BOSC
BOARD OF SCIENTIFIC COUNSELORS

February 28, 2023

Christopher Frey, Ph.D.
Assistant Administrator
Office of Research and Development
U.S. Environmental Protection Agency

Dear Dr. Frey:

On behalf of the Executive Committee of the Board of Scientific Counselors (BOSC) we are pleased to provide you a review report addressing five charge questions posed after a review of the EPA's draft entitled, "The New Chemicals Collaborative Research Program: Modernizing the Process and Bringing Innovative Science to Evaluate New Chemicals Under TSCA" ("White Paper").

The New Chemicals Collaborative Research Program (NCCRP) is a joint activity of EPA's Office of Research and Development (ORD) and the Office of Pollution Prevention and Toxics (OPPT) to be an integrative research program within the Agency's 2023-2026 Chemical Safety for Sustainability Strategic Research Action Plan. The panel met in October 2022 culminating in an Executive Committee meeting in December 2022. This report represents the cumulative effort of the NCCRP review panel and the Executive Committee.

We anticipate that this report will assist ORD in developing and applying innovative approaches to address the requirements of the Toxic Substances Control Act (TSCA) for the review of new chemicals. We will be happy to provide any additional information concerning the review or answers to any questions you may have, and we look forward to working with you in the future on these programs.

Sincerely,

Paul Gilman, Ph.D.
Chair, BOSC

Lucinda Johnson, Ph.D.
Vice Chair, BOSC

BOSC

BOARD OF SCIENTIFIC COUNSELORS

REPORT OF THE U.S. ENVIRONMENTAL PROTECTION AGENCY BOARD OF SCIENTIFIC COUNSELORS NEW CHEMICALS COLLABORATIVE RESEARCH PROGRAM (NCCRP) PANEL

RESPONSES TO CHARGE QUESTIONS

BOSC Executive Committee

Paul Gilman, Ph.D. (Chair)
Covanta

Lucinda Johnson, Ph.D. (Vice-Chair)
University of Minnesota Duluth

Tracey Woodruff, Ph.D., MPH {recused}
University of California, San Francisco

Bart Croes, PE, MS
California Air Resources Board (retired)

Gilbert Gee, Ph.D.
University of California, Los Angeles

Jay Golden, Ph.D.
Syracuse Univrsity

Katharine Jacobs, MLA
University of Arizona

Daland Juberg, Ph.D., MS
Juberg Toxicology Consulting LLC

Rainer Lohman, Ph.D.
The University of Rhode Island

Jaime Madrigano, Sc.D., MPH
Johns Hopkins University

Pamela McElwee, Ph.D.
Rutgers University

Jayne Morrow, Ph.D.
National Institute of Standards and Technology

Olga Naidenko, Ph.D.
Environmental Working Group

Barret Ristroph, JD, Ph.D.
Ristroph Law, Planning, and Research

Mahmoud Saleh, Ph.D.
Texas Southern University

Derek Shendell, D.Env., MPH
Rutgers University School of Public Health

Justin Teeguarden, Ph.D.
Pacific Northwest National Laboratory

Dana Tulis, MEVE
U.S. Coast Guard

Stephen Weisberg, Ph.D.
Southern California Coastal Water Research Project Authority

Kevin Teichman, Ph.D.
Georgetown University

John White, Ph.D.
Louisiana State University

George Thurston, Sc.D.
New York University School of Medicine

Laureen Monica Boles, MCP
E4 Progress Planning and Engineering

Crystal Upperman, Ph.D., MPA
Deloitte

G. Allen Burton, Ph.D.
University of Michigan

Richard Becker, Ph.D.
American Chemistry Council Long-Range Research Initiative

Michelle Crimi, Ph.D., MS
Clarkson University

Ellen Mantus, Ph.D.
Health Effects Institute

BOSC New Chemicals Collaborative Research Program Panel

Hilary Adragna, Ph.D. *W.L. Gore & Associates*	Jennifer McPartland, Ph.D. *Beautycounter*	Holger Behrsing, Ph.D. *Institute for In Vitro Sciences, Inc.*
Richard Becker, Ph.D. (Panel Co-Chair) *American Chemistry Council Long-Range Research Initiative*	Gina Solomon, M.D., M.P.H. *Public Health Institute*	Sue Marty, Ph.D. *Dow, Inc.*
George Cobb, Ph.D. *Baylor University*	Cheryl Murphy, Ph.D. *Michigan State University*	Justin Colacino, Ph.D. *University of Michigan School of Public Health*
Justin Teeguarden, Ph.D. (Panel Co-Chair) *Pacific Northwest National Laboratory*	Mark Nelms *RTI*	Kristie Sullivan, M.P.H. *Physicians Committee for Responsible Medicine*
Marissa Smith, Ph.D. *WA Dept. Ecology*	Thomas Luechtefeld *Insilica*	

EPA Contact

Tom Tracy, Designated Federal Officer

October 24-25, 2022

CONTENTS

LIST OF ACRONYMS

ACE	Air, Climate, and Energy Subcommittee
AI	Artificial Intelligence
AOP	Adverse Outcome Pathways
API	Application Programming Interface
BOSC	Board of Scientific Counselors
CBI	Confidential Business Information
CDC	Centers for Disease Control and Prevention
DMSO	dimethyl sulfoxide
DOE	Department of Energy
DRP	detailed review paper
ECHA	The European Chemicals Agency
EJSCREEN	Environmental Justice Screening and Mapping Tool
EPA	Environmental Protection Agency
eSTAR	Emerging Systems Toxicology for Assessment of Risk
FACA	Federal Advisory Committee Act
FAERS	FDA Adverse Event Reporting System
FDA	Food and Drug Administration
GC	Gas Chromatography
GEO	Gene Expression Omnibus
GIVIMP	Good *In Vitro* Method Practices
HTTK	High-Throughput Toxicokinetics
HTTr	High-Throughput Transcriptomics
HTPP	High-Throughput Phenotypic Profiling
IATA	Integrated Approaches to Testing and Assessment
ICE	Integrated Chemical Environment
ITRC	Interstate Technology and Regulatory Council
IUCLID	International Uniform Chemical Information Database
IVIVE	*in vitro* to *in vivo* extrapolation
LC	Liquid Chromatography
MS	Mass Spectrometry
NAM	New Approach Methods
NAS	National Academy of Sciences
NCCRP	New Chemicals Collaborative Research Program
NICEATM	NTP Interagency Center for the Evaluation of Alternative Toxicological Methods
NIH	National Institute for Health
NSF	National Science Foundation
OECD	Organization for Economic Cooperation and Development
OPPT	Office of Pollution Prevention and Toxics
ORD	Office of Research and Development
PFAS	Per- and polyfluoroalkyl substances

QSAR	Quantitative structure-activity relationship
(Q)SAR	A collective term signifying QSARs and SARs collectively
QSUR	Quantitative Structure Use Relationships
RACT	Research Area Coordination Teams
SAR	Structure-activity relationship
SMARTS	Simplified Molecular-input line-entry system Arbitrary Target specification
TSCA	Toxic Substances Control Act
UVCB	Unknown or variable composition, complex reaction products, or biological materials

INTRODUCTION

The New Chemicals Collaborative Research Program (NCCRP) is a joint activity of EPA's Office of Research and Development (ORD) and the Office of Pollution Prevention and Toxics (OPPT)[1] to develop and apply innovative approaches to address the requirements of the Toxic Substances Control Act (TSCA) for the review of new chemicals. TSCA requires EPA to review all new chemical substances (i.e., those not yet in commerce) to make determinations regarding potential risks to human health and the environment before manufacturing can commence. With hundreds of new chemical notices submitted to OPPT per year and limited hazard and exposure information, addressing these statutory requirements with sound science, transparency, and consistency, while meeting tight statutory deadlines for decisions, requires continued evolution of scientific methods, approaches, and tools. Bringing innovative science to modernize the new chemicals evaluation procedures will help overcome information gaps and help OPPT meet TSCA statutory requirements in a timely, effective, and efficient manner. While the NCCRP is focused on developing methods for the TSCA new chemicals program, many, if not all, of these methods will likely also prove useful to help fill both hazard and exposure data and information needs for existing TSCA chemicals.

The NCCRP was announced in February 2022. A public meeting followed in April 2022. The NCCRP has been designed by ORD and OPPT to be an integrative research program within the Agency's 2023-2026 Chemical Safety for Sustainability Strategic Research Action Plan.[2] The NCCRP is described in detail in the October 2022 report from EPA to the BOSC entitled, "The New Chemicals Collaborative Research Program: Modernizing the Process and Bringing Innovative Science to Evaluate New Chemicals Under TSCA."[3] The research program described in this EPA Report is the focus of this review by the BOSC.

The NCCRP is a focused research program which reflects the translation and extension of successful developments that have emerged from 15 years of computational toxicology research conducted by ORD staff and partners. The NCCRP actualizes the vision and objectives of the CompTox Blueprint[4] and EPA's New Approach Methods (NAMs) Work Plan[5] by developing NAMs to provide data and information needs for OPPT's new chemicals program. Importantly, the research conducted in the NCCRP will also contribute to establishing the requisite degree of scientific confidence needed for these methods to be used in regulatory decision making in OPPT. Through the NCCRP, ORD is working with the OPPT to advance five key Research Areas:

1. Updating and refining chemical category formation approaches and improving read-across inference methods.

2. Developing and expanding databases containing TSCA chemical information.

3. Developing and refining predictive models for physicochemical properties, environmental fate/transport, hazard, exposure, and toxicokinetics;

.....

[1] OPPT is a division of the Office of Chemical Safety and Pollution Prevention (OCSPP); OPPT is the program office in EPA that administers TSCA

[2] https://www.epa.gov/system/files/documents/2022-10/CSS%20FY23-26%20StRAP_EPA-ORD_October%202022_508.pdf

[3] BOSC Review Draft, October 2022. https://www.epa.gov/system/files/documents/2022-10/White_Paper_New%20Chemicals%20Collaborative%20Research%20Program_BOSC_Final_24Oct2022.pdf.

[4] Thomas et. al., 2019. The Next Generation Blueprint of Computational Toxicology at the U.S. Environmental Protection Agency. Toxicological Sciences, Volume 169, Issue 2, June 2019, Pages 317–332, https://academic.oup.com/toxsci/article/169/2/317/5369737.

[5] EPA New Approach Methods Work Plan, December 2021. https://www.epa.gov/system/files/documents/2021-11/nams-work-plan_11_15_21_508-tagged.pdf.

4. Integrating and applying *in vitro* new approach methodologies (NAMs) to biologically profile substances; and

5. Developing a TSCA new chemicals decision support tool that utilizes curated data and integrates lines of evidence across many chemical, computational, and biological profiling platforms.

The NCCRP is somewhat unique for ORD in that it has been designed, in collaboration with the OPPT, to explicitly focus on research and development of specific scientific tools and methods needed to modernize the approaches for evaluating chemicals in EPA's New Chemicals Program under TSCA. It is vital, therefore, that the NCCRP include Research Area Coordination Teams (RACTs)[6] comprised of ORD scientists and OPPT scientists. Such RACTs will ensure this applied research program is designed and conducted in a manner that will deliver the specific scientific work products needed by OPPT. In this same vein, from the outset, the NCCRP would benefit from incorporating technology transfer activities as an integral component of each research project. As noted in the NCCRP report to the BOSC, this focused research program has been specifically designed to address OPPT's regulatory needs and bolster ORD's efforts to develop NAMs."[77] Therefore, it's critical that the NCCRP research activities include actions to help integrate these modernized approaches into the toolbox of methods used by the OPPT and other end users for the evaluation of new chemicals. Accordingly, to meet this shared responsibility of ORD and OPPT, activities should be built into the NCCRP, such as education, training, and outreach to end users for each research tool or methodology, as appropriate.

The identified strengths, suggestions, and recommendations herein are informed by a review of the EPA's draft entitled, "The New Chemicals Collaborative Research Program: Modernizing the Process and Bringing Innovative Science to Evaluate New Chemicals Under TSCA" ("White Paper"), the EPA's presentations to the Committee, available scientific literature, and Committee members' experiences using a variety of NAM tools including those developed or used by the EPA.

In this report, Committee members provide specific Recommendations for priority actions by EPA as the Agency moves forward with implementing the NCCRP. These Recommendations should be of the highest priority. The Committee also provides numerous Suggestions. The Committee's judgement regarding the priority for these Suggestions and estimates of the level of effort for each Suggestion are also provided to aid decision making. However, these Suggestions are subordinate to the Recommendations. Accordingly, Suggestions should be viewed as information for EPA to take under consideration, whereas Recommendations should be viewed as activities that the Committee agreed reflected the most critical opportunities to improve the NCCRP or address important weaknesses in the NCCRP. These Recommendations would be incorporated into the NCCRP as it is refined and implemented by EPA

.....

[6] The RACT "..develops goals and objectives for the Output and establishes criteria for the work needed to accomplish it. ORD researchers propose research Products, which the RACT reviews and refines to ensure Products will meet the goals and objectives of the Output and reflect the timing and specific needs of [the] EPA program [OPPT's New Chemicals Program]…" Strategic Research Action Plan, Fiscal Years 2023-2026, Chemical Safety for Sustainability Research Program, EPA/600/R-22/238 | October 2022, https://www.epa.gov/system/files/documents/2022-10/CSS%20FY23-26%20StRAP_EPA-ORD_October%202022_508.pdf.

[7] The New Chemicals Collaborative Research Program: Modernizing the Process and Bringing Innovative Science to Evaluate New Chemicals Under TSCA; page 5. https://www.epa.gov/system/files/documents/2022-10/White_Paper_New%20Chemicals%20Collaborative%20Research%20Program_BOSC_Final_24Oct2022.pdf.

CHARGE QUESTIONS AND CONTEXT

The NCCRP Panel was charged with five questions as follows:

Q.1: As described in Research Area 1 of the accompanying White Paper (pages 16-20), planned research activities are focused on updating and refining the chemical categories and read across methods used by OPPT.

Please comment on whether there are other approaches or chemical characteristics that could be considered when developing the categories and analog identification methodologies.

Q.2: As described in Research Area 2 of the accompanying White Paper (pages 20-28), planned research activities are focused on expansion and further development of existing public databases in ORD containing chemistry, hazard, exposure, and toxicokinetic information relevant to TSCA chemicals.

Please comment on this effort, including in your feedback useful sources of chemical information that could be incorporated into the curation efforts.

Q.3: As described in Research Area 3 of the accompanying White Paper (pages 28-33), planned research activities are focused on developing, refining, and evaluating (Q)SAR and other predictive models for physical-chemical properties, environmental fate/transport, hazard, exposure, and toxicokinetics.

a. Please comment on the (Q)SAR and predictive modeling proposed, as well as the proposed informatics platform for management of input data and development and management of (Q)SAR and other predictive models. In your comments, please address whether there are additional (Q)SAR models, approaches, or other informatics platform features that could be considered.

b. Please comment on any additional features that could be considered in the evaluation of these models, applicability domain(s), and associated documentation.

Q.4: As described in Research Area 4 of the accompanying White Paper (pages 33-40), planned research activities are focused on developing and evaluating a suite of *in vitro* NAMs that could be used by external stakeholders for testing and data submissions under TSCA, as well as potentially informing and/or expanding new chemical categories.

Please comment on the initial screening strategy proposed. Please include in your comments, other assays and/or endpoints to consider for the research plan.

Q.5: In the Background of the accompanying White Paper (pages 5-16), information on challenges in new chemical assessment and the vision statement for the NCCRP are presented. The primary vision of the NCCRP is to modernize the process for evaluating new chemicals under TSCA by supporting the evolution of OPPT's use of new and existing methods, approaches, and tools using innovative science.

Research Areas may address the issues identified in the Background and vision statement. Please also include potential additional research areas for EPA to consider.

PANEL RESPONSES TO CHARGE QUESTIONS

Charge Question 1

Q.1. As described in Research Area 1 of the accompanying White Paper (pages 16-20), planned research activities are focused on updating and refining the chemical categories and read across methods used by OPPT.

Please comment on whether there are other approaches or chemical characteristics that could be considered when developing the categories and analog identification methodologies.

Narrative

ORD and OPPT are to be commended for including, as a critical pillar of the New Chemicals Collaborative Research Program, research focused on modernizing the methods used by OPPT to group chemicals into categories and the procedures to conduct read-across (i.e., inference prediction modeling to extrapolate data / information from a similar substance to the substance undergoing review). The OPPT new chemicals program currently relies heavily upon grouping of chemicals into categories and read-across to fulfill needs for data/information to evaluate potential hazards, exposures and risks of new chemical submissions. As we understand it, OPPT currently relies upon expert judgement procedures for grouping similar chemicals into a category (or sub-category) by applying an OPPT chemical similarity grouping guidance document[8] that was last updated in 2010. While expert scientific judgment has, in the past, often played a large role in many scientific interpretation processes, such practices can be problematic due to lack of transparency, difficulties in reproducibility, and concerns over subjectivity and bias. In addition, it is our understanding that the current toxicity inference approaches used by OPPT rely almost exclusively on extrapolating traditional toxicity testing data obtained from laboratory animal studies.

The diverse data streams proposed to support chemical clustering and rapid hazard assessment have tremendous promise to improve the ability to estimate toxicity over traditional methods, however, the use of these new technologies should be fit for the purpose of the assessment. While similarity in structure is one important attribute to evaluate when grouping chemicals into a category, structural alerts alone are not likely to be sufficient. Data from new lines of evidence, in particular approaches that use mechanistic NAM assays to explore similarities in biological response pathways (i.e., biological activity profiling), can provide critical information for grouping. Over the past 15 years, there have been considerable advances made in scientific understanding of biological pathways and how chemicals interact with biological systems (e.g., the Adverse Outcome Pathway (AOP) framework). This knowledge has been instrumental in enabling the development of advanced mechanistic assays (NAMs) and improved computational profiling methods. New methods for dosimetry, such as *in vitro* to *in vivo* extrapolation (IVIVE), and improvements in exposure science and exposure modeling have also been brought to the forefront during this time period. It's important for these efforts to evaluate, and incorporate into GenRA, as appropriate, the advances in read-across methodologies developed by other organizations or research programs, such as the Organisation for Economic Co-operation and Development (OECD)[9], the European Commissions' H2020 project EU-ToxRisk[10], Ambit[11], the UL Cheminformatics Tool Kit[12], etc. By working together on the New Chemicals Collaborative Research Program, ORD and OPPT can bring this knowledge and these methods forward to design and conduct the research needed to develop, evaluate, and establish scientific confidence in more objective,

[8] https://www.epa.gov/sites/default/files/2014-10/documents/ncp_chemical_categories_august_2010_version_0.pdf.

advanced and transparent approaches for grouping similar chemicals and inference modeling to address data / information needs.

Strengths

- Incorporation of a wide set of attributes: By using many different attributes and methods, the breadth of this research coupled to the systematic approach will improve the objectivity and transparency of the data and procedures used to inform chemical category formation and the basis for similarities for read-across.
- Integration of computational and biological profiles: This research area explicitly includes approaches to evaluate and integrate computational and biological activity profiles, toxicokinetics, metabolite formation, persistence, etc. This is expected to create a richer understanding of similarities and differences.
- Increasing transparency and reproducibility: The GenRA method is an online automated web application that is expected to 1) improve transparency and reproducibility in category formation and read-across, and 2) increase understanding and communication of uncertainties.
- Improving objectivity: The explicit procedures envisioned to be actualized in GenRA are expected to reduce subjective, expert judgment and unconscious / conscious bias.
- Improving chemical category determinations: Converting the structural information that underlies the existing new chemical categories (NCC) into a machine-readable form (e.g., SMARTS patterns) will help to make the process of reviewing whether a new chemical fits into an NCC more systematic, transparent, and reproducible, thereby improving confidence in the predictions.
- Increasing understanding of domains of applicability: Understanding how well the chemicals in the TSCA non-confidential list fit within the domain of applicability for the different (Q)SAR models ORD uses is important to help make determinations as to the suitability of the predictions. This could also help to guide NAM-based testing to expand the applicability domain of the (Q)SAR models and improve confidence in the predictions.

Suggestions

- The title of this research area should be changed from “Update and Refine Chemical Categories” to “Modernizing Chemical Categories and Improving Inference Modeling to Meet Data / Information Needs.” In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a low degree of effort by EPA.
- To build understanding and confidence in the new chemical grouping methods and modernized read-across methods, from the outset these research activities should keep the end users in mind. For example, ORD and OPPT should consider including outreach, education, and training to EPA staff and external stakeholders as specific activities and integrate these into the overall research program milestones. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- ORD and OPPT should consider exploring approaches for data integration and visualization to help document and communicate similarities and differences across compounds for all the attributes evaluated in GenRA. Examples include spider/radar plots and 3-D techniques – techniques that could

[9] OECD https://www.oecd.org/chemicalsafety/risk-assessment/iata/
[10] EU ToxRisk https://www.eu-toxrisk.eu/page/en/case-studies.php
[11] AMBIT http://cefic-lri.org/toolbox/ambit/
[12] UL https://www.ul.com/news/sniffing-out-hazardous-chemicals

facilitate side-by-side, or overlay, comparisons. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a low degree of effort by EPA.

- EPA should explore the potential to use of Quantitative Structure Use Relationships (QSURs) and advanced high-throughput exposure models to inform category formation, read-across, and screening level risk evaluations. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a medium degree of effort by EPA.
- Taking into consideration the New Chemicals Program approaches[13], including procedures for requiring new data and information, consideration should be given to designing the modernized New Chemicals evaluation procedures in a tiered manner, in which the first tiers utilize predictive *in silico* tools to quickly identify potential toxicity, group chemicals for read-across, predict potential exposures, and then additional information can be incorporated as necessary to build the weight of evidence to support read-across. This could include incorporation of approaches to efficiently predict approximate metabolite abundance, activation or breakdown to a reactive species, or detoxification to inform chemical grouping. Overall, this tiered approach should be designed to be adaptable to different exposure and use scenarios. For example, modernized clustering algorithms can be used to quickly identify analogues and support read-across using tools such as GenRA when the chemical of interest is within the domain of applicability of the models. However, for chemicals that do not lie well within the domain of applicability, addition of bioactivity data from the rapid screening assays and mechanistic biological pathway knowledge (e.g., AOPs) could improve hazard estimation. Structuring this process as a flexible, tiered approach should encourage assessors to focus on the best tools for the particular risk decision at hand. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a low degree of effort by EPA.
- To facilitate transparency and reproducibility, clear decision criteria need to be defined for each grouping or read-across tool. Explicit data interpretation procedures for model results and a structured decision analysis framework for determining when additional analysis or specific additional testing should be considered will be important for ensuring these new methods are used to their best effect. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a high degree of effort by EPA.
- The clusters for the TSCA active inventory should be periodically (e.g., every 4 years) updated based upon the availability of new or updated data/knowledge on the chemicals in the inventory or new methods to clustering. For example, if the model(s) that calculate physicochemical properties used to cluster chemicals is updated, it would be expected to be able to make predictions for an expanded set of chemicals**.** In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a high degree of effort by EPA.
- It will be important to clearly define the domain of applicability, as well as the areas of uncertainty, to ensure appropriate use of the new tools. Chemicals that are not likely to be well-addressed by a particular model should be clearly flagged, and explicit data interpretation procedures provided for alternative assessment approaches. It will be particularly important to address difficult-to-test substances and complex mixtures (e.g., UVCBs). These products are increasing in the marketplace and pose many challenging regulatory issues, such as naturally varying mixture contents, percentage content of compound to regulate, and interactive toxicant effects. The European Chemicals Agency (ECHA) and fragrance industry are currently developing guidelines. These issues are larger than a single agency or research program. Leveraging the broader regulatory science community through communities of practice and crowd-sourcing solutions may help facilitate improvement in these

[13] https://www.epa.gov/sites/default/files/2021-04/documents/new_chems_working_approach_-_12.20.19_final_with_disclaimer.pdf

areas. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a high degree of effort by EPA.

- While publication in peer reviewed journals can be a critical step to broader scientific acceptance of new and improved methods, the publication process can often delay public dissemination of EPA work products, which can slow down and unnecessarily impede acceptance and use. These delays need to be avoided. This can be accomplished by incorporating into the project design alternative methods for independent scientific engagement and/or peer review (see EPA's Peer Review Handbook and, e.g., SciPinion) that can be combined with stakeholder engagement. Examples include *ad hoc* presentations of interim work products and updated plans, periodically focused webinars, planned peer engagement on specific activities, dedicated sections of the annual EPA NAMs workshop, etc. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- The importance of metabolism and degradation materials/pathways in chemical toxicity must continue to be considered within prediction-based risk assessment approaches. We are aware of efforts within the EPA as well as the broader research community to begin to address this challenge. As the science progresses, opportunities should be explored to incorporate prediction of metabolites and degradation products. Given the nascent state of the science, significant resources are currently required for metabolite identification, abundance, and bioactivity determinations. Therefore, EPA should continue to monitor developments in this space and incorporate newer methods into read-across approaches when these applications are determined to be fit for purpose for OPPT's new chemicals program. In considering potential actions for this suggestion, the Committee suggests this is a low priority that would require a high degree of effort by EPA.

Recommendations

The Panel offers the following recommendations:

Recommendation 1.1: The Committee recommends ORD, in conjunction with OPPT, design, conduct, and publicly disseminate case studies evaluating the performance of the current OPPT categories compared to the new approaches, such as GenRA, to support a read-across assessment where analog toxicity data are compared to target chemical toxicity data that are initially blinded to the assessor. Case studies should include several situations (e.g., where an understanding of metabolism is critical for establishing suitable analogs, where bioactivity data are limited, where small changes in chemistry have the potential to have significant impact on toxicity). Given the widespread interest in read-across in the global regulatory science community, it would be informative for these EPA case studies to include comparisons to read-across methodologies developed by other organizations (cited above), and document, as appropriate, alignment with ECHA's Read Across Assessment Framework.[14] These case study activities will help document scientific confidence in the newer approaches and support transitions from the existing OPPT approaches to the newer read-across approaches (e.g., GenRA).

EPA Response:
Thank you for the comment. We agree that case studies are useful in advancing the science and building confidence in new approaches. Case studies are integral to the broader CSS research program and will continue to be for the work being done under the NCCRP collaboration. Specifically, additional case studies to build confidence in the use of GenRA in regulatory toxicology will be considered as part of

[14] https://echa.europa.eu/documents/10162/13628/raaf_en.pdf/614e5d61-891d-4154-8a47-87efebd1851a

ongoing and planned work as well as through collaborations with external partners at the European Chemicals Agency.

Recommendation 1.2: The Committee recommends ORD and OPPT explore the potential to use CBI data within the GenRA and other inference models for grouping and read-across. One option to explore would be using federated learning[15] with differential privacy data methods, or similar technologies, that allow the private data to be retained and protected locally while still enabling the data to be used in model development (for example, the Machine Learning Ledger Orchestration for Drug Discovery collaboration[16]). Another option to consider would be developing a protected in-house user downloadable app (e.g., like the OECD toolbox download) to enable data use while protecting CBI. This is a particularly important research activity that may improve approaches for new chemicals that fall outside the current domains of non-CBI databases.

EPA Response:

Through the NCCRP collaboration, ORD and OPPT will explore a wide variety of public sources and legacy OPPT TSCA files to support modeling efforts. Planned activities include identifying, extracting, curating and cataloging available data on chemistry, hazard, fate, and exposure from different TSCA databases and holdings, which may include TSCA CBI. As TSCA data management systems evolve, the NCCRP efforts will adapt to these changes, while continuing to leverage the TSCA databases and holdings. Safeguards for CBI will be maintained, as appropriate, throughout the curation and use of the data. In addition, as described in Research Area 5, OPPT and ORD will work collaboratively both to increase the amount of data previously submitted to OPPT for use within a CBI-protected environment, and to allow for the use of ORD developed modeling tools within a CBI-protected environment.

Recommendation 1.3: The Committee recommends that, in addition to having a RACT, ORD and OPPT should establish a process and schedule for jointly evaluating the scientific confidence and readiness of these NAMs for updating the new chemical grouping and read-across methods that are intended to be used by OPPT's new chemicals program. A set schedule is needed to ensure the review process is keeping pace with advances in science and knowledge, to focus the next round of research, and to provide the certainty needed for the Agency and stakeholders to implement these methodologies efficiently and confidently. This would also ensure predictability in the application of program guidance for a set time period. One schedule to consider is alignment with the StRAP cycle. For example, the schedule for this scientific confidence and readiness review could be sequenced to finish at a point in time where the results of the review and recommendations for additional research serve as input into development of the next StRAP.

EPA Response:

The idea of a mechanism to update the readiness of NAMs for new chemical evaluations such as updating chemical grouping and read-across methods is a valuable suggestion. EPA's strategy for establishing scientific confidence in NAMs and demonstrating application to regulatory decisions is described under

[15] "Federated learning (also known as collaborative learning) is a machine learning technique that trains an algorithm across multiple decentralized edge devices or servers holding local data samples, without exchanging them." https://en.wikipedia.org/wiki/Federated_learning
[16] https://www.melloddy.eu/general-assembly

Objective III of the EPA NAMs Work Plan. A multi-prong strategy includes characterizing the scientific quality and relevance of existing vertebrate animal tests, developing an EPA scientific confidence framework, developing recommended reporting requirements, and demonstrating the application of the NAMs to regulatory decisions through case studies. ORD and OPPT will continue to follow this approach and other strategies described in the EPA NAMs Work Plan as work progresses under the NCCRP. As the work described in the NAMs Work Plan and the work of the NCCRP collaboration are being conducted within the CSS StRAP, associated reviews to evaluate the science and ensure alignment with OPPT needs will be part of the ORD StRAP review cycle. EPA (OPPT and ORD) will plan for regularly scheduled joint "Work-In-Progress" discussions for awareness and adoption, as appropriate. In addition, we expect to host and/or present results regularly at scientific meetings and relevant stakeholder meetings.

Charge Question 2

Q.2. As described in Research Area 2 of the accompanying White Paper (pages 20-28), planned research activities are focused on expansion and further development of existing public databases in ORD containing chemistry, hazard, exposure, and toxicokinetic information relevant to TSCA chemicals.

Please comment on this effort, including in your feedback useful sources of chemical information that could be incorporated into the curation efforts.

Narrative

Data relevant to TSCA chemicals are available in a wide range of public sources along with legacy OPPT TSCA files. Many of these legacy TSCA data are not in a digital form that can be currently accessed. Moreover, data that exist in publicly available databases may not exist in a form where they are easily and reproducibly queried and integrated. There is also a vast amount of existing chemical information in peer reviewed and "gray" literature that is currently not easily accessible. To address these complex challenges, ORD and OPPT seek to develop and expand databases containing TSCA relevant information. Plans described in Research Area 2 include continued extraction and curation of existing data on physical-chemical properties, environmental fate, hazard, and exposure. Plans are also outlined to map information in existing ORD databases to standardized reporting templates, storing the linked information in an International Uniform Chemical Information Database (IUCLID). Developing robust and comprehensive databases that digitize and merge this existing information will be essential for rigorous predictive evaluation of new chemicals under TSCA. If successful, the proposed plan will enable the reproducible development and refining of (Q)SAR models, inform the development of new chemical categories, and provide readily accessible data for analogs in the read-across evaluation of new chemicals.

In general, the strategies laid out by NCCRP are robust and well thought out. Digitization of legacy OPPT TSCA data in a machine searchable format will enable these data (potentially including CBI information) to be incorporated in new chemical characterization in a transparent manner. By integrating existing databases on physicochemical properties and environmental fate properties, household product chemical composition and function, multimedia monitoring data, ecological hazard, human health hazards, and toxicokinetic data, OPPT will be able to leverage vast amounts of existing data, assisting EPA in their legislative mandate for timely new chemical evaluation. The development and integration of literature mining techniques will potentially allow for the incorporation of relevant chemical information

from the published and gray literature. We commend OPPT and ORD for their commitment to open-source reproducible science. As part of our suggestions, in the table below, we provide a list of databases not used by the EPA which may provide additional information relevant to toxicological evaluation. We additionally make suggestions related to best practices for data submission, curation, and harmonization. Finally, we make recommendations related to replication, quality control, and validation to improve or ensure reproducible and transferable methods.

Strengths

- Single source of truth: Standardization of database vocabularies to an intentionally compatible format will ease the use of data in more applications and create more transparency in the evidence used for downstream applications.
- Data source versioning: Versioning and storing of source databases will help to maintain their data as part of a larger data store and help guarantee the longevity of those data as well as the reproducibility of analyses of the data.
- Comprehensive set of databases identified: Proposed databases will capture relevant information on chemical identity and structure, physiochemical and fate properties, health hazard/toxicodynamics data, human exposure data, and toxicokinetics.

Suggestions

- It is important to consider the Data Life Cycle. Source databases will deprecate, lose support over time, or, possibly, be identified as having quality control issues. A protocol should exist to handle source data deprecation / removal. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a low degree of effort by EPA.
- Quality control approaches are critical. Care should be taken to create a tracking system to unambiguously associate source studies with aggregated report data to prevent data duplication and avoid impact on Weight of Evidence analyses. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- The Committee suggests EPA explore approaches to capture and link data provenance (i.e., history for each piece of data) in the databases it develops. When models are created from the constructed data store, it should be possible to reference which source data were used to construct the model. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- A large list of life science databases with open-source scripts to extract and build versioned parquet tables is available at https://github.com/orgs/biobricks-ai/repositories. In addition, in the table below, the Committee provides a list of recommended databases for EPA to consider for possible incorporation into the Agency's curation efforts. In considering potential actions for this suggestion, the Committee suggests this is a high priority, but cautions that incorporating all of these databases would require a high degree of effort by EPA.

Table: Databases recommended to U.S. EPA to consider for the NCCRP.

Name	Link	Description
Chemical Identity & Properties		

Name	Link	Description
PFAS Tox Database	https://pfastoxdatabase.org/	Collaborative group of university and non-profit based scientists to support comparators
Interstate Technology and Regulatory Council (ITRC)	https://pfas-1.itrcweb.org/	Technical resources for addressing environmental Releases of PFAS; small database of structure / physical / chemical / toxicology data.
ChemIDPlus	https://chem.nlm.nih.gov/chemidplus	Contains chemical, physical, and some hazard/toxicology information.
Zinc20	https://pubs.acs.org/doi/10.1021/acs.jcim.0c00675	Billions of small molecules specifications.
Human Metabolome Database	https://hmdb.ca/	Small molecule metabolites. This includes drug bank (drugs/metabolites relevant to some PFAS like fluoxetine and detergents – antimicrobials).
***In Vitro* Hazard Data**		
LINCS L1000	https://lincsproject.org/LINCS/data/overview	Compilation of gene expression profiles.
The Cell Image Library	https://doi.org/10.1093/gigascience/giw014 and http://www.cellimagelibrary.org/home	Morphological profiles of 30,000 small molecules via cell painting.
Gene Expression Omnibus	https://www.ncbi.nlm.nih.gov/geo/	A public functional genomics database - array and sequence data.
***In Vivo* Hazard Data**		
FDA Adverse Event Reporting System (FAERS)	http://open.fda.gov/data/faers/	FDA Adverse Event Reporting System
Chembl	http://www.ebi.ac.uk/chembl/	Manually curated database of bioactive molecules; combines chemical / bioactivity / genomic data.

Name	Link	Description
Clinvar	http://ncbi.nlm.nih.gov/clinvar/	Aggregated information on genomic variation / human health relationships.
PharmGKB	https://www.pharmgkb.org	PharmGKB is a comprehensive resource that curates knowledge about the impact of genetic variation on drug response for clinicians and researchers.
Integrated Chemical Environment (ICE)	https://ice.ntp.niehs.nih.gov/	Data sets curated for targeted toxicity endpoints by NICEATM and others.
Comparative Toxicogenomics Database	http://ctdbase.org	Curated associations between chemicals, pathways, diseases, exposures, organisms, genes, and anatomy.
Echemportal	https://www.echemportal.org	Chemical hazard classifications from 30+ data participants.

- When possible, open-source tools should be used for the referenced document review workflows. Open-source tools enable greater transparency and replicability. In considering potential actions for this suggestion, the Committee suggests this is a high priority, however it's not possible to gauge the specific degree of effort this would require by EPA.
- In addition to a harmonization of chemical identifiers, there is a need to harmonize any entities that associate chemicals with values. Understanding which tests are indicated for different regulatory needs and designing models that merge the outputs of different assays is challenging when there are ambiguous relationships between test protocols, assays, and chemical properties. There are ontologies that attempt to hierarchically name assays (bioassayontology.org). Adoption of an existing method, or creation of a new method, to both unambiguously identify tests and identify relationships between tests is suggested. For example, knowing which assays are referenced by which ECHA or OECD guidelines and where those guidelines are referenced in hazard classifications requires controlled vocabularies for assays, guidelines, classifications, and their relationships. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- While the health outcome databases appropriately focus on experimentally derived toxicology data, a focus on mining the existing literature for epidemiological data linking exposures and health outcomes could be considered. This is particularly relevant in the case of some PFAS, where toxicokinetics and toxicodynamics are very different in humans than in commonly used rodent models. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a high degree of effort by EPA.

- There is a need in the modeling ecosystem for comparative validation. When new computational models are constructed to estimate NCCRP endpoints, their use should be benchmarked and compared to existing tools to verify their effectiveness (i.e., accuracy, etc.). A large, confidential validation set,[17] that is not publicly shared, could be used periodically as a consistent method of comparison for new models. It is critical to define model uncertainties and their potential impact on decision making. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a high degree of effort by EPA.
- If there is a plan to impute or fill in missing chemical property gaps, the method of imputation should be clear and the use of estimates to build new estimates should be limited to reduce error propagation. EPA should indicate the maximum percentage of data imputation that is acceptable. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a medium degree of effort by EPA.
- Several of the suggested source databases allow for public submission and depositing of new data. For example, the Gene Expression Omnibus (GEO) database allows researchers to deposit raw and processed high throughput sequencing and array-based data identifying molecular signatures of chemical exposures. If EPA is planning on allowing direct public submissions to EPA databases, specific guidance would need to be developed. Such guidance, if developed, would need to specify, among other things, how to add new data to the constructed system, and whether there are tools to deposit data directly. In considering potential actions for this suggestion, the Committee suggests this is a low priority that would require a low degree of effort by EPA.
- Given the low safe use levels for some chemicals, stakeholder concerns (NGOs, public) regarding exposure, it may be useful to expand those exposure scenarios in the CPDat to include dermal exposure scenarios in clothing and occupational personal protective equipment. This is particularly relevant given the EPA's emphasis on equity, environmental justice, and cumulative impacts.[18] As the EPA is compiling existing exposure data through the Multimedia Monitoring Database, the agency could consider making these data public and easily accessible, which could help to build trust with environmental justice and fence line communities. In considering potential actions for this suggestion, the Committee suggests this is a low priority that would require a high degree of effort by EPA.
- Some data sources are beginning to adopt semi-automated curation methods that use AI tools to automate the extraction of structured data from unstructured sources. Automated methods can introduce unknown biases and sources of error. When possible, these data should be flagged and be separable from non-automated approaches. In considering potential actions for this suggestion, the Committee suggests this is a low priority that would require a medium degree of effort by EPA.

Recommendations

The Panel offers the following recommendations:

……

[17] A confidential (private) validation set can be a valuable tool for generating independent comparative evaluations of NAM performance. When publicly known data on chemical-properties (or responses elicited by chemicals in assays) are used for validation, model developers could easily achieve strong performance by aligning their models with published values (overfitting). This diminishes the value of the models by potentially reducing their ability to generalize to new data. One way that this might be addressed by EPA is for the Agency to consider using both public and confidential (private) validation sets, with the confidential (private) sets consisting of CBI data. This would allow NAM developers to build and test their models on highly curated public data, while also providing the Agency to create high quality evaluations on CBI data. It would also be beneficial to periodically update the validation sets to correct errors and align with evolving goals' updates also make it more difficult to overfit models.

[18] See Washburn *et al*, 2005: https://pubmed.ncbi.nlm.nih.gov/15984763/

Recommendation 2.1: Implementing a system for easy replication has high value and relatively low added effort. Accordingly, the Committee recommends EPA should include a programmatic method to easily download a versioned copy of all of the open access data. This will allow stakeholders to better align their analyses with best practices created in NCCRP. A single bulk download is a less costly and more maintainable way to distribute the created data than application programming interface (APIs), which create uptime and versioning issues and create additional work for developers. A bulk download that can be accessed via tools like ftp, rclone, wget, and curl, will make it easier for developers to use the created data. When a dataset is very large, serving data in a method that allows efficient mirroring (and reduces redundant downloading) is recommended.

EPA Response:

Thank you for this comment. EPA is committed to making non-CBI data available publicly in a manner which allows for efficient reuse of the data. Throughout the NCCRP collaboration, we will regularly evaluate which methods for data delivery and data access are the most feasible and useful. ORD values public release of data and delivers data publicly through user interfaces such as the CompTox Chemicals Dashboard, downloadable data and tools, as well as APIs, which are anticipated to be updated and expanded throughout the StRAP cycle. ORD will continue to investigate efficient and effective data delivery tools and appreciates these suggestions.

Recommendation 2.2: The Committee recommends development of documented standard operating procedures for quality control that should be implemented in place of ad hoc methods. Development of automated quality control processes to identify outliers, data conflicts, and or likely sources of error should be considered to reduce the cost and effort of these procedures. If missing data will be imputed, the methods of imputation should follow a defined protocol and imputed values flagged. Automated quality control tests are high value but also significant effort. Thus, the choice to act on this recommendation, and the design and implementation of such activities will need to be carefully thought through.

EPA Response:

EPA appreciates this feedback. Scientific investigations within ORD abide by the principles and requirements stated in ORD's Quality Management Plan (QMP) for Scientific Research. The QMP details the requirements for a Quality Assurance Project Plan (QAPP), which is required to be established for each ORD research project. The QAPP details the necessary quality assurance activities that must be implemented to ensure quality and accuracy of the results of the work performed. Appropriate quality control processes, which may include standard operating procedures and automated methods, are determined through consultation with an assigned Quality Assurance Manager (QAM).

Recommendation 2.3: The Committee recommends EPA undertake the creation of standard validation sets for the evaluation of NAMs. These validation sets, constructed by the Agency using highly curated public data, would be used to fairly, and quantitatively, evaluate NAMs. In addition to developing public validation sets, the Agency should consider developing confidential (private) validation sets. Although

confidential (private) validation sets are not necessary or required, their value as a fair comparator increases[19] and the capacity for NAM developers to overfit AI models or construct *in vitro* models specifically to perform well on validation decreases. However, managing validation sets could create significant value for the NAM ecosystem, but present a high effort, high maintenance, and high responsibility deliverable. Accordingly, the choice to act on this recommendation, and the design and implementation of such activities will need to be carefully thought through.

EPA Response:

Thank you very much for your feedback, these are valuable recommendations to consider as we determine how best to implement the NCCRP. Within ORD the goal is to publish open-source models with available training and test sets such that model performance can be replicated by stakeholders to the extent practicable.

Charge Question 3

Q.3: As described in Research Area 3 of the accompanying White Paper (pages 28-33), planned research activities are focused on developing, refining, and evaluating (Q)SAR and other predictive models for physical-chemical properties, environmental fate/transport, hazard, exposure, and toxicokinetics.

a. Please comment on the (Q)SAR and predictive modeling proposed, as well as the proposed informatics platform for management of input data and development and management of (Q)SAR and other predictive models. In your comments, please address whether there are additional (Q)SAR models, approaches, or other informatics platform features that could be considered.

b. Please comment on any additional features that could be considered in the evaluation of these models, applicability domain(s), and association documentation.

Narrative

In its review and response to this charge question, the Committee considered the strengths and possible weaknesses of the (Q)SAR and QSUR approaches presented, alternative approaches and additional approaches and activities with the potential to augment these Quantitative structure-activity relationship (QSAR/QSUR) methods. The Committee also considered various forms of uncertainty in QSAR/QSUR approaches and how to characterize and report on them.

The Committee commends ORD and OPPT on an ambitious and groundbreaking approach to advance chemical assessments within the EPA and perhaps more broadly. Goals presented in the "white paper" are clearly stated and, if properly funded, have significant potential to achieve the desired effect of streamlining and improving chemical hazard assessments. Improvements that are planned in QSARs for physical chemical processes, fate and transport, and toxicological modes of actions and mechanisms are well described and reasonable. The use of QSURs was considered innovative and reasonable. The

[19] Op. cit. footnote 17.

Committee identified a need for confirmatory empirical (not *in silico*) data to ground truth model output for a subset of existing compounds.

While the research plan outlined in the NCCRP white paper largely focuses on developing and actualizing advanced methods for human health safety evaluations, improving ecotoxicology methods is equally important. The Committee supports the plan's outlined in the NCCRP white paper to 1) expand curation of *in vivo* ecotoxicology hazard data for improving the ECOTOX database, 2) incorporate tiered modified ecotoxicology studies and HTTr for ecotoxicology (EcoHTTr) as part of the proposed initial *in vitro* NAM screening battery, and 3) conduct the proposed studies of up to 60 chemicals (representing five chemical structural domains of interest for which ecological toxicity data and/or understanding of the applicability of currently (Q)SAR models is limited) with the goal of improving ecological effects QSAR modeling.

The Committee is impressed with the plans for the (Q)SAR and predictive modeling proposed, as well as the proposed informatics platform, and found the WebTEST tool for (Q)SARs to be a significant strength for the EPA, primarily as an organizing platform to integrate data and modeling efforts. The modeling directions (QSUR, HTTK, fate and transport) are all appropriately aligned to stakeholder needs and will be useful tools that are publicly available to assist data-poor decisions. The Committee also lauds the proposed expansion of the framework (to the OPEn structure-activity Relationship App - OPERA) and the incorporation of QSUR as a novel tool that could greatly improve exposure assessments.

The Committee structured our suggestions and recommendations so that tasks that could have significant impacts in the near term and that do not require substantial investment are listed first. Suggestions and recommendations that are more visionary and challenging are listed at the end. These tasks could require several iterations, review and engagement with the scientific community and stakeholders before reaching final form, but the Committee believes that these are appropriate directions for the Agency to follow.

Strengths

- Developing a data and computing infrastructure that integrates machine readable data and modeling platforms will have a large long-term impact by enabling efficient use of expanding/evolving models and growing data sets. This activity strongly complements other activities such as the WebTEST tool and generation of toxicity data itself.
- The WebTEST platform is a significant strength and should remain a priority, because it improves access and usability and enables community QSAR modeling building.
- The selection of QSAR model targets (QSUR, HTTK, fate and transport, toxicity, etc.) is clearly aligned with and supports stakeholder needs for decision making/risk assessment.
- The addition of QSUR is innovative and has the potential to have high impact on other activities like use cases for exposure assessment.
- Requiring that QSAR/QSUR models are publicly available, including the associated training sets, algorithms, and validation work assures transparency, improves confidence, and allows all such models to be properly tested and benchmarked.
- Clearly articulating the expectation that QSAR approaches are developed for application in data-poor environments will ensure appropriate methods are developed and appropriate testing/verification/assessment approaches are created.

- Expanding EPISuite to OPERA will be a strength, given the added functionality of the OPERA platform. Specifically linking structural characteristics to important mechanisms of toxicity will facilitate the direction of ORD's activities in the QSAR space.

Suggestions

- The Committee suggests that EPA consider developing and/or articulating the Agency's plan for horizon scanning to assure that emerging published QSAR models are added to the EPA model suite over time. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- The Committee suggests the EPA examine the pros and cons of requiring that computational models be open source. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- The Committee suggests EPA consider exploring the appropriateness of new machine learning approaches designed for sparse data sets (e.g., few shot machine learning methods[20]) for QSAR modeling. Traditionally developed for image analysis, this method may or may not be of value here. EPA could explore this by consulting with experts in this technology or by conducting a literature review. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a low degree of effort by EPA.
- Given that chemical purity data is already collected by GC or LC MS analysis, the Committee suggests EPA evaluate the value of implementing a method[21] for measurement of partition coefficient during these same GC and LC MS runs and implement if EPA judges the value justifies the investment. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a medium degree of effort by EPA.
- As the EPA moves from development of open-source QSAR models using open-source data to use of data and models protected by CBI, the Committee suggests it is important for the Agency to develop appropriate standards and criteria for utilization of those data and models. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a medium degree of effort by EPA.
- The Committee suggest EPA consider tracking the opportunity that molecular dynamic simulation models (quantum chemistry models from institutions like DOE[22] and NSF) might offer for improving the accuracy of prediction of chemical properties or calculation of additional chemical properties useful for QSAR, categorization or QSUR. Molecular dynamic models might also be able to adjust ligand-binding models developed for one species (e.g., estrogen, human) to another species where the receptor exists in a different internal environment (e.g., pH, temperature, etc.). In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a medium degree of effort by EPA.
- As efforts to develop databases of known metabolites mature, the Committee suggests that the EPA develop a framework or method for incorporating assessment of known metabolites/breakdown products/transformation as classes of compounds for QSAR modeling and incorporate QSAR or other models that predict metabolites/breakdown products/transformation products for later exposure and toxicity QSAR modeling. Additionally, impurities that arise through specific manufacturing procedures may pose concerns, and like breakdown products or metabolites, may in

[20] "Few-Shot Learning new categories of data (that the pre-trained model has not seen during training) using only a few labeled samples per class. It falls under the paradigm of meta-learning (meta-learning means learning to learn)." https://blog.paperspace.com/few-shot-learning/.

[21] See: OECD. 2022. Test No. 117: Partition Coefficient (n-octanol/water), HPLC Method. Organization Économique Cooperation and Development. 11 pp. https://doi.org/10.1787/9789264069824-en

[22] https://www.energy.gov/science/bes/articles/breakthrough-reported-machine-learning-enhanced-quantum-chemistry.

the future become more amenable to being predicted with greater confidence. These suggestions should be considered for implementation in the longer term. In considering potential actions for this suggestion, the Committee suggests this is a low priority at the present time that would require a high degree of effort by EPA.

Recommendations

The Panel offers the following recommendations:

Recommendation 3.1: The Committee recommends EPA expand tools/approaches for reporting on confidence in QSAR model predictions including measures of variance and uncertainty (e.g., domain of applicability, strength of training data) and provide documentation how those measures of variability and uncertainty are calculated, including the actual code.

EPA Response:

We appreciate this recommendation and understand the importance of reporting on measures of confidence in all model predictions. Development and application of (Q)SAR models under NCCRP will follow the OECD principles for validation of (Q)SAR, which suggests that consideration of (Q)SARs for regulatory purposes is facilitated by, amongst other attributes, appropriate measures of its goodness-of-fit, robustness, and predictivity. ORD will continue to evaluate appropriate metrics for reporting on confidence in QSAR model predictions throughout the NCCRP collaboration.

Recommendation 3.2: The Committee recommends EPA establish and implement methods, if feasible, for including a "flag" in toxicity databases for compounds that cause non-specific effects (e.g., surfactants and facile reactants), or other flags (e.g., related to overfitted dose-response curves in some *in vitro* data sets) to assure that these problems do not adversely and unknowingly affect QSAR modeling.

EPA Response:

Thank you for the comment. We appreciate this suggestion and will consider how we can ensure that various properties of both the data and the modeling do not adversely or unknowingly affect modeling predictions. Indeed, ToxCast data already includes flagging for concentration-response curves that may have resulted from over-fitting or other aberrant behaviors to enable data users to filter data appropriately for use. Ongoing work to develop and expand chemical lists, including surfactants, will continue and enable users to create improved datasets for modeling. Ongoing work in ORD to create views of the data or to assist users in filtering data for various applications will continue to evolve.

Recommendation 3.3: To support the value and impact of the WebTEST resource, the Committee recommends EPA: a) engage the regulatory science community in one or more workshops to provide feedback on performance and usability and solicit suggestions for further development and b) develop and deploy a semi-automated (easy to access and utilize by the community) workflow for model evaluation that is quantitative, transparent, consistent and offers comparative benchmarking.

EPA Response:

Thank you for this recommendation. Throughout the NCCRP collaboration, EPA will be exploring various means of engaging with the regulatory science community, as a valued means of receiving feedback. We appreciate your recommendation in this context, specifically with regards to WebTEST. Early development of WebTEST2.0 is public for use and comment and will be presented at other venues including conferences and publications. EPA will continue to work towards a transparent, quantitative, consistent workflow that can be easily accessed and deployed, starting with public release of these tools, as well as the code and input data.

Recommendation 3.4: As efforts to expand toxicity databases to address gaps in domains of applicability conclude, the Committee suggests EPA identify the next priority areas where toxicity data need to be expanded to improve the ability to develop QSAR and related models that support ecotoxicity assessments (e.g., terrestrial toxicity). This recommendation should be considered for implementation in the longer term (e.g., as part of the next CSS StRAP cycle).

EPA Response:

The EPA appreciates this input. As noted, a regular and critical step in preparing for the next ORD StRAP cycle is an evaluation and identification of priority research areas, which will include specific needs related to the NCCRP collaboration.

Charge Question 4

Q.3: As described in Research Area 4 of the accompanying White Paper (pages 33-40), planned research activities are focused on developing and evaluating a suite of *in vitro* NAMs that could be used by external stakeholders for testing and data submissions under TSCA, as well as potentially informing and/or expanding new chemical categories.

Please include in your comments, other assays and/or endpoints to consider for the research plan.

Narrative

EPA's proposal outlines a comprehensive NAM-based program to screen new chemicals for safety in accordance with the Lautenberg Chemical Safety for the 21st Century Act. The proposed approach follows the path identified in the EPA CompTox Blueprint, including: 1) broad-based analyses for chemical interactions with numerous molecular/protein targets (discrete target or generalized/multi-target effects) to cover a wide breadth of potential chemical-biological target interactions; and 2) targeted analyses to predict potential adverse outcomes.

As our knowledge of biological pathways underpinning human and ecological health improves, so should the appropriateness and availability of NAM-based approaches. Meanwhile, Integrated Approaches to Testing and Assessment (IATA), as well as disorder and disease models, that are fit for purpose can be used to enhance current NAM predictions of chemical toxicity to support consistent evaluation of data within a weight-of-evidence approach. For purposes of risk-based screening assessment of new chemical submissions, EPA has focused on *in silico* and *in vitro* tools, primarily used in high-throughput modes, this approach is appropriate to support EPA's requirement under TSCA to review new chemical submissions with limited available toxicity and exposure data. These screening methodologies for human health and ecotoxicology are equally important. The Committee supports the NCCRP, and encourages incorporation of these advanced human health and ecological screening methods into IATA approaches that include exposure information and IVIVE to provide contextual dosimetry.

Strengths

- Both human health and ecotoxicological assessments are included in the defined NAMs approach.

- The integration of data streams in an IATA-based approach strengthens subsequent conclusions, particularly when cheminformatic fingerprints/QSARs are combined with broad and targeted NAM assessments to evaluate data consistency.
- The broad coverage of potential toxicity pathways allows greater confidence in NAM-based assessments. For example, the proposed human health assessment uses:
    - Broad-based, high-content screening approaches to examine numerous chemical-biological target interactions [i.e., using HTPP and HTTr for respiratory toxicity to identify whether a chemical may act at a discrete molecular target (specific MIE) or produce generalized stress responses due to multiple molecular targets (non-specific response)].
    - Targeted screening approaches to provide information on specific MIEs, key events or hazard-related processes [i.e., SafetyPharm, DevTox Germ Layer Reporter assay, genotoxicity assays – micronucleus test and Ames] coupled with *in vitro* assays to refine HTTK for improved IVIVE.
- These multiple data streams will help identify molecular/protein targets, inform potential hazard identification, and improve dosimetry estimates, which can be evaluated for consistency and biological plausibility and thus can improve confidence.
- One of the key challenges is interpreting bioactivity information in a manner that is commensurate with the information it conveys. Bioactivity of a substance, as measured by a response in an early key event in an adverse outcome pathway may, or may not, be indicative of the ability of a substance to lead to an adverse health effect. This will depend on the scientific confidence that's been established in predicting along the causal chain of key event relationships in that pathway. The research described in the NCCRP will provide important data and knowledge to help develop these linkages from early key events to later events and to adverse outcomes.
- Incorporation of developmental toxicity potential (DevTox assay) is advantageous to provide data on a 'high concern hazard' that is typically not available for new chemicals.
- Expanding available data using human ALI respiratory cell and/or precision-cut lung slice cultures and HTTr data will provide valuable information on the performance of these tools to predict potential inhalation hazards.
- The proposal to screen 200-300 candidate chemicals is important for gaining scientific confidence in the application of EPA-identified suites of NAMs, particularly if the candidate chemicals are selected to fill *in vitro* and *in silico* gaps and improve applicability domains. Additionally, this exercise will help identify opportunities to continuously improve NAM specificity and sensitivity (e.g., DevToxGLR currently at 58% specificity).
- Applying Eco-HTTr for a subset of chemical structures for which ecotoxicity data and QSAR applicability is limited will help delineate the domain of applicability of the method, while generating data that improves mechanistic understanding of ecotoxicity.
- EPA's focus on analytical quality control of identity and purity of candidate chemicals is critical.

Suggestions

- EPA should review and suggest plausible methods for poorly soluble or non-aerosolizable chemicals (e.g., microvolume dosing in DMSO using applicable instrumentation and NAMs). When considering potential actions for this suggestion, the Committee indicates this as a high priority that would require a low degree of effort by EPA.
- TSCA requires that risk evaluations of new and existing chemicals consider potentially exposed and susceptible subpopulations such as infants, children, pregnant women, workers, and the elderly ("vulnerable subpopulations"). The Committee suggests that the NCCRP explicitly describe how a suite of selected *in vitro* NAMs considers (or does not consider) these vulnerable subpopulations and

continuously work toward better accounting for such subpopulations. This is also in line with the Agency's increasing emphasis on equity, environmental justice, and cumulative impacts. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a low degree of effort by EPA.

- The application of biological test systems to obtain endpoint-specific data should be conducted using standardized approaches that have been optimized as part of the fit-for-purpose determination. While the culture and assay methodology for more conventional *in vitro* test systems may have been well-defined, this may not be the case for the newer, more complex systems such as ALI and microphysiological systems (and their exposure and dosimetry methods). Where possible, EPA could partner with various stakeholder organizations to facilitate method development/standardization, which would allow additional expert input, other funding sources, and accelerated timelines. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- Before proceeding to invest significantly more resources into the DevTox GLR-Endo assay development and standardization research, ORD and OPPT should work together to develop and collect feedback from the regulatory science community on a detailed review paper (DRP) on the state of the science of developmental toxicity NAMs, including comparisons of the predictive performance, strengths and limitations of the DevTox GLR-Endo assay compared to, for example, the Murine Embryonic Stem Cell Assay and the Zebrafish embryo developmental toxicology assay. This DRP would be expected to provide scientific justification to support investing in research in the most promising fit-for-purpose assays suited for OPPT's new chemicals program decision context. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- The Committee suggests that EPA articulate how considerations like route of exposure (e.g., inhalation, dermal, oral), bioavailability, metabolism requirements (e.g., formation of active metabolites), etc. will affect NAM requirements. For example, if the relevant route of exposure is dermal and the compound is poorly absorbed, should the NAM data requirements be the same? In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a high degree of effort by EPA.
- The Committee suggests that EPA expand the chemical domain for cell painting (HTPP) to better represent the TSCA universe by applying a cheminformatics approach to ensure appropriate chemical diversity. It would be useful for EPA to compare data generated from HTTr vs. HTPP in terms of identification/correlation of bioactivity profiles (cell phenotypic changes vs. expression changes) and bioactive PoD concentrations. As HTPP matures, EPA should develop a DRP that includes endpoints examined, translation to *in vivo* adverse effects, grouping endpoint data to identify positive responses, sensitivity of HTPP vs. other broad-based NAM screening approaches, impact of cell type, availability of orthogonal assays, and domain of chemicals tested. In addition, there is a new HESI/Broad Institute Emerging Systems Toxicology for Assessment of Risk (eSTAR) project to use cell painting and transcriptomics to evaluate liver toxicity. ORD and OPPT could join this group to provide their experience and gather stakeholder input on the use of these technologies. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a high degree of effort by EPA.
- EPA should consider new approaches to assess genotoxicity to ensure that the selected methodologies are the most appropriate. While conventional assays may have provided substantial guidance on the evaluation for the genotoxic potential of materials, these methods are often cumbersome and time consuming. How are newer (and potentially more predictive) assays being incorporated into the overall testing scheme to replace the older assays? In considering potential

actions for this suggestion, the Committee suggests this is a high priority that would require a high degree of effort by EPA.

- EPA should consider including human precision-cut lung slice cultures (where airway contractility may be evaluated as a phenotypic endpoint), along with the other identified complex, heterocellular 3D experimental models that offer high content phenotypic responses. Recent advances in preservation and increased throughput have made these more accessible and allow for larger scale and repeat donor-based studies. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a medium degree of effort by EPA.
- The Committee suggests EPA continue to evolve the BioTransformer (OECD Toolbox) program to address the likelihood of metabolite formation. BioTransformer (OECD Toolbox), which is used to predict liver-generated metabolite, can predict metabolites that have not been observed in guideline studies. For example, HTTK and HTTr data generation on metabolism and comparison with other *in silico* tools to predict metabolism will be valuable to better understand the relevance of these predictions. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a high degree of effort by EPA.
- The Committee recognizes the potential of the DevTox GLR assay endpoints (biomarkers for differentiation of the endoderm, mesoderm and ectoderm germ layers), but is concerned that this assay represents a very limited portion of development. The Kapraun/Wambaugh HTTK computational model for pregnancy[23] simulates gestational week 13 until parturition, whereas the gastrulation step measured in the DevTox GLR assay occurs at weeks 3-4 of pregnancy. The Committee suggests that the NCCRP include developmental toxicity assays that span longer, equally relevant periods of gestation. EPA may consider NAMs such as ReproTracker® and devTOXquickPredict™ in this effort. Furthermore, the Committee recommends that each of these applicable gestational stages be incorporated into HTTK models to allow IVIVE. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a high degree of effort by EPA.
- The Committee suggests EPA undertake a periodic review of the development of NAMs that address complex endpoints. The technology regarding all of the more complex biological test systems (e.g., 3D reconstructed tissues, organ on a chip, precision-cut lung slices) is rapidly evolving. Accordingly, EPA should routinely assess at appropriate points in time in the StRAP cycle the NAMs in the NCCRP initiative, and where necessary modify these. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- As a longer-term suggestion, the Committee suggests that the NCCRP consider how and when higher order NAMs (e.g., zebrafish embryos, planaria, *C. elegans*) would support effective assessment of integrated endpoints including neurobehavior. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a low degree of effort by EPA.

.....

[23] Reproductive Toxicology, Volume 113, October 2022, Pages 172-188; https://www.sciencedirect.com/science/article/abs/pii/S0890623822001381.

Recommendations

The Panel offers the following recommendations:

Recommendation 4.1: The Committee recommends that EPA's NCCRP conduct dedicated reviews of the program (perhaps aligned with the StRAP cycle) to assess progress, opportunities, and challenges with implementation, including an opportunity for stakeholders and the public to provide input and feedback. This will be especially valuable for further refinement and application of more innovative NAMs like HTPP.

EPA Response:

Thank you for this suggestion. EPA is committed to transparency and engagement with stakeholders and the public in all aspects of our work, including under the NCCRP collaboration. EPA is committed to reviewing the outcome of projects under the NCCRP, relative to the initial project objectives, to assess progress and identify additional opportunities for collaboration. In addition to engagement opportunities specific to the NCCRP, we encourage stakeholders and the public to engage with EPA on NAMs through the public webinars (*e.g.*, Computational Toxicology Communities of Practice, training program(s)), and regular NAMs Workshop events which EPA has committed to conduct under the EPA NAMs Workplan.

Recommendation 4.2: The Committee recommends that the Agency optimize and standardizes NAM development using Good *In Vitro* Method Practices (GIVIMP)[24], which would aid in their acceptance and transferability.

EPA Response:

Thank you for this recommendation. A principal tenet of ORD/EPA research is transparency, which comes with good quality assurance/quality control practices. As previously noted, scientific investigations within the Office of Research and Development abide by the principles and requirements stated in ORD's QMP for Scientific Research. Research strictly adheres to a QAPP, which is required to be established for each ORD research project. The QAPP details the necessary quality assurance activities that must be implemented to ensure quality and accuracy of the results of the work performed. The principles set forth in GIVIMP are paralleled in the development of EPA ORD QAPPs and operating procedures for *in vitro* screening.

Recommendation 4.3: The Committee recommends that research aimed at defining a suite of *in vitro* NAMs to inform new chemical reviews account for potentially exposed or susceptible subpopulations specifically as it relates to relevant, differential biological considerations across the population (e.g., variance in toxicokinetics, disease states, age). Information on exposure should be integrated with the *in vitro* NAM results to better inform the risk determinations required in the TSCA new chemicals program.

EPA Response:

We appreciate this thoughtful input which is consistent with our current plans to include contextual demographic and life stage information as we develop a suite of *in vitro* NAMs and apply toxicokinetic models to estimate human equivalent doses.

[24] https://www.oecd.org/env/guidance-document-on-good-in-vitro-method-practices-givimp-9789264304796-en.htm

Charge Question 5

Q.5: In the Background of the accompanying White Paper (pages 5-16), information on challenges in new chemical assessment and the vision statement for the NCCRP are presented. The primary vision of the NCCRP is to modernize the process for evaluating new chemicals under TSCA by supporting the evolution of OPPT's use of new and existing methods, approaches, and tools using innovative science.

Please comment on the extent to which the Research Areas may address the issues identified in the Background and vision statement. Please also include potential additional research areas for EPA to consider.

Narrative

EPA is required to make timely evaluations of data-poor chemicals that are newly entering the market, using the best available science and methods, in a way that is as transparent as possible. To address these challenges, EPA has developed a collaborative research program between OPPT and ORD. This collaborative research program proposes four research areas: chemical categories and read-across; database development and growth; predictive models for hazard, kinetics, and exposure; and *in vitro* NAMs. A fifth research area focuses on decision support tools and is minimally included in this review as methods are still under development.

This charge question asks the BOSC to comment broadly on the extent to which the challenges EPA identifies in the background and research statement are addressed by the research proposal.

These challenges include:

- High volume of submissions (average of 500 new chemical submissions per year).
- Need for rapid decision making (general requirement for EPA to decide in 90 days).
- Lack of information available on chemicals including human and environmental hazard data as well as use and exposure data.
- Requirement to make (and justify) a formal decision for all new chemicals.
- Substantial informatic needs to making and documenting decisions.
- Promoting transparency when possible while maintaining CBI on a large percentage of the new chemicals.

The NCCRP research program leverages efforts in ORD to help fill data gaps and manage information and builds on years of extensive work by EPA and others in the regulatory science community to develop predictive toxicology and exposure science tools. The NCCRP has many strengths that meet the challenges discussed above. Updating, expanding, and developing new chemical categories and furthering the development and refinement of QSAR and predictive models will help fill data gaps. Using predictive models when other data are not available will help EPA make science-based decisions efficiently, which is critical to address the time, information and resource limitations described above. The BOSC is pleased to see that the long-term investment within ORD in predictive tools is finding application within the Agency.

The NAS Tox21 report and prior BOSC reviews have recommended efforts to incorporate human and epidemiologic data into the development, and refinement, of predictive toxicology tools. Furthering

these recommendations, we are pleased to note the ORD case study with PFAS and electronic health records. Additional work on incorporating clinical data, occupational health data, and molecular epidemiologic data into some of the decision-support tools, as feasible, would increase confidence in the tools and more clearly link AOP networks with potential human relevance. The Committee recognizes that this is a long-term objective and encourages further work toward this goal.

The Committee is concerned that the significant percentage of new chemicals with CBI claims may pose a challenge for evaluating the effectiveness of the tools that ORD is developing. If the CBI chemicals differ systematically from the non-CBI chemicals, then the plan to validate the tools using only non-CBI information could result in failures to recognize issues or limitations with the tools. For example, if a large percentage of CBI claims are for polymers, ORD should test the tools on a representative array of polymer structures. For both polymers and UVCBs, it is also critical to model how they change over time and ensure that the entire mixture is evaluated.

Strengths

- The proposed research program is well-tailored to rapidly evaluating chemicals that have little or no toxicity/exposure data.
- The program also leverages resources and skills that the ORD team has already developed and prioritizes and operationalizes cross-agency connections and collaboration.
- The modeling of potential use and exposure is an important component of the research program.
- Another strength of the research program involves data generation, notably that:
    - The approaches will be assessed with about 200-300 chemicals. This relatively large number of chemicals will go a long way towards addressing scientific confidence in assay performance (sensitivity, specificity, reproducibility) and establishing domains of applicability. A greater degree of uncertainty and lower scientific confidence would result if fewer chemicals were to be evaluated.
    - Data will be generated to inform IVIVE and kinetics modeling. This is important because IVIVE allows the concentrations of substances producing *in vitro* responses to be converted to corresponding *in vivo* exposure levels to support risk-based interpretations; and
    - The inclusion of methods to evaluate the effects of chemicals that may be inhaled will significantly advance respiratory tract NAM development. Inhalation exposure to vapors, aerosols, and particulates is an important route of exposure in humans. Developing NAMs that expose cells at the air-liquid interface to simulate *in vivo* inhalation exposures is technologically challenging. While considerable progress has been made in this area over the last several years, much more research is needed to develop the methods and datasets required to use these assays with confidence as a replacement for *in vivo* lab animal inhalation tests.
- The vision to put multiple data streams together into a unified usable decision support tool is ambitious and clearly needed.

Suggestions

- The Committee suggests EPA consider approaches for ground-truthing the ORD exposure models using existing biomonitoring datasets (e.g., from CDC, NIH), including, where feasible, biomonitoring using non-targeted assessment. This could be similar to the multimedia monitoring database for environmental chemical data. In considering potential actions for this suggestion, the Committee suggests this is a high priority that would require a medium degree of effort by EPA.
- The Committee suggests EPA explore opportunities early on to develop or disseminate analytical methods for environmental monitoring for newly introduced chemicals. The availability of analytical

methods for environmental monitoring is also important for continued evaluation of exposure once a chemical is on the market. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a high degree of effort by EPA.

- ORD should consider longer-term goals of developing tools to predict how the toxicity of UVCBs change throughout their lifecycle, from manufacture through disposal, due to shifts in the composition of the mixtures. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a high degree of effort by EPA.
- Longer-term work should include the development of exposure models, when feasible, to include and predict unintended exposures from activities like recycling consumer products into new products, potable water reuse, and composting. In considering potential actions for this suggestion, the Committee suggests this is a medium priority that would require a high degree of effort by EPA.

Recommendations

The Panel offers the following recommendations:

Recommendation 5.1: The Committee recommends EPA consider ways to integrate human data into databases and tools when possible, including clinical, occupational, and other epidemiological study data, especially in the context of AOP networks. This information could provide a link between mechanistic results and human outcomes and help to benchmark tools EPA is developing.

EPA Response:

We appreciate your feedback. Throughout implementation of the NCCRP, EPA will consider the most appropriate methods for benchmarking tools currently under development and associating mechanistic results with human outcomes. If human data are used, EPA will adhere to federal policies on the protection of human subjects.

Recommendation 5.2: The Committee recommends EPA assemble (or develop) the training or reference chemical sets used for developing and evaluating methods and models for both human health and ecological effects such that they mirror the characteristics of CBI and non-CBI chemicals that OPPT typically receives, including polymers and UVCBs. EPA should also try to identify and address relevant impurities and byproducts, including residual monomers and oligomers. This will help to ensure that the methods are applicable to the chemistries OPPT is typically addressing.

EPA Response:

EPA thanks the Committee for this recommendation. ORD and OPPT are committed to working collaboratively under the NCCRP to ensure that methods and models developed in ORD are applicable to the chemistries OPPT is typically evaluating under their new chemicals program. Additional chemical curation and cheminformatic work to link chemical representations that may be within a mixture or UVCB is one avenue of ongoing research in ORD to address some of these issues. Continued communication between ORD and OPPT colleagues with consideration of the chemistries evaluated by the new chemicals program at OPPT will support progress being made on evaluating and expanding applicability domains of *in silico* and *in vitro* approaches, or delineating cases in which more traditional approaches may be needed to fill data gaps.

SUMMARY LIST OF RECOMMENDATIONS

Charge Question 1: As described in Research Area 1 of the accompanying White Paper (pages 16-20), planned research activities are focused on updating and refining the chemical categories and read across methods used by OPPT.

Please comment on whether there are other approaches or chemical characteristics that could be considered when developing the categories and analog identification methodologies.

- **Recommendation 1.1:** The Committee recommends ORD, in conjunction with OPPT, design, conduct, and publicly disseminate case studies evaluating the performance of the current OPPT categories compared to the new approaches, such as GenRA, to support a read-across

assessment where analog toxicity data are compared to target chemical toxicity data that are initially blinded to the assessor. Case studies should include several situations (e.g., where an understanding of metabolism is critical for establishing suitable analogs, where bioactivity data are limited, where small changes in chemistry have the potential to have significant impact on toxicity). Given the widespread interest in read-across in the global regulatory science community, it would be informative for these EPA case studies to include comparisons to read-across methodologies developed by other organizations (cited above), and document, as appropriate, alignment with ECHA's Read Across Assessment Framework.[14] These case study activities will help document scientific confidence in the newer approaches and support transitions from the existing OPPT approaches to the newer read-across approaches (e.g., GenRA).

- **Recommendation 1.2:** The Committee recommends ORD and OPPT explore the potential to use CBI data within the GenRA and other inference models for grouping and read-across. One option to explore would be using federated learning[15] with differential privacy data methods, or similar technologies, that allow the private data to be retained and protected locally while still enabling the data to be used in model development (for example, the Machine Learning Ledger Orchestration for Drug Discovery collaboration[16]). Another option to consider would be developing a protected in-house user downloadable app (e.g., like the OECD toolbox download) to enable data use while protecting CBI. This is a particularly important research activity that may improve approaches for new chemicals that fall outside the current domains of non-CBI databases.
- **Recommendation 1.3:** The Committee recommends that, in addition to having a Research Area Coordination Team (RACT), ORD and OPPT should establish a process and schedule for jointly evaluating the scientific confidence and readiness of these NAMs for updating the new chemical grouping and read-across methods that are intended to be used by OPPT's new chemicals program. A set schedule is needed to ensure the review process is keeping pace with advances in science and knowledge, to focus the next round of research, and to provide the certainty needed for the Agency and stakeholders to efficiently and confidently implement these methodologies. This would also ensure predictability in the application of program guidance for a set time period. One schedule to consider is alignment with the StRAP cycle. For example, the schedule for this scientific confidence and readiness review could be sequenced to finish at a point in time where the results of the review and recommendations for additional research serve as input into development of the next StRAP.

Charge Question 2: As described in Research Area 2 of the accompanying White Paper (pages 20-28), planned research activities are focused on expansion and further development of existing public databases in ORD containing chemistry, hazard, exposure, and toxicokinetic information relevant to TSCA chemicals.

Please comment on this effort, including in your feedback useful sources of chemical information that could be incorporated into the curation efforts.

- **Recommendation 2.1**: Implementing a system for rapid replication has high value and relatively low added effort. Accordingly, the Committee recommends EPA should include a programmatic method to easily download a versioned copy of all of the open access data. This will allow stakeholders to better align their analyses with best practices created in NCCRP. A single bulk download is a less costly and more maintainable way to distribute the created data than APIs, which create uptime and versioning issues and create additional work for developers. A bulk

download that can be accessed via tools like ftp, rclone, wget, and curl, will make it easier for developers to use the created data. When a dataset is very large, serving data in a method that allows efficient mirroring (and reduces redundant downloading) is recommended.

- **Recommendation 2.2:** The Committee recommends development of documented standard operating procedures for quality control that should be implemented in place of ad hoc methods. Development of automated quality control processes to identify outliers, data conflicts, and or likely sources of error should be considered to reduce the cost and effort of these procedures. If missing data will be imputed, the methods of imputation should follow a defined protocol and imputed values flagged. EPA should indicate the maximum percentage of data imputation that is acceptable. Automated quality control tests are high value but also significant effort. Thus, the choice to act on this recommendation, and the design and implementation of such activities will need to be carefully thought through.
- **Recommendation 2.3:** The Committee recommends EPA undertake the creation of standard validation sets for the evaluation of NAMs. These validation sets, constructed by the Agency using highly curated public data, would be used to fairly, and quantitatively, evaluate NAMs. In addition to developing public validation sets, the Agency should consider developing confidential (private) validation sets. Although confidential (private) validation sets are not necessary or required, their value as a fair comparator increases[19] and the capacity for NAM developers to overfit AI models or construct *in vitro* models specifically to perform well on validation decreases. However, managing validation sets could create significant value for the NAM ecosystem, but present a high effort, high maintenance, and high responsibility deliverable. Accordingly, the choice to act on this recommendation, and the design and implementation of such activities will need to be carefully thought through.

Charge Question 3: As described in Research Area 3 of the accompanying White Paper (pages 28-33), planned research activities are focused on developing, refining, and evaluating (Q)SAR and other predictive models for physical-chemical properties, environmental fate/transport, hazard, exposure, and toxicokinetics.

a. Please comment on the (Q)SAR and predictive modeling proposed, as well as the proposed informatics platform for management of input data and development and management of (Q)SAR and other predictive models. In your comments, please address whether there are additional (Q)SAR models, approaches, or other informatics platform features that could be considered.

b. Please comment on any additional features that could be considered in the evaluation of these models, applicability domain(s), and association documentation.

- **Recommendation 3.1:** The Committee recommends EPA expand tools/approaches for reporting on confidence in QSAR model predictions including measures of variance and uncertainty (e.g., domain of applicability, strength of training data) and provide documentation how those measures of variability and uncertainty are calculated, including the actual code.
- **Recommendation 3.2:** The Committee recommends EPA establish and implement methods, if feasible, for including a "flag" in toxicity databases for compounds that cause non-specific effects (e.g., surfactants and facile reactants), or other flags (e.g., related to overfitted dose-response curves in some *in vitro* data sets) to assure that these problems do not adversely and unknowingly affect QSAR modeling.
- **Recommendation 3.3:** To support the value and impact of the WebTEST resource, the Committee recommends EPA: a) engage the regulatory science community in one or more workshops to provide feedback on performance and usability and solicit suggestions for further

development and b) develop and deploy a semi-automated (easy to access and utilize by the community) workflow for model evaluation that is quantitative, transparent, consistent and offers comparative benchmarking.

- **Recommendation 3.4:** As efforts to expand toxicity databases to address gaps in domains of applicability conclude, the Committee suggests EPA identify the next priority areas where toxicity data need to be expanded to improve the ability to develop QSAR and related models that support ecotoxicity assessments (e.g., terrestrial toxicity). This recommendation should be considered for implementation in the longer term (e.g., as part of the next CSS StRAP cycle).

Charge Question 4: As described in Research Area 4 of the accompanying White Paper (pages 33-40), planned research activities are focused on developing and evaluating a suite of *in vitro* NAMs that could be used by external stakeholders for testing and data submissions under TSCA, as well as potentially informing and/or expanding new chemical categories.

Please include in your comments, other assays and/or endpoints to consider for the research plan.

- **Recommendation 4.1:** The Committee recommends that EPA's NCCRP conduct dedicated reviews of the program (perhaps aligned with the StRAP cycle) to assess progress, opportunities, and challenges with implementation, including an opportunity for stakeholders and the public to provide input and feedback. This will be especially valuable for further refinement and application of more innovative NAMs like HTPP.
- **Recommendation 4.2:** The Committee recommends that the Agency optimize and standardizes NAM development using Good *In Vitro* Method Practices (GIVIMP)[24], which would aid in their acceptance and transferability.
- **Recommendation 4.3:** The Committee recommends that research aimed at defining a suite of *in vitro* NAMs to inform new chemical reviews account for potentially exposed or susceptible subpopulations specifically as it relates to relevant, differential biological considerations across the population (e.g., variance in toxicokinetics, disease states, age). Information on exposure should be integrated with the *in vitro* NAM results to better inform the risk determinations required in the TSCA new chemicals program.

Charge Question 5: In the Background of the accompanying White Paper (pages 5-16), information on challenges in new chemical assessment and the vision statement for the NCCRP are presented. The primary vision of the NCCRP is to modernize the process for evaluating new chemicals under TSCA by supporting the evolution of OPPT's use of new and existing methods, approaches, and tools using innovative science.

Please comment on the extent to which the Research Areas may address the issues identified in the Background and vision statement. Please also include potential additional research areas for EPA to consider.

- **Recommendation 5.1:** The Committee recommends EPA consider ways to integrate human data into databases and tools when possible, including clinical, occupational, and other epidemiological study data, especially in the context of AOP networks. This information could provide a link between mechanistic results and human outcomes and help to benchmark tools EPA is developing.
- **Recommendation 5.2:** The Committee recommends EPA assemble (or develop) the training or reference chemical sets used for developing and evaluating methods and models such that they mirror the characteristics of CBI and non-CBI chemicals that OPPT typically receives, including polymers and UVCBs. EPA should also try to identify and address relevant impurities and

byproducts, including residual monomers and oligomers. This will help to ensure that the methods are applicable to the chemistries OPPT is typically addressing.

APPENDIX A: MEETING AGENDA

Monday, October 24, 2022

TIME (EDT)	AGENDA ACTIVITY	PRESENTER
12:00 pm	Meeting Kickoff, Federal Advisory Committee Act provisions and expectations, Logistics	Tom Tracy (OSAPE) Designated Federal Official
12:10 pm	Opening Remarks from OCSPP	Michal Freedhoff (OCSPP) Assistant Administrator for the Office of Chemical Safety and Pollution Prevention (OCSPP)
12:20 pm	Opening Remarks and Introduction of panel members	Paul Gilman BOSC Executive Committee Chair Richard Becker Justin Teeguarden Panel Co-Chairs
12:35 pm	Agenda Overview	Annette Guiseppi-Elie (ORD) Acting National Program Director, Chemical Safety for Sustainability (CSS)
12:50 pm	Challenges in New Chemical Assessment Under the Amended Toxic Substances Control Act (40 min presentation + 20 min Q&A)	Louis "Gino" Scarano (OPPT) Senior Science Advisor, Office of Pollution Prevention and Toxics (OPPT)
1:50 pm	Break	
2:00 pm	Overview of the New Chemicals Collaborative Research Program (NCCRP) (80 min presentation + 25 min Q&A)	Katie Paul Friedman (ORD) Toxicologist, Center for Computational Toxicology and Exposure (CCTE)
3:45 pm	Overview of Ad Hoc Panel Charge	Annette Guiseppi-Elie (ORD) Acting National Program Director, Chemical Safety for Sustainability (CSS)
4:00 pm	Public comment	Richard Becker Justin Teeguarden Panel Co-Chairs
6:00 pm	Closing (ORD) *May occur earlier, dependent on public comment period*	Bruce Rodan (ORD) Associate Director for Science for ORD

Tuesday, October 25, 2022

TIME (EDT)	AGENDA ACTIVITY	PRESENTER(S)
12:00 pm	Convene Meeting	Richard Becker Justin Teeguarden Panel Co-Chairs
12:05 pm	Opening Remarks from ORD	Maureen Gwinn (ORD) Principal Deputy Assistant Administrator for Research and Development
12:15 pm	Overview of Day 2 by Moderator/Re-state charge	Annette Guiseppi-Elie (ORD) Acting National Program Director, Chemical Safety for Sustainability (CSS)
12:30 pm	Open Comment and Question Period from the Panel	Richard Becker Justin Teeguarden Panel Co-Chairs
1:30 pm	Break out group discussion of charge questions (closed sessions)	Richard Becker Justin Teeguarden Panel Co-Chairs
3:00 pm	Break	
3:30 pm	Continue group discussion of charge questions (closed sessions)	Richard Becker Justin Teeguarden Panel Co-Chairs
5:30 pm	Committee questions and next steps	Full Committee and EPA Participants
6:00 pm	Meeting adjournment	Bruce Rodan (ORD) Associate Director for Science for ORD

APPENDIX B: MATERIALS

Material Provided in Advance of the Meeting

- Agenda
- Charge questions

Material Provided During or After the Meeting

- PowerPoint presentation slides presented during the meeting
- ORD responses to BOSC follow-up questions