Building Confidence in New Evidence Streams for Human Health Risk Assessment: Lessons Learned from Laboratory Mammalian Toxicity Tests

The mission of the U.S. Environmental Protection Agency (EPA) is to protect human health and the environment. Some pollutants (e.g., lead, benzene, and ozone) are well studied, enabling EPA to characterize their human health hazards and risks. However, for many more chemicals in commerce and in the environment, such as per- and polyfluoroalkyl substances, there are little or no data on their potential health effects . As a result, only a small fraction of chemicals and other toxicants currently in the environment have undergone formal assessment of hazards and risks by the EPA.

Historically, the most influential data have come from laboratory mammalian studies, which still form the foundation of most human health assessments. Over the past decade it has become increasingly clear that relying primarily on laboratory mammalian studies limits the ability to assess the human health hazards and potential risk of the tens of thousands of chemicals to which people may be exposed. Further, there are concerns about costs, timeliness, animal welfare, and adequate characterization of the spectrum of effects observed in humans.

New approach methods (NAMs) in toxicology offer opportunities to surmount these limitations. For instance, NAMs could potentially inform timely decision-making when no data are available from laboratory mammalian toxicity tests or epidemiological studies. NAMs may also characterize subtle health perturbations, better encompass genetic diversity, and account for nonchemical stressors, informing better protection of susceptible and vulnerable populations. As a result, multiple efforts both in the United States and internationally have sought to develop NAMs for risk assessment purposes.

Although the promise and need for NAMs is clear, many barriers to their use remain. Despite recommendations that date to 2007 (*Toxicity Testing in the 21st Century: A Vision and a Strategy*, NASEM 2007), few concrete examples exist today of NAM data applications in hazard or risk assessment decisions by EPA or other authoritative bodies. This report aims to bridge this notable gap between the potential of NAMs and their practical application in human health risk assessment. It draws lessons learned from laboratory mammalian toxicity tests to help inform approaches for building scientific confidence in NAMs and for incorporating such data into risk assessment and decision-making. Overall, the recommendations aim to ensure a seamless handoff from the evaluation of NAM-based testing strategies in the laboratory to the incorporation of NAM data into modern, systematic-review-based risk assessments.

NEED FOR EPA TO BROADEN ITS DEFINITION OF NAMs

EPA's definition of "new approach methods" (NAMs) is "any technology, methodology, approach, or combination that can provide information on chemical hazard and risk assessment to avoid the use of animal testing". The goal to "avoid the use of animal testing" is inappropriate given that NAMs encompass studies in integrated in vivo systems, including in nonmammalian species such as zebrafish, and that some NAMs studies require whole animals as the source of cells. The report recommends that EPA broaden the definition of NAM to encompass the full range of strategies and approaches shown in Figure 1, all of which can be informative for human health risk assessment.

VARIABILITY AND CONCORDANCE

A key charge to the report's authoring committee was to review the scientific literature and conduct information-gathering sessions, including two public workshops, related to variability and concordance of mammalian toxicity tests, which have been the primary basis for most existing chemical assessments. The committee noted that minimizing variability—which refers to differences in attributes within and across studies —may limit the understanding of the distribution of toxic response, and therefore the generalizability of a study's results. Thus, the report recommends that EPA generally refrain from identifying a threshold of acceptable variability across all NAMs based on laboratory mammalian studies.

For concordance of adverse health effects—which refers to the similarity in toxic responses across species--the evidence the committee reviewed showed laboratory mammalian toxicity tests can generally identify human health hazards for a range of adverse health outcomes, but not necessarily the extent of response at a given dose. However, laboratory mammalian toxicity testing has been less successful for complex health endpoints, such as developmental neurotoxicity and mammary gland effects. This is due in part to the lack of alignment of methods and endpoints across experimental species and humans. The report provides guidance to EPA on how to evaluate concordance but cautions against using laboratory mammalian toxicity tests as the sole factor for determining the acceptance of NAMs.

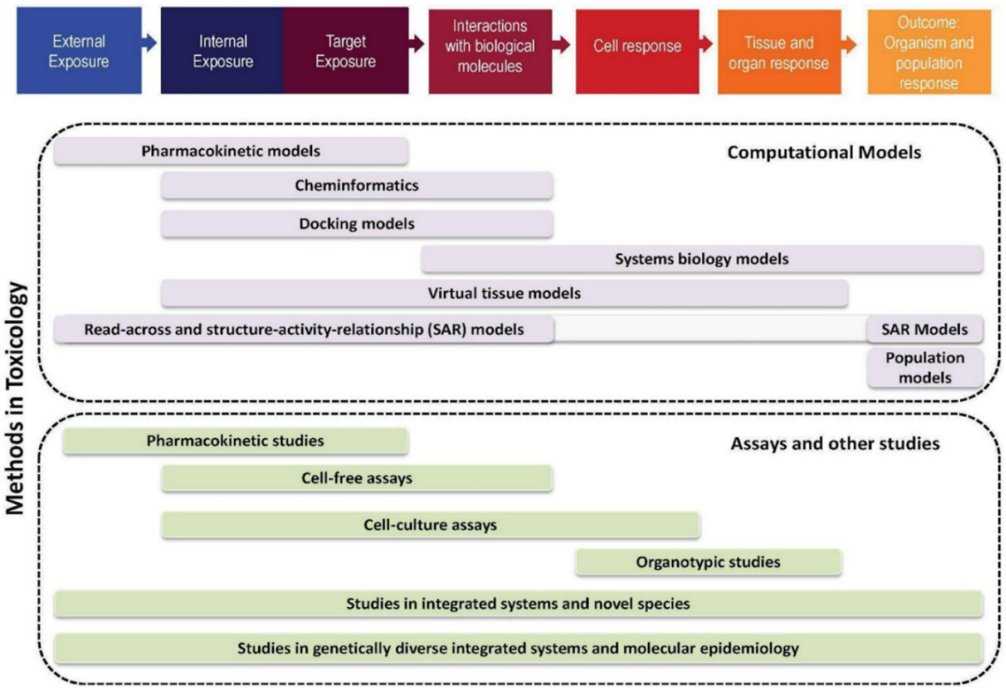

FIGURE 1: A continuum of computational models and biological assays can provide information for human health risk assessment and are relevant to considerations for building scientific confidence in NAMs.

SOURCE: NASEM 2017

BRIDGING DIFFERENT CONTEXTS FOR EVALUATING SCIENTIFIC CONFIDENCE

Structured, systematic-review approaches are considered best practices for evaluating data for human health hazard identification and dose-response. However, these methods were designed to evaluate existing evidence, and thus have generally focused on data from human epidemiology and laboratory mammalian toxicity studies. In contrast, scientific confidence frameworks for NAMs have focused on evaluation of assay design and with the goal of determining if a NAM will generate data acceptable for use in hazard identification and dose-response assessment.

The committee aimed to integrate and bridge these different contexts, to enable a seamless handoff between them. To do so, the committee identified key components of a scientific confidence framework for NAMs, listed in Box 2 and mapped them to established approaches for systematic review (Figure 2).

DEFINING THE INTENDED PURPOSE AND CONTEXT OF USE OF A NAM

A key aspect of the committee's recommendations to build a bridge from NAMs to application in human health risk assessment involves use of a population, exposure, comparator, and outcomes (PECO) statement. A cornerstone of systematic review approaches, the PECO statement clarifies the question being addressed and promotes transparency. For instance, laboratory mammalian toxicity tests are generally intended as surrogates for a corresponding "target human" PECO for the same biological tissue or system. However, PECO statements are not currently routinely used for in silico, in vitro, and nonmammalian toxicity tests. This limits their direct applicability in risk assessments. 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.

BOX 2. KEY COMPONENTS OF SCIENTIFIC CONFIDENCE FRAMEWORK FOR NAMs

1. **Intended purpose and context of use** relates to the specific question being addressed, commonly framed as a population, exposure, comparator, and outcome (PECO) statement.
2. **Internal validity** relates to the extent to which systematic error (bias) can influence the extent to which a study answers its research question correctly.
3. **External validity** refers to whether the study is addressing the relevant research question and the extent to which results from a study can be applied (generalized) to other situations, groups, or contexts.
4. **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.
5. **Transparency** refers to there being adequate information available in order to fully evaluate (1)–(4). Transparency may be more challenging for proprietary assays where some information may be kept confidential.

BUILDING CONFIDENCE IN NAMs FOR HUMAN HEALTH RISK ASSESSMENTS

Figure 2 illustrates the interface of concepts related to evaluating scientific confidence of NAM-based testing strategies with the steps in systematic review-based human health hazard assessment. The committee provides recommendations, along with examples, for each component or step in both processes, emphasizing opportunities for collaboration with the National Toxicology Program and other organizations. Going forward, EPA should develop and utilize a framework for hazard identification and deriving toxicity values protective of public health that does not require human epidemiologic or laboratory mammalian toxicity data. In so doing, the EPA should continue to follow previous NASEM recommendations related to systematic review and risk assessment. Overall, this approach will ensure a seamless handoff between development of NAM-based testing strategies and evaluation of scientific confidence of NAM data for individual chemicals.

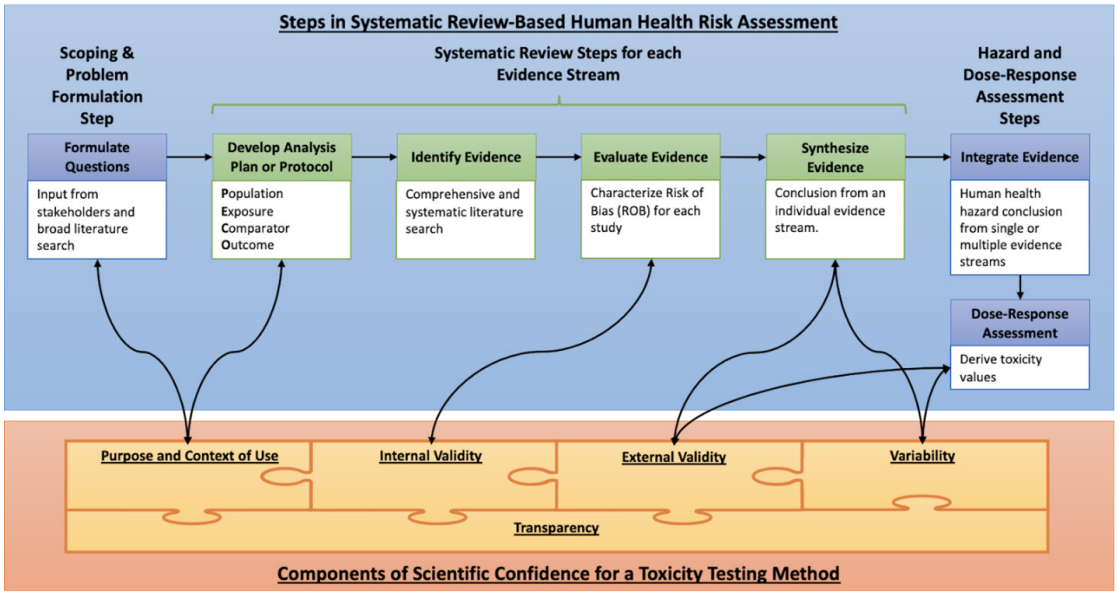

FIGURE 2: Interface between components of scientific confidence for a toxicity testing method and human health hazard and risk assessment.

SOURCE: Adapted from NASEM 2021

COMMITTEE

Weihsueh A. Chiu (Chair), Texas A&M University; **Kim Boekelheide,** Brown University School of Medicine; **Patience Browne,** OECD; **Holly Davies,** Washington State Department of Health; **Corie A. Ellison,** Procter and Gamble; **Marie C. Fortin**, Jazz Pharmaceuticals; **Nicole Kleinstreuer,** NIEHS; **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.

STUDY 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; **Corrine Lutz,** Senior Program Officer; **Thomasina Lyles,** Senior Program Assistant.

This Consensus Study Report Highlights was prepared by the Board on Environmental Studies and Toxicology based on the Consensus Study Report *Building Confidence in New Evidence Streams for Human Health Risk Assessment: Lessons Learned from Laboratory Mammalian Toxicity Tests (2023).*

The study was sponsored by the EPA. Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

This Consensus Study Report is available from the National Academies Press (800) 624-6242 | http://www.nap.edu | http://www.nationalacademies.org

To read the full report, please visit http://www.nationalacademies.org/BEST

Division on Earth and Life Studies

NATIONAL ACADEMIES *Sciences Engineering Medicine*