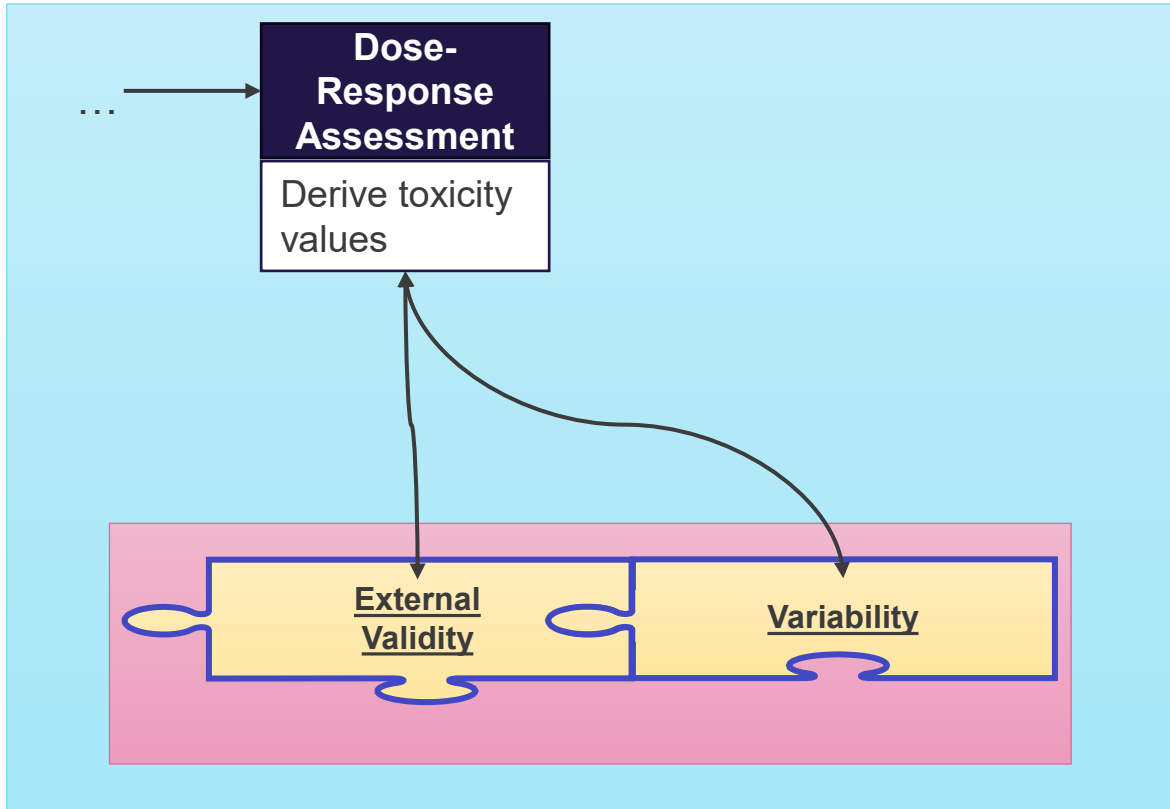


Components of scientific confidence for NAMs **map directly** to several considerations in evidence integration.

Scientific Confidence Domain	Corresponding Evidence Integration Consideration(s)
<ul style="list-style-type: none"> External Validity 	<ul style="list-style-type: none"> Human Relevance Cross-Stream Coherence Biological Plausibility
<ul style="list-style-type: none"> Biological and Experimental Variability 	<ul style="list-style-type: none"> Susceptible Populations and Lifestages

Note: Care needs to be taken to not “double count” aspects of scientific confidence in both Evidence Synthesis and Evidence Integration (especially External Validity).

Building Scientific Confidence in NAMs: *Dose-Response*



Key considerations for deriving toxicity values from NAMs:

- **External Validity** informs necessary **quantitative** adjustments from experimental systems to humans.
- **Experimental and Biological Variability** provides insight into uncertainty and variability in susceptibility **within the human population**.

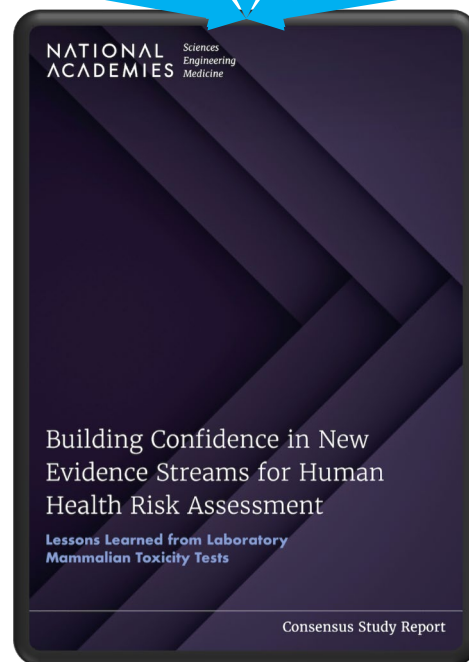
Building Scientific Confidence in NAMs for Human Health Risk Assessment Applications



Déjà vu all over again...



- Decreasing reliance solely on apical endpoints and “guideline” studies
- Increasing use of *in vitro* and computational approaches
- Increasing role of systematic review-based evidence assessment methods
- Increasing coverage of susceptible and vulnerable populations



- Broaden definition of “NAM”
- Build bridge between design of NAMs assays and testing strategies and use of NAMs in hazard and dose-response
 - Parallel PECO to interface NAMs with systematic review methods
 - Structured evaluation of NAMs internal and external validity
 - Evidence synthesis and integration with NAMs
 - Prioritize assays increasing coverage of biological variability

Goals of Committee Recommendations



Provide a path that builds confidence in NAMs data and approaches, from start to finish.



Prepare for a future when NAMs may be the sole basis for human health risk assessment and risk management decisions.



Ultimately, address many long-standing risk assessment challenges – from lack of data for most chemicals to better coverage of susceptible and vulnerable populations – and thereby better protect public health.

Extra Slides

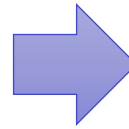
Committee's Statement of Task

- Review the variability and relevance of existing laboratory mammalian toxicity tests for **human health risk assessment** to inform the **development of approaches for validation and establishing scientific confidence in using New Approach Methods (NAMs)** and provide recommendations on expectations associated with NAMs when they cannot be compared with human studies.
- The work of the study committee will be informed by an initial public workshop organized by a subgroup of the committee, by a literature review that addresses the variability and human relevance of current laboratory mammalian toxicity tests and approaches to validation and establishing scientific confidence in using NAMs, and by public information gathering meetings organized by the study committee.

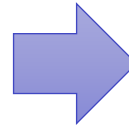


Summary Responses to Charge Questions

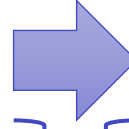
1. Does the committee assess the literature review and data provided as reflecting a comprehensive, workable, objective, and transparent process?
2. Given the results of the literature review and workshops, what are the implications of the qualitative and quantitative variability of laboratory mammalian toxicity studies when using them to establish the performance of NAMs?
3. What do the literature review and workshops indicate about concordance between laboratory mammalian models and humans in the adverse effects following chemical exposure and how might this frame expectations of NAMs when they cannot be compared directly with human studies?
4. The Committee shall impart expert advice on addressing the two related issues that were left unresolved in the 2017 NRC report:
 - a. Evaluation of the validity of assays that are not intended as one-to-one replacements for in vivo toxicity assays; and
 - b. Assessment of the concordance of data from assays that use cells or proteins of human origin with toxicity data that are virtually all derived from animal models.
5. Based on the conclusions from 1 – 4 above, how may the Committee foresee this information being incorporated into a new or the existing validation paradigm or scientific confidence framework so that EPA can ensure that NAMs are equivalent to or better than the animal tests replaced?



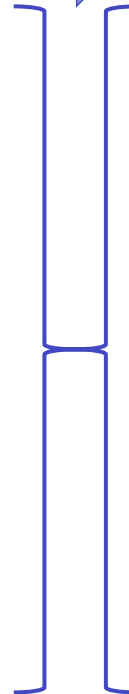
- Limited number of higher quality systematic reviews on either variability or concordance, based on comprehensive and transparent process



- Variability is not a fundamentally negative attribute
- Thresholds / benchmarking NAMs based on variability of mammalian studies **not** recommended

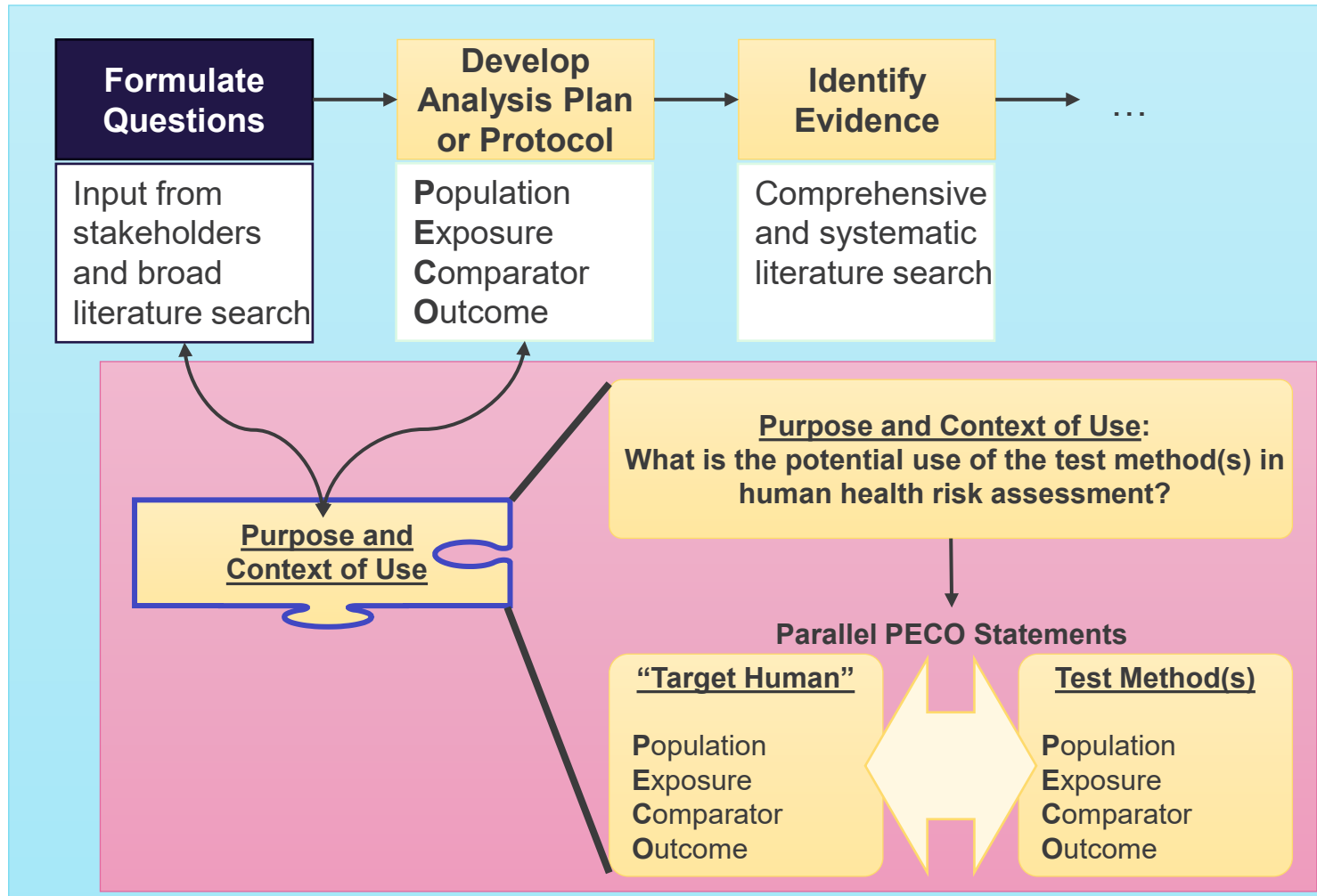


- Evaluate concordance with existing or new systematic and authoritative reviews



- **Key components of scientific confidence:**
 1. **Intended purpose and context of use** (“Parallel PECO”)
 2. **Internal validity** (Risk of Bias)
 3. **External validity** (Biological [P and O], Exposure, Concordance)
 4. **Biological and experimental variability**
 5. **Transparency**
- **Public health-protective framework needed for seamless handoff between**
 - **Scientific confidence evaluation of assay design and NAM-based testing strategies**
 - **Systematic review-based evaluation of NAMs data in human health hazard or risk assessment for particular chemical(s)**

How to Incorporate NAMs in *Scoping and Problem Formulation*?

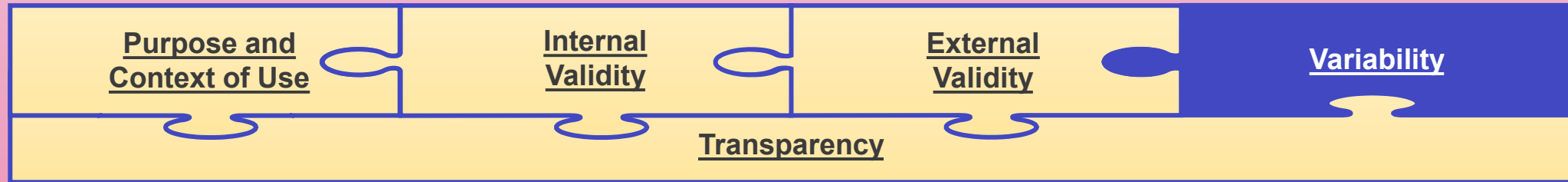


The use of “**parallel PECO**” statements as part of the **Purpose and Context of Use** of a NAMs provides a way to directly incorporate a NAMs during the **Scoping and Problem Formulation** step.

Specifically, the “target human” PECO facilitates considering **NAMs as a “evidence stream” that can undergo systematic review.**

External Validity in Relation to *Variability*

Components of Scientific Confidence for a Toxicity Testing Method



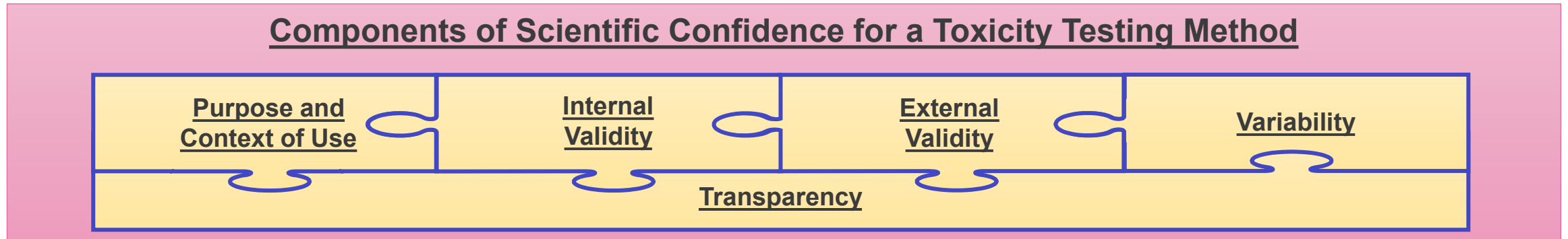
Biological and experimental variability

- **Biological variability** is defined as the true differences in attributes due to heterogeneity or diversity. Therefore, biological variability cannot be eliminated, but can be better characterized or controlled via rigorous experimental design.
- **Experimental variability** encompasses inter- and intra-laboratory variability, repeatability, and all aspects of reproducibility.

Recommendation 5.10: In its evaluation of test methods, EPA should prioritize increasing external validity (discussed above) through broader coverage of biological variability. One strategy that may be useful could be to use a battery of assays to encompass greater biological variability, while designing each assay so as to minimize experimental variability.

Recommendation 5.11: For any test method intended for use in risk assessment, whether in vivo, in vitro, in silico, or otherwise, particularly in a context where there are no other data (laboratory mammalian or human data), EPA's tolerance of variability should be driven by an analysis of the different levels and types of variability and of their impact on the test method's internal and external validity (discussed above). This analysis should also take into account the test method's purpose and context of use.

Building Scientific Confidence in NAMs from Start to Finish



Recommendation 5.12: EPA should establish the acceptability of NAMs-based testing strategies based on each specific purpose and context of use. EPA should be transparent as to the level of scientific confidence that results from examining the NAMs’s internal validity, external validity, and variability.

Recommendation 5.13: For the regulated community, the EPA’s goal should be to provide lists of acceptable NAM-based testing strategies under different purposes and contexts of use in order to establish confidence that NAM-derived data submissions to the agency will be integrated into decision-making (discussed in next section). This could be accomplished through EPA working with partners in the U.S. government and appropriate international organizations to develop a harmonized registry of toxicity testing methods documenting purpose and context of use (including parallel PECO statements), internal validity, external validity, and variability.