The New Chemicals Collaborative Research Program: Modernizing the Process and Bringing Innovative Science to Evaluate New Chemicals Under TSCA

A Summary Report to the Board of Scientific Counselors (BOSC) on an integrative research plan within the 2023-2026 Chemical Safety for Sustainability Strategic Research Action Plan

October 2022

Office of Research and Development
Office of Chemical Safety and Pollution Prevention
U.S. Environmental Protection Agency

Contents

Executive Summary

The US EPA Office of Research and Development (ORD) has been developing and evaluating new approach methodologies (NAMs) and decision support tools for toxicology and exposure to build a next generation risk assessment toolbox. In a new joint effort referred to as the New Chemicals Collaborative Research Program (NCCRP), the US EPA Office of Chemical Safety and Pollution Prevention (OCSPP) and ORD are working together to bring 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) and, as amended in 2016, make one of several affirmative determinations regarding risks to human health and the environment. Based on the determination, EPA may need take further action to prevent unreasonable risks before manufacturing for the chemical can commence. With hundreds of new chemical notices 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. Modernizing the new chemicals review process and bringing innovative science to inform risk assessment and decision making will help overcome information gaps and help the Agency meet its statutory requirements in a timely, effective, and efficient manner.

Under the NCCRP, ORD is working with the Office of Pollution Prevention and Toxics (OPPT) within OCSPP to advance five key Research Areas: (1) updates and refinements to chemical analog and category approaches; (2) development and expansion of databases containing TSCA chemical information; (3) development and refinement of predictive models for physicochemical properties, environmental fate/transport, hazard, exposure, and toxicokinetics; (4) integration and application of *in vitro* NAMs; and (5) development of a TSCA new chemicals decision support tool that utilizes curated data. Each of these five Research Areas represents translation and extension of computational toxicology research that has been in development under the vision of the CompTox BluePrint (Thomas et al., 2019) and the EPA NAM Work Plan (USEPA, 2021b), which together form a strategic roadmap for: developing and integrating NAMs to fill information gaps; establishing scientific confidence of NAM application to regulatory toxicology; and engaging with stakeholders. The NCCRP was announced in February 2022 followed by a public meeting in April 2022 (USEPA, 2022a). ORD has aligned research planning for the NCCRP with the ORD Chemical Safety for Sustainability (CSS) Strategic Research Action Plan (StRAP) for 2023-2026, which has been reviewed by ORD management, representatives of OCSPP, the Board of Scientific Counselors (BOSC) executive committee, and other stakeholders. Alignment with the StRAP ensures that research within the NCCRP supports broader ORD objectives, including coordination of NAM and interactive tool development, as well as coordination of resources. **In this report, details about the research proposed in the CSS StRAP that are relevant for the NCCRP will be summarized and are the focus of this document and review by the BOSC.**

Research within the scope of the NCCRP to address new chemical assessment is expected to have broad applicability. As such, research will extend beyond this four year StRAP cycle and may involve future collaborations with other relevant ORD research programs such as Health and Environmental Risk Assessment, other federal institutions (*e.g.*, the Division of Translational Toxicology, formerly known as the National Toxicology Program, at the National Institute of Environmental Health Sciences in the National Institutes of Health), and regulatory toxicology experts at other agencies such as the European Chemicals Agency and Health Canada. Potential engagement with other ORD research programs and external collaborators will leverage additional expertise and resources. **These**

collaborations and the OPPT implementation details are beyond the scope of this report and are thus not part of the BOSC review.

Acronyms

Acronym	*Explanation*
AOP	Adverse Outcome Pathway
BOSC	Board of Scientific Counselors
CBI	Confidential Business Information
CpDat	Chemicals and Products Database
CSS	Chemical Safety for Sustainability [a research program within ORD]
CvT	Concentration versus Time Database
DSSTox	Distributed Structure-Searchable Toxicity Database [for chemistry information]
ECOTOX	ECOTOXicology Knowledgebase
EPA	Environmental Protection Agency
HERA	Health and Environmental Risk Assessment [a research program within ORD]
HTTK	High-throughput Toxicokinetics
HTPP	High-throughput Phenotypic Profiling [also known as Cell Painting]
IATA	Integrated Approaches to Testing and Assessment
IUCLID	International Uniform Chemical Information Database
IVIVE	*In Vitro* to *In Vivo* Extrapolation
MIE	Molecular Initiating Event
MMDB	Multimedia Monitoring Database
NAM	New Approach Methodology
NCCs	New Chemical Categories [see *Chemical Categories Used to Review New Chemicals under TSCA*]
NCD	New Chemicals Division [in OPPT]
NCCRP	New Chemicals Collaborative Research Program
OCSPP	Office of Chemical Safety and Pollution Prevention (OCSPP)
OHT	Organisation for Economic Co-operation and Development [OECD] Harmonized Template
OPPT	Office of Pollution Prevention and Toxics
ORD	Office of Research and Development
PFAS	Per- and Polyfluoroalkyl Substances
POD	Point-of-Departure
(Q)SAR	(Quantitative) Structure-Activity Relationship; some are more or less quantitative
(Q)SUR	Quantitative Structure-Use Relationship
StRAP	Strategic Research Action Plan
TEST/WebTEST	Toxicity Estimation Software Tool/Web Toxicity Estimation Software Tool
ToxRefDB	Toxicity Reference Database
ToxValDB	Toxicity Value Database
TSCA	Toxic Substances Control Act
UVCBs	Chemical substances of Unknown or Variable Composition, Complex Reaction Products and Biological Materials

Background

The US EPA Office of Research and Development (ORD) and the Office of Chemical Safety and Pollution Prevention (OCSPP) have been working collaboratively for many years to apply new approach methodologies (NAMs) to regulatory toxicology needs. Most recently, ORD engaged with the Office of Pollution Prevention and Toxics (OPPT) on the *National PFAS Testing Strategy* (USEPA, 2021a) and *A Proof of Concept Case Study Integrating Publicly Available Information to Screen Candidates for Chemical Prioritization under the Toxic Substances Control Act (TSCA)* (USEPA, 2021c). The New Chemicals Collaborative Research Program (NCCRP) is an ambitious planned collaboration that will enable next generation risk assessment while addressing OPPT's regulatory needs and bolstering ORD's efforts to develop NAMs. More specifically, the NCCRP seeks to both rapidly modernize available approaches, including decision support tools, for new chemicals evaluation and also impact the engineering of the databases, models, and tools that ORD is building for multiple stakeholders to execute the vision of the CompTox BluePrint (Thomas et al., 2019) and the EPA NAMs WorkPlan (USEPA, 2021b). If successful, with each 4-year research cycle, the NCCRP will enable progress in ORD and OPPT toward meeting their respective goals to advance chemical risk assessment. These goals include greater acceptance and scientific confidence in NAMs applied within the NCCRP; greater understanding of the future needs of NAM development; and decision support tools that provide consistent, but iteratively improving, access to and integration of myriad data sources with chemical information, including data derived from NAMs.

In this section, the regulatory toxicology challenges posed by TSCA, how strategic research planning to address these challenges is proceeding, and the launch of the proposed NCCRP will be presented as background prior to discussion of the planned research.

New Chemical Risk Assessment Challenges

The regulatory toxicology challenges faced by OPPT guide both immediate and long-term goals for the NCCRP. On June 22, 2016, TSCA was amended by the Frank R. Lautenberg Chemical Safety for the 21st Century Act.[1] At the US EPA, OPPT, within OCSPP, is responsible for carrying out the mandates of TSCA, including provisions requiring the review, determination of unreasonable risk, and subsequent management of any identified risks associated with both existing (those already in the marketplace) and new (those that manufacturers intend to bring to market) chemicals. OPPT's New Chemicals Division (NCD) is responsible for the review of new chemicals prior to introduction of a new chemical into U.S. commerce (via either import or domestic manufacturing). NCD received an average of 500 new chemical

[1] The Frank R. Lautenberg Chemical Safety for the 21st Century Act | US EPA

notices per year since 2010 and a total of over 50,000 since 1979. Reference in this report to "new chemical," "submissions" or "notices," is generally meant to be inclusive of all potential notice types under TSCA Section 5 requiring review, such as significant new use notices[2], low volume and low release and exposure exemption notices, and pre-manufacture notices (PMNs). Depending on the type of notice, the statute or appropriate regulations generally require EPA to make determinations within 30 to 90 days of notice receipt. Details on the full review process can be found on EPA's website.[3]

A first key challenge posed by new chemical assessment is the dearth of information available, as many new chemicals have little to no chemical-specific information available. To address data gaps for both human and environmental hazards, and exposure, OPPT has led the world in the use of (quantitative) structure-activity relationships ((Q)SARs) and other predictive models and tools coupled with the use of category-based approaches.[4] The methods, approaches, and tools developed over the past four decades have been used to carry out tens of thousands of evaluations under TSCA. ORD plans to augment the currently available (Q)SAR, read-across, and predictive approaches with new chemical groupings, systemized read-across approaches, and (Q)SARs developed and evaluated using internationally recognized and established validation principles (OECD, 2007; OECD, 2014).

A second major challenge has been implementing the changes in statutory requirements under TSCA for new chemicals, which now mandates reviews and determinations for all new chemicals and thus increased efforts and resources required in making determinations as well as documenting and supporting those determinations. Prior to the 2016 amendments, EPA could evaluate whether to "drop" a new chemical from further review and determined that for roughly 80% of annual new chemical submissions a full determination or further detailed analyses were not necessary. A submitter could then commence manufacture of the new chemical upon expiration of the review period without restriction. In addition to requiring affirmative determinations for all new chemicals, the 2016 amendments introduced several new possible determinations that EPA could make, including that the chemical is "not likely" to present an unreasonable risk or that the available information is "insufficient" to permit a reasoned evaluation. In total, the statute now sets forth five possible determinations:

[2] Significant new use notices may not necessarily pertain to new chemical substances but are nonetheless part of the new chemicals program and are submitted under TSCA Section 5. See: https://www.epa.gov/reviewing-new-chemicals-under-toxic-substances-control-act-tsca/filing-significant-new-use-notice

[3] https://www.epa.gov/reviewing-new-chemicals-under-toxic-substances-control-act-tsca

[4] Predictive Models and Tools for Assessing Chemicals under TSCA

1. The chemical substance or significant new use presents an unreasonable risk of injury to human health or the environment;
2. The information available is insufficient to permit a reasoned evaluation of the health and environmental effects of the chemical substance or significant new use;
3. In the absence of sufficient information to make an evaluation, the chemical substance or significant new use may present an unreasonable risk of injury to health or the environment;
4. The chemical substance is or will be produced in substantial quantities and the substance either enters or may reasonably be anticipated to enter the environment in substantial quantities or there is or may be significant or substantial exposure to the chemical; or,
5. The chemical or significant new use is not likely to present an unreasonable risk of injury to human health or the environment.[5]

Further, the 2016 amendments to TSCA explicitly require EPA to review new chemicals under the "conditions of use" – a phrase defined in the law to include the circumstances under which the chemical is, "intended, known or reasonably foreseen to be manufactured, processed, distributed in commerce, used or disposed of." The identification of conditions of use – particularly those that are "reasonably foreseen" – presents a unique challenge for the New Chemicals Program, given the data submitted by the manufacturer regarding the intended use are often limited, let alone any future use. NAMs developed or used by ORD and other stakeholders may have the potential added benefits of addressing additional hazard data gaps, identification of conditions of use, and furnishing more information for making the required determination.

OPPT has long used predictive models and other non-vertebrate methods for evaluating new chemical submissions. The 2016 amendments to TSCA reinforced the need for more predictive models and non-vertebrate methods via addition of Section 4(h) entitled, *Reduction of Testing in Vertebrates*, which requires that: "The Administrator shall reduce and replace, to the extent practicable, scientifically justified, and consistent with the policies of this title, the use of vertebrate animals in the testing of chemical substances or mixtures under this title…" (Section 4(h)(1)). This subsection further requires that prior to EPA making a request or adopting a requirement for testing using vertebrate animals, to consider, as appropriate and to the extent practicable and scientifically justified, reasonably available

[5] Section 2604(a)(3) at 15 USC Chapter 53

existing information, including toxicity information, computational toxicology and bioinformatics, and high-throughput screening methods and prediction tools (Section 4(h)(1)(A)).[6] Section 4(h)(2)(A) required OPPT to release a Strategic Plan in 2018[7] to promote the development and implementation of alternative test methods and strategies. As suggested in Section 4(h), new chemical reviews need to continue to incorporate new, innovative methods, approaches, and tools to maintain a modern and efficient process.

The requirement to review and make an affirmative determination on each new chemical submission for their conditions of use, the accompanying need to support each determination with a robust assessment, the 90-day review requirement, and TSCA's direction regarding non-animal testing, all underscore the need for updated approaches for new chemical assessments. Additionally, incorporating additional NAMs could increase both efficiency and transparency.

A third major challenge area can be summarized as a substantial informatic need, in which increased computational accessibility of data and modernized chemical information management is required to efficiently perform assessments. Inherently, some TSCA information may be claimed as confidential business information (CBI), and current public availability of existing chemical data to inform computational approaches to chemical categories, read-across, and (Q)SAR development may be limited. The development of human health and ecological risk assessments for new chemicals has relied heavily on the limited information and data provided in a new chemical submission, often associated with CBI claims. Providing additional public data and utilizing modernized tools will increase both the transparency in decisions and the amount of information available to support new chemical determinations. Since the 2016 amendments to TSCA, efforts have been underway in OPPT to make information claimed as TSCA CBI publicly available where the Agency has determined that the information is not entitled to confidential treatment.[8] OPPT and ORD plan to work in tandem to increase

[6] Section 4(h)(1)(B) contains further requirements, including "encouraging and facilitating— (i) the use of scientifically valid test methods and strategies that reduce or replace the use of vertebrate animals while providing information of equivalent or better scientific quality and relevance that will support regulatory decisions under this title; (ii) the grouping of 2 or more chemical substances into scientifically appropriate categories in cases in which testing of a chemical substance would provide scientifically valid and useful information on other chemical substances in the category; and (iii) the formation of industry consortia to jointly conduct testing to avoid unnecessary duplication of tests, provided that such consortia make all information from such testing available to the Administrator."

[7] See: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/strategic-plan-reduce-use-vertebrate-animals-chemical

[8] Go to: Confidential Business Information under TSCA | US EPA

the computational accessibility of non-confidential information, including chemical-specific information that ORD already curates and compiles into publicly available databases.

Bringing together existing chemical data for TSCA-relevant chemicals and NAM information will require long-term work, extending beyond 2023-2026, to modernize access and utilization of heterogeneous data coming from disparate sources. OPPT is working to modernize its digital chemical information system that supports the entirety of the TSCA program. Although the effort to modernize its chemical information system is internal to OPPT, the modernization effort is necessary for full implementation of the tools developed under the NCCRP and may be coordinated with collaborative research activities under the NCCRP. Within work relevant to the NCCRP, ORD plans to implement an International Uniform Chemical Information Database (IUCLID)[9] environment for existing curated databases and public document sources. This will begin with curated *in vivo* hazard data but will extend to as many data types as possible. Building an IUCLID compatible data environment within ORD and OPPT contributes to a long-term Agency goal of using information from multiple sources (e.g., public and internal) and supports development of a decision support tool for new chemical assessments that can accelerate the pace of risk evaluations under TSCA.

Strategic Research Planning in ORD

Research within ORD is guided by the EPA Strategic Plan,[10] which delineates clear goals for Agency decisions and actions. Based this Plan, ORD delivers research to meet both short- and long-term Agency needs, to inform Agency decisions, and to support the needs of tribal, state, and community partners. ORD coordinates this research through four-year planning cycles within six National Research Programs, including Chemical Safety for Sustainability (CSS), which includes research relevant to the NCCRP.

Due to the high visibility of this cross-cutting research, technical perspectives from the ORD federal advisory committee, the BOSC, are being sought. A draft of the CSS StRAP for FY23-26, also known as StRAP4, was already presented to the BOSC executive committee for review,[11] with an emphasis on the Topic, Research Area,[12] and Output level details, which follow a hierarchical order. The

[9] See https://iuclid6.echa.europa.eu/project-iuclid-6
[10] See: https://www.epa.gov/planandbudget/strategicplan
[11] https://www.epa.gov/system/files/documents/2022-04/epa-ord_css-fy23-26-draft-strap_3-28-2022.pdf
[12] Note that "Research Area" in the StRAP is not equivalent to "Research Area" in the NCCRP. Research Area in the StRAP groups scientific expertise at a high level. Research Area in the NCCRP refers to more specific collections of research to be performed and are described herein as Research Areas 1-5.

broad Topics provide overarching research foci, followed by Research Areas which group the science expertise and research that will be assembled to address partner needs. Outputs provide even more detail about the results that will be achieved under each Research Area. Outputs are composed of Products, which are the tangible deliverables of the National Research Program. In the report herein, details about the relevant Outputs and narrative summaries of multiple Products will be provided to give an overarching view of the coordinated research relevant to the NCCRP across CSS in StRAP4. This review of the planned ORD research relevant to the NCCRP by the BOSC provides another opportunity to obtain stakeholder perspectives. Previously, a public meeting announcing the NCCRP sought feedback on the five research areas proposed (and outlined below in Figure 1 and Table 1). Future implementation details within OPPT, based on progress and application of research relevant to the NCCRP, will be presented to other federal advisory committees that provide advice on TSCA-relevant work at the EPA.

Much, but not all, of the proposed research within ORD that is relevant to the NCCRP is consolidated in CSS StRAP Output 408.4, *Strengthening the Science to Support New Chemicals Evaluation*. This Output includes research that addresses the needs of EPA's programs and regions, states, tribes, and external partners as well as identified cross-cutting research priorities for CSS: developing a tiered testing strategy, building confidence in NAMs, increasing data availability and accessibility, and contributing to decision support and translation. The innovative science required to address the risk assessment of new chemicals necessitates coordination and work across the CSS portfolio, beyond CSS 408.4 (see Figure 2), and may involve future collaborations with other relevant ORD National Research Programs such as the Health and Environmental Risk Assessment research program.

New Chemicals Collaborative Research Program (NCCRP)

The NCCRP will likely involve other federal institutions (*e.g.*, the Division of Translational Toxicology, formerly known as the National Toxicology Program, at the National Institute of Environmental Health Sciences in the National Institutes of Health) as well as collaborations with other regulatory entities such as Health Canada and the European Chemicals Agency to leverage the expertise and resources of these entities to address TSCA-specific needs as well as to enhance broad applicability of the research. The results of the effort are expected to increase the efficiency of new chemical reviews, but more importantly bring innovative science to new chemicals assessments and decisions for protecting human health and the environment using the authority under TSCA Section 5.

ORD's Center for Computational Toxicology and Exposure (CCTE) and Center for Public Health and Environmental Assessment (CPHEA) have been working closely with OPPT to develop this overarching research plan and coordinate activities. Additionally, internal and external partners will be consulted for input and research contributions. Previously, an overview of the NCCRP was released for public comment (USEPA, 2022a), with comments received in a docket.[13] While the focus of planned research relevant to the NCCRP falls within the next StRAP4 (FY23-26), the collaboration needed to support modernization and innovation for new chemicals assessment will likely extend beyond completion of StRAP4.

Problem and Vision Statement

Refinement of and updates to methods, approaches, and tools used by OPPT to evaluate new chemicals are critical to continuing to ensure the safety of new chemicals prior to their entrance into US commerce and that decisions regarding the risk posed by new chemicals to human health and the environment are supported by the best available science. Any changes should align with statutory deadlines, be operational in a data poor environment, make effective use of new data sources and approaches, and be transparent to the extent practicable (given that TSCA CBI may be used in the development of these approaches). The 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 through the use of innovative science.

[13] See: https://www.epa.gov/reviewing-new-chemicals-under-toxic-substances-control-act-tsca/new-chemicals-collaborative

Proposed NCCRP Research Areas

The five proposed Research Areas are described and summarized in Figure 1 and Table 1.

Figure 1. Interconnectivity of the NCCRP Research Areas.

These five Research Areas are interconnected efforts to ultimately integrate NAMs and computationally accessible data into a decision support tool that can be iteratively improved to support new chemical assessments. Research areas 1, 2, 3, and 4, bounded by a dashed rectangle, are all interrelated; computationally accessible data in Research Area 2 feeds into Research Areas 1 and 3, and informs data gaps to be addressed in Research Area 4. Research Areas 1, 3, and 4 supply data back to the database environment. Together, Research Areas 1-4 supply information to be used in Research Area 5 (development of a decision support tool). Research Areas: 1 = Update and refine chemical categories; 2 = Develop and expand databases containing TSCA chemical information; 3 = Develop and refine (Q)SAR and predictive models for physicochemical properties, environmental fate/transport, hazard, exposure, and toxicokinetics; 4 = Explore ways to integrate and apply *in vivo* NAMs in new chemical assessments; 5 = Develop a decision support tool to modernize the process.

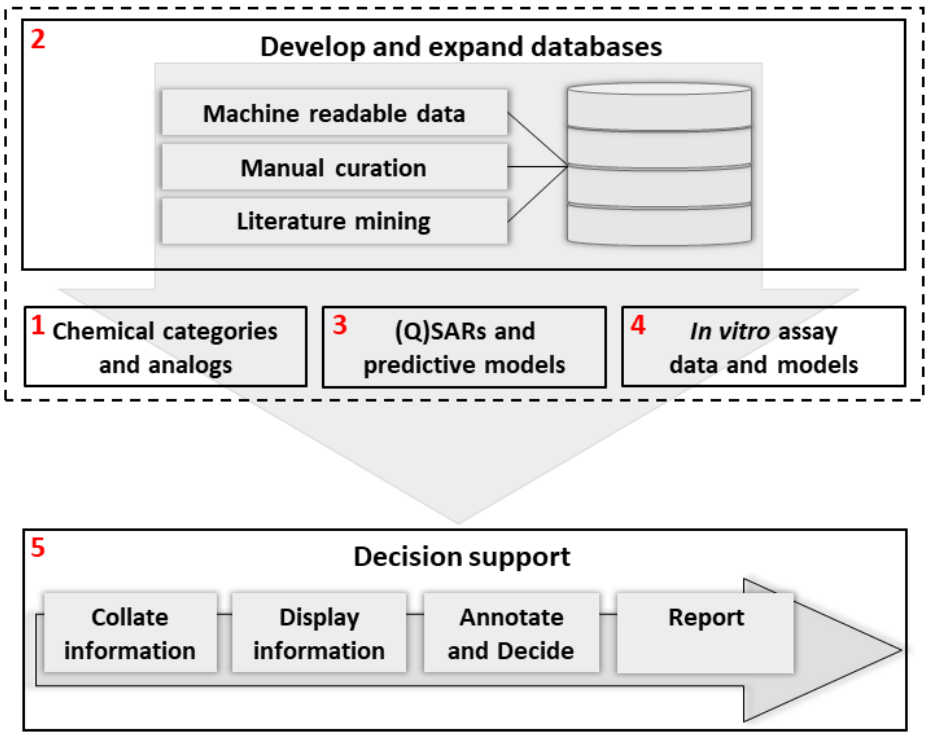


Table 1. Proposed NCCRP Research Areas.

Each Research Area is addressing key problems, and through applied research approaches, will yield one or more outcomes that have the potential to impact new chemical decisions in OPPT.

	Research Area	Problem	Approach	Expected Outcome(s)
1	Update and Refine Chemical Categories	Currently 56 TSCA categories, last updated 2010	Systematically define chemical categories and analogs for read-across using structural (and other) boundaries; physical-chemical properties; structural alerts for hazard, fate, exposure, and/or functional uses; existing hazard data; and/or, *in vitro* mechanistic and toxicokinetic data from NAMs	This will increase the efficiency of new chemical reviews and promote the use of the best available data to protect human health and the environment.
2	Develop and Expand Databases Containing TSCA Chemical Information	Existing TSCA information is not computationally accessible or easily searchable	Extract and curate available TSCA CBI study information Continue extraction and curation of physical-chemical property, environmental fate, hazard, and exposure information (non-CBI) in ORD databases Map information in ORD databases to standardized reporting templates and store in an International Uniform Chemical Information Database (IUCLID)	The TSCA CBI information will be combined with publicly available sources to expand the amount of information available, enhancing chemical reviews and enabling efficient sharing of chemical information across EPA. Safeguards for CBI will be maintained as appropriate in this process.
3	Develop and Refine (Q)SAR and Predictive Models for Physical-Chemical Properties, Environmental Fate/Transport, Hazard, Exposure, and Toxicokinetics	Currently used models are not always publicly accessible, easy to update with additional chemicals, or the best performing for all chemistries	Develop and update (Q)SAR and predictive models using existing data and curated data from Research Area 2 Evaluate models to determine the best suite for use by OPPT for regulatory purposes	Updated models that reflect the best available science, increased transparency, and a process for updating these models as science allows.

4	Explore Ways to Integrate and Apply *In Vitro* NAMs in New Chemical Assessments	Reduction in the use of vertebrate animals in accordance with TSCA Section 4(h) Many new chemical submissions are data poor Amended TSCA requires affirmative determination regarding unreasonable risk	Develop and evaluate a suite of *in vitro* NAMs for informing new chemical evaluations Use mechanistic and toxicokinetic *in vitro* NAMs to inform and refine chemical categories in Research Area 1	A suite of *in vitro* NAMs that could be used by external stakeholders for testing and data submissions under TSCA as well as informing and expanding new chemical categories
5	Develop a TSCA New Chemicals Decision Support Tool to Modernize the Process	Searching, collating, and integrating data for new chemical assessments is inefficient and costly	Build proof of concept software workflow that integrates all data streams in a new chemical risk decision context	A decision support tool that will efficiently integrate all the data streams (e.g., chemistry, fate, exposures, hazards) into a final risk assessment and transparently document the decisions and assumptions made. This will facilitate the new chemicals program tracking decisions over time and evaluating consistency within and across chemistries.

Figure 2. NCCRP Research Areas rely on CSS research.

The five NCCRP Research Areas cut across several CSS StRAP Outputs previously reviewed by the BOSC, ORD, and stakeholders. Blue boxes represent overlap of NCCRP research activities with CSS StRAP Outputs.

CSS Research Area	Output	1 Update and refine chemical categories	2 Develop and expand databases	3 Develop and refine QSAR and predictive models	4 Explore ways to integrate and apply NAMs	5 Develop a decision support tool
High-throughput toxicology (HTT)	401.1				■	
	401.2		■		■	
Rapid Exposure and Dosimetry (RED)	402.1		■			
	402.2		■			
	402.3			■		
Ecotoxicological Assessment and Modeling (ETAM)	406.3		■			
	406.5				■	
Chemical Characterization and Informatics (CCI)	407.1		■			
	407.3	■	■	■		
	407.4	■				
	407.5	■	■			
Integration, Translation, and Knowledge Delivery (ITK)	408.1		■			■
	408.3					■
	408.4	■	■	■	■	■

Proposed Research Relevant to the NCCRP

1. Update and Refine Chemical Categories

As TSCA new chemical notices are typically data poor, OPPT has historically relied heavily on the use of chemical categories[14] and read-across[15] as methods to fill data gaps, particularly for hazard characterization. OPPT currently uses the 2010 version of the *New Chemicals Program under TSCA Chemical Categories* document which identifies 56 chemical categories[16] based on chemical class, referred to herein as new chemical categories (NCCs). When OPPT evaluates a new chemical, determining if it belongs in an existing NCC is important for evaluating human health or environmental effects.

In Research Area 1, ORD and OPPT are proposing to develop a systematic, transparent, and reproducible approach for modernizing both chemical categories and read-across methods. Research will identify scientific information to support development or refinement of chemical categories and read-across methods, such as: structural (and other) boundaries; physicochemical properties; structural alerts for hazard, fate, exposure, and/or functional uses; mechanistic and toxicokinetic data from NAMs; and/or, existing hazard data. The new approach will document the data used to inform chemical categories as well as the basis of any similarity or read-across applications in a systematic manner.

The proposed approach will increase the efficiency of new chemical reviews and promote the use of the best available data to protect human health and the environment. Further, application of a chemical category approach itself should result in greater confidence in inferences made for a given

[14] These categories are used for analysis and risk management of individual new chemicals, and thus do not implicate Section 26(c) of TSCA, which allows EPA to take action with respect to a category of chemical substances. Nonetheless, the new chemicals categories use similar principles to categories under Section 26(c), which may be applied to, "a group of chemical substances the members of which are similar in molecular structure, in physical, chemical, or biological properties, in use, or in mode of entrance into the human body or into the environment, or the members of which are in some other way suitable for classification as such for purposes of [TSCA.]" The Organization for Economic Cooperation and Development (OECD) defines a category as "(C)hemicals whose physicochemical, toxicological and ecotoxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity may be considered as a group, or 'category'…" (p.11 in OECD, 2017).

[15] Read-across is defined as a data gap filling technique that relies on an analog or category approach, with analogs or categories defined on the basis of similarity of structure, properties, or other information. To "read across" is to apply data from a tested chemical for a particular property or effect to similar untested chemicals. See https://dx.doi.org/10.14573/altex.1410071 and https://dx.doi.org/10.1016/j.yrtph.2016.05.008 for further discussion.

[16] See *Chemical Categories Used to Review New Chemicals under TSCA*

target chemical (OECD, 2014), assuming that the categories applied are robust, which depends on category size (number of members) and the amount of data available for each category member. This research builds upon ongoing research and further motivates applied cheminformatic research within ORD to support CSS goals. The research covers several CSS StRAP Outputs listed in Table 2.

Table 2. StRAP Outputs Relevant to Update and Refine Chemical Categories

Relevant StRAP Output	Relevant StRAP Output Title
CSS.407.3	Develop new and improve existing structure activity relationship models to support risk assessment
CSS.407.4	Advancing chemical categorization approaches for aiding the interpretation and prediction of bioassay and toxicity outcomes
CSS.407.5	Advancing the use of structural, mechanistic, and toxicokinetic data to support categorization and classification of Per- and Polyfluoroalkyl substances (PFAS)
CSS.408.4	Strengthening the science to support new chemicals evaluation

A. Chemical category modernization approach

The 56 existing NCCs are characterized largely by structural features and in some cases by physicochemical properties. The key goals of collaborative research in this area are to implement the chemical categories in a transparent and reproducible manner that would permit updates with new information, such as additional structure descriptors, physicochemical data, or NAM data. Further, planned research will investigate to what extent new categories are needed to capture substances in the TSCA active inventory that could not be readily assigned to one of the 56 existing NCCs. This research will bridge between the current NCCs and development of an easily updated approach to chemical grouping.

First, the chemical structure information built into the current NCCs will be turned into a machine-readable format, such as system arbitrary target specification (referred to as SMARTS), to enable substructure searching and mapping to other types of structural descriptors, such as ToxPrints (Yang et al., 2015). This research will enable computational approaches to chemical grouping based on one or more types of structural descriptor(s) as well as other pertinent information. The TSCA non-confidential active chemical inventory will be profiled using the newly codified NCCs to assign them into their respective categories. Chemical categories may be developed by a combination of one or more of the following: use of more structural descriptors, physicochemical properties, predicted metabolism, *in vitro* mechanistic and toxicokinetic, and/or *in vivo* toxicity data, pending resources and available data. Finally, with research completed to better understand the chemical structure space encompassed within the

TSCA non-confidential active chemical inventory, research will evaluate to what extent the chemicals on this inventory fall within the applicability domain for existing ORD (Q)SAR models such as the Toxicity Estimation Software Tool (TEST) (USEPA, 2020b), with an online version referred to as WebTEST1.0, or other structural alert schemes (either existing or in development) to better characterize limitations in the ability of those models to make robust and reliable predictions. This will help target further data curation efforts for chemistry information in Research Area 2 aimed at trying to increase the applicability domain of structure alerts and models. The insights gained will help tailor the combination of NCCs and models as a proof-of-principle scheme that is most informative for *in silico* evaluation of the TSCA active inventory.

B. Expansion and application of systematic read-across

Generalized read-across (GenRA) (Helman et al., 2019a; Helman et al., 2018; Helman et al., 2019b; Shah et al., 2016; Shah et al., 2021), is a systematic, data-driven read-across approach developed by ORD. GenRA has been implemented as part of a read-across workflow within a web application initiated via chemical search or chemical structure drawing in the CompTox Chemicals Dashboard[17] and also as a Python package, genra-py, to facilitate batch processing with user-specific datasets. Though the GenRA approach has been applied to systematically evaluate *in vivo* toxicity datasets represented by potency values or binary hazard outcomes, the GenRA web application is currently structured to make binary (positive or negative) *in vivo* toxicity estimates based only on data curated into the Toxicity Reference Database (ToxRefDB). Local neighborhoods characterized by different fingerprints based on chemistry and/or bioactivity information, *e.g.*, circular Morgan fingerprints (Rogers and Hahn, 2010), ToxCast bioactivity fingerprints based on positive or negative assay responses, or a hybrid combination of both, can be used to identify candidate source analogs from which a GenRA prediction is derived for the toxicity outcomes of interest. GenRA is a tool that could be potentially implemented by the OPPT NCD in the near-term, and such translation efforts will be part of the NCCRP. GenRA can also be enhanced to better meet the needs of the NCCRP over the course of StRAP4, benefitting not only new chemical assessment, but also other applications of GenRA.

Key planned research to enhance GenRA capabilities, and also meet OPPT needs for transparent and reproducible read-across, will proceed in a number of areas, including: evaluating the impact of hybrid features on GenRA performance; extending similarity contexts to additional types of bioactivity data;

[17] https://comptox.epa.gov/dashboard/

evaluating the contribution of metabolism data to inform analog identification and evaluation; and, additional case studies to build confidence in the use of GenRA.

Within the current GenRA web application, a hybrid fingerprint can be constructed to search for candidate source analogs. This hybrid can currently be created by using up to three different fingerprint types with associated percentage weightings, *e.g.* 60% contribution from Morgan fingerprints versus 40% ToxCast bioactivity fingerprints. However, guidance as to what might be an optimal weight to use for each fingerprint type, and the extent to which this differs depending on the type of chemical (*e.g.*, structural class, functional groups) or toxicity endpoints being predicted *(e.g.,* liver toxicity versus kidney effects), has not been systematically evaluated. Such an analysis will characterize the relative contribution that different weighting schemes may play in predicting toxicity outcomes.

Additional information to define target to analog similarity, such as *in vitro* bioactivity, could be very informative for analog selection. An *in vitro* NAM such as high-throughput phenotypic profiling (HTPP) is broad in biological coverage, high-throughput, and multi-dimensional, and may be useful in understanding the bioactivity fingerprints of chemicals that share common biological targets (Nyffeler et al., 2022; Nyffeler et al., 2020; Willis et al., 2020). Other multi-dimensional assay suites, such as a safety pharmacology panel (Bowes et al., 2012; Ietswaart et al., 2020; Smit et al., 2021; Valentin et al., 2018), may also provide valuable mechanistic fingerprint data to evaluate the similarity of potential analogs for a target chemical. Additional research within GenRA aims to quantify the potential contribution of these bioactivity data in inferring toxicity in read-across applications.

Analog identification would be further informed by understanding related metabolites of a target. Initial planned research includes *in silico* predictions of liver-generated metabolites using a selection of prediction tools (*e.g.,* BioTransformer, OECD Toolbox) across a large and diverse chemical data set (*e.g.,* ToxCast chemical library). An accessible database of these metabolite predictions provides a foundation for investigation of chemical structural similarity and common metabolic pathways to better inform chemical categories and analog identification. Chemical similarity in related metabolite production, whether that be by virtue of the similarity in transformation profile, the sequence of transformations, or the structural similarity in the predicted metabolites themselves, will be determined. A possible extension to GenRA will aim to complement ongoing work in other areas of CSS to collect experimental metabolism information.

Finally, additional case studies will bolster confidence in the application of GenRA in regulatory toxicology. GenRA performance on *in vivo* toxicity datasets has been systematically evaluated using standard performance metrics, such as a coefficient of determination, to understand the goodness of fit for GenRA predictions. However, a comparison with expert driven read-across cases has not been performed primarily because a database of expert-driven read-across selections is not available. A concerted effort will be made to identify read-across case studies either reported in the literature or under the auspices of other EPA or OECD activities and extract relevant information, including: the strategy taken to identify candidate source analogs, the rationale for selecting source analogs, the toxicity endpoint being predicted, the underlying toxicity data for the source analogs, and the data gap filling technique used.

2. Develop and Expand Databases Containing TSCA Chemical Information

In addition to the information submitted for a new chemical, information on other TSCA-relevant chemicals may be found in a wide variety of public sources as well as in legacy OPPT TSCA files (which may include TSCA CBI). However, many of the public sources as well as the TSCA data are not in a digital form that can be efficiently searched, analyzed, and used to develop and refine (Q)SAR models, inform the refinement and development of chemical categories, and provide data for analogs in read-across evaluations of new chemicals. ORD is proposing to continue expanding existing ORD databases and curation efforts to structure data including physicochemical and environmental fate properties (ChemProp); household product chemical composition and function (CPDat); multimedia monitoring data (MMDB); ecological hazard (ECOTOX); human health hazard (ToxVal, ToxRefDB); and toxicokinetics (CvT, HTTK). The information in the ORD databases will be mapped to available standardized reporting templates (starting with hazard data), stored in IUCLID as appropriate (see Research Area 5), and made publicly available. Literature mining tools for information retrieval and extraction will also be refined and further developed to attempt to rapidly screen the open literature and gray literature, *i.e.* information not available from commercial publishers, for relevant information on chemicals and associated analogs, with publication of the tools and approaches employed.

In addition, OPPT will plan to identify, extract, curate and catalog available data on chemistry, hazard, fate, and exposure from different TSCA databases and holdings (which may include TSCA CBI). This will include digitizing existing physical records (largely paper and some microfiche) to capture all relevant TSCA information for a given chemical substance. Available TSCA information will be combined

with publicly available sources, such as information from ORD databases, to expand the amount of information available, thereby enhancing chemical reviews and enabling efficient sharing of chemical information across EPA. Safeguards for CBI will be maintained as appropriate in this process. This initiative to digitize TSCA information contributes to a long-term goal of maintaining and utilizing fully computationally accessible data within OPPT. Based on the size of this task, ongoing work within OPPT to achieve this extends beyond the initial collaborative research plan described here.

The proposed research in Research Area 2 and Research Area 5 will result in EPA having increased interoperability between relational databases containing TSCA-relevant physicochemical properties, environmental fate/transport, hazard, and exposure information to ensure the efficient searching of existing chemical information. Additional curation efforts will expand available chemical information for developing chemical categories, (Q)SAR, and other predictive models and will enable efficient sharing of chemical information within EPA. This research builds upon ongoing data curation activities and further motivates curation and database engineering tasks to support TSCA new chemical assessments and collaborations with internal and external stakeholders that utilize harmonized data formats, particularly for hazard data (Table 3).

Table 3. StRAP Outputs Relevant to Data Curation.

Data type curated	*Relevant StRAP4 Output*	*Relevant StRAP4 Output Title*
Chemistry and properties	CSS.407.1	Generate and curate data relevant to chemical substances, structures, samples, and properties
Chemistry and properties	CSS.407.3	Develop new and improve existing structure activity relationship models to support risk assessment
Chemistry and properties	CSS.407.5	Advancing the use of structural, mechanistic, and toxicokinetic data to support categorization and classification of Per- and Polyfluoroalkyl substances (PFAS)
***In vivo* hazard in human health relevant models**	CSS.401.2	Provide structured and computationally accessible data to support tiered toxicity testing
	CSS.408.4	Strengthening the science to support new chemicals evaluation
***In vivo* hazard in ecologically relevant species**	CSS.406.3	Identify, assemble, and curate toxicity data for ecologically relevant species for risk assessment (ECOTOX)
Monitoring and release data to support exposure assessment and modeling	CSS.402.1	Collect and curate exposure-relevant data
Toxicokinetic data	CSS.402.2	High-throughput toxicokinetic (HTTK) tools to support *in vitro* to *in vivo* extrapolation

A. Chemical structure, physicochemical and environmental fate properties

Expanded curation of chemical identity, physicochemical, and environmental fate properties make more of the chemical landscape accessible for chemical category formation, read-across, and predictive modeling. As more chemicals are added to the TSCA active nonconfidential inventory, or chemistries with limited available information are identified, more structure, physicochemical, and environmental fate property data curation is needed to support decision making.

1. **Chemical identity and structure (Distributed structure-searchable toxicity, DSSTox)**: Each time a new chemical submission is reviewed, an initial step may include connecting the submitted chemical structure to existing information or to existing information for analogs (on the basis of attributes such as structure). Thus, chemical identity and structure curation that expands along with expansions of the TSCA chemical inventory will support new chemical evaluation. Ongoing work in ORD to expand and update the DSSTox database (Grulke et al., 2019) informs the CompTox Chemicals Dashboard (Williams et al., 2017) and internal chem- and bio-informatic

databases and models that may be relevant for application within the NCCRP. For instance, (Q)SAR development and surfacing of curated hazard and exposure information all rely upon accurately curated chemistry information. DSSTox chemical information spans over 1.2 million unique substances (as of July 2022) with definitions of chemical structure and with curated identifier linkages. Accurate structure-data linkages, in turn, provide the quality foundation for chemical screening projects (such as ToxCast and high-throughput screening efforts), read-across, non-targeted analysis, and structure-based modeling efforts across CSS research programs, including models useful to the NCCRP. As the TSCA active inventory of chemicals grows each year, expansion of the DSSTox database to include these chemicals as well as existing and emerging chemicals of interest for modeling applications is essential. This DSSTox expansion supports the evolving needs of the Agency and providing programmatic access to any data that can be linked to a DSSTox identifier. New chemical submissions under TSCA may be for defined or complex mixtures, and chemistry curation can provide solutions for better linking appropriate data to these mixtures to facilitate read-across or other downstream predictions. ORD efforts to curate special chemistries relevant to TSCA, such as PFAS, require manual curation using Markush-type structure representations (as exemplified by the PFAS Category list: https://comptox.epa.gov/dashboard/chemical-lists/EPAPFASCAT) or creation of linkages to defined structural components, in order to enable connection of these chemical structures to the web-accessible compendium of structure-linked chemical knowledge. In addition to providing linkage of families of chemicals that have similar structural features, chemistry curation of Markush representations helps address challenges in assigning defined structures to complex mixtures, including UVCBs (*i.e.,* chemical substances of Unknown or Variable Composition, Complex Reaction Products and Biological Materials). Indeed, UVCBs constitute a significant percentage of TSCA (*i.e.,* over 40% of the non-confidential TSCA chemical inventory list on the Dashboard, https://comptox.epa.gov/dashboard/chemical-lists/TSCA_ACTIVE_NCTI_0320, cannot be assigned defined structures). These chemistry curation efforts provide a foundational source of chemistry information to be used in research and workflow applications for the NCCRP.

2. **Physicochemical, fate, and toxicity properties:** Within ORD, the development and evaluation of (Q)SAR models for physicochemical, fate, and toxicity properties have relied on curated data from external sources. Making forward predictions of these properties for new chemical submissions may improve with more data from existing TSCA-relevant chemicals. In this

proposed research, additional property data relevant to training and test sets for (Q)SARs developed for or applied to the NCCRP will be curated. For example, (Q)SAR prediction models from the Toxicity Estimation Software Tool (TEST) were built on datasets compiled before 2012 and have training data that are distinct and generally smaller than those used to build Open Structure-Activity/Property Relationship Application (OPERA) models (Mansouri et al., 2018; Mansouri et al., 2016). Recent efforts greatly expanded available property data, including data for PFAS. As more TSCA-relevant chemicals are added to DSSTox, or as gaps in chemical category coverage are identified through Research Area 1, additional physicochemical, fate, and toxicity property data will be prioritized for curation.

B. *In vivo* hazard data

Expanded curation of *in vivo* hazard data will be performed to make more information available on potential analogs for target new chemical submissions under TSCA. For example, *in silico* approaches such as GenRA rely upon quality curation and computationally accessible *in vivo* hazard data in ToxRefDB. *In vivo* hazard data curation will be prioritized to capture chemistries of interest based on learnings from chemical categories in Research Area 1, information from OPPT, and available data. These curated data can be described as human health hazard relevant (for ToxRefDB and/or ToxValDB) or ecologically relevant (for ECOTOX). These data curation efforts are described in more detail below.

1. **Human health hazard data (ToxRefDB and ToxValDB):** In ORD, legacy *in vivo* human health hazard information is stored in two different relational databases: the Toxicity Reference Database (ToxRefDB) and the Toxicity Value Database (ToxValDB). These two databases continue to be expanded through manual, user interface-driven curation workflow and scraping of web-accessible data, respectively. As an example, TSCA-relevant chemical data may be added in ToxRefDB and/or ToxValDB via curation of full text source documents and aggregation of summarized data available on the public website, ChemView.[18] More hazard information, particularly from guideline studies or guideline-like studies, will be added based on identification of source documents or data for TSCA-relevant chemicals and resources available for curation. Expanding the chemical and study coverage for these databases for TSCA-relevant chemicals will improve *in silico* approaches, including (Q)SAR and read-across, for new chemical evaluation.

[18] https://chemview.epa.gov/chemview/

a. ToxRefDB (Watford et al., 2019) contains *in vivo* study data from over 5900 guideline or guideline-like studies for over 1100 chemicals. This is largely comprised of curated animal study data from repeat dose studies conducted according to Health Effects Series 870 guidelines, and many of these studies (over 3,000 of them) come from registrant-submitted toxicity studies known as data evaluation records (DERs) from the U.S. EPA's Office of Pesticide Programs (OPP). While employing a controlled vocabulary for enhanced data quality, ToxRefDB serves as a resource for study design, quantitative dose response, and endpoint testing status information given guideline specifications. To enable high quality and consistent extractions, the document extraction is performed by 1-2 study curators with manager review within the Data Collection Tool (DCT). ToxRefDB is summarized with calculated point-of-departure values at the chemical and study level for inclusion in the summary-level database, ToxValDB.

b. ToxValDB includes data on thousands of chemicals from tens of thousands of records, with an emphasis on quantitative estimates of relevant points-of-departure from *in vivo* toxicology studies, such as no- and low-observable adverse effect levels, screening levels, reference doses, tolerable daily intake, etc. The source data originates from multiple public datasets, databases (*i.e.*, with data already digitized), and the open literature. Each dataset is reshaped to a standard source data format and then all source data streams are integrated into the main ToxValDB database. Data is reviewed for quality within the source data tables. In addition to the main *in vivo* quantitative data, ToxValDB also contains data on cancer slope factors, genotoxicity assays, and acute toxicity information (*e.g.*, skin and eye irritation and skin sensitization). All data is surfaced on the CompTox Chemicals Dashboard.[19]

2. **Ecologically-relevant hazard data (ECOTOXicology Knowledgebase, ECOTOX):** ECOTOX (Olker et al., 2022) is a comprehensive, publicly available knowledgebase providing single chemical environmental toxicity data on aquatic life, terrestrial plants, and wildlife. Hazard data are extracted and added to ECOTOX quarterly and improvement and expansion of the ECOTOX controlled vocabulary is ongoing. Addition of data to ECOTOX may include TSCA-relevant chemicals, as well as emerging contaminants, chemicals associated with assessment of Endangered Species Act concerns, and/or chemicals of interest to the Office of Pesticide

[19] CompTox Chemicals Dashboard: https://comptox.epa.gov/dashboard/

Programs. Additional curation of TSCA-relevant chemicals will enhance opportunities for read-across and augmentation of training data in (Q)SAR models for aquatic toxicity. ECOTOX has its own user interface[20] and is also surfaced on the CompTox Chemicals Dashboard.

C. Exposure data

Expanding curation of the use context for chemicals supports development of quantitative structure use relationships that may be important for predicting potential uses of new chemicals. Additional curation of chemical occurrence in environmental matrices can inform predictions of chemical release to environmental compartments for data-poor chemicals. And finally, curated *in vivo* toxicokinetic data can support evaluations of confidence in high-throughput toxicokinetic (HTTK) modeling and curated intrinsic clearance and plasma protein binding data can further inform generic toxicokinetic modeling approaches applied for *in vitro* to *in vivo* extrapolation (IVIVE).

1. **Chemicals and Products Database (CPDat):** How a chemical is used in consumer, occupational, or industrial context(s) determines relevant exposure pathways for risk evaluation. As a result, exposure research relies on well-documented and accessible datasets of chemical use information curated from publicly available sources. These curated chemical use data can support key needs of new chemical assessments under TSCA with respect to development and refinement of systematic approaches for predicting potential conditions of use for submitted substances (*e.g.,* via the development of Quantitative Structure Use Relationship, or (Q)SUR, models) (Phillips et al., 2017); estimating exposure via poorly characterized release scenarios; characterizing variability in population exposures; and, addressing data-poor chemicals by creating generic exposure scenarios by use. ORD developed a data management and curation application, Factotum, to facilitate the rapid collection and distribution of high-quality chemical and exposure-related data from public documents via curation, quality assurance, visualization, and data delivery tools, which are released as the Chemicals and Products Database (Dionisio et al., 2018; USEPA, 2020a). Factotum has also enabled manual and machine-learning based curation of data to new or updated harmonized chemical use tags, product categories, and OECD function categories. Additional planned curation efforts will augment consumer product information curation with data on the occurrence of data-poor chemistries, including PFAS and

[20]ECOTOX user interface: https://cfpub.epa.gov/ecotox/

UVCBs, in products within the home by analyzing data from measures of presence and emissivity of these chemicals from household products and articles. The corpus of curated chemical use information will help fill gaps in inferring exposure scenarios for new chemicals.

2. **Environmental occurrence (Multimedia Monitoring Database, MMDB):** Chemical monitoring data are the gold-standard for exposure data in risk assessment, and curation of additional chemical monitoring data would inform predictive models of potential environmental. A recent OPPT and ORD collaboration produced a machine-readable database, MMDB, which comprises chemical measurements in environmental media collected and curated from publicly available government databases and reports as well as additional media occurrence data (in consumer articles, indoor air and dust, and biological media) curated from scientific literature (Isaacs et al., 2022; USEPA, 2022b). MMDB requires standardization and aggregation of monitoring data for each media to harmonized units, where possible. Planned curation work potentially includes, but is not limited to, quantitative or qualitative non-targeted analysis data from EPA or other studies; new data on chemicals measured in biosolids; PCB concentrations in consumer products; residential exposure data from EPA field studies; and monitoring data for PFAS compounds obtained from extracted reports or other sources. Currently, these data are considered public, but are generally inaccessible. A number of these regulatory monitoring data streams are not currently in MMDB and will provide a unique and valuable contribution to the database. In collaboration with the Center for Environmental Solutions and Emergency Response (CESER), this product will also produce a production-quality database (tentatively called StEWIDB) to integrate their Standard Emissions and Waste Inventories (StEWI) with Factotum (for example, via migration to an updated open-source SQL system and development of workflows for chemical curation) for delivery of release data to internal or external stakeholders via existing ORD tools. The environmental occurrence data curated into these databases (MMDB, Factotum, and StEWIDB) will inform the development of much-needed predictive models of chemical releases into environmental compartments assessment of TSCA-relevant chemicals.
3. **Toxicokinetic data for internal exposure (Concentration versus Time Database, CvTdb, and high-throughput toxicokinetics, HTTK):** Toxicokinetic data when combined with generic toxicokinetic models provide a quantitative link between potential human exposures and bioactive concentrations in *in vitro* screening. In addition, toxicokinetic data has the potential to inform analog selection (*e.g.*, selection of analogs with similar toxicokinetic properties) and

refine chemical categories based on similar toxicokinetic properties among chemicals. Data curated in CvTdb (*in vivo* tissue Concentration vs. Time database) enable quantitative evaluation of confidence of HTTK predictions for many chemicals and provides data to refine generic HTTK models (Sayre et al., 2020). Work planned in StRAP4 includes expansion of curated data for key exposure routes (*e.g.*, dermal and inhalation exposures) for TSCA-relevant chemicals that will be useful in evaluating the generic HTTK models of dermal and inhalation exposures. Carefully quantifying the chemical- and scenario-specific uncertainty in data and models allows decision makers to consider the use of HTTK for IVIVE, as well as any other situation where extrapolation using TK may be useful. This task is not possible without generating and continually expanding the CvTdb resource, particularly to include routes of exposure and chemistries relevant to TSCA for the NCCRP.

3. Develop and Refine (Q)SAR and Predictive Models for Physicochemical Properties, Environmental Fate/Transport, Hazard, Exposure, and Toxicokinetics

OPPT has developed and applied a large suite of (Q)SAR and other predictive models, including expert systems, to estimate physicochemical properties, exposure, environmental fate/transport, carcinogenic hazard, and ecological hazard.[21] OPPT and ORD are proposing to update and/or improve existing OPPT (Q)SAR and predictive models and enable regular model updates. The data used to develop and update (Q)SAR and predictive models will be derived from the curated public and TSCA databases described above in Research Area 2. ORD and OPPT will evaluate all appropriate models, including evaluation of the data used to build models and model performance against measured data, to ultimately determine the best suite of models for use by OPPT for regulatory purposes.[22]

The OECD principles for validation of (Q)SAR (OECD, 2007) suggest that consideration of (Q)SARs for regulatory purposes is facilitated by association of the (Q)SAR with information regarding the defined endpoint predicted; a clear (and reproducible) algorithm; a defined domain of applicability; appropriate measures of its goodness-of-fit, robustness, and predictivity; and, a mechanistic interpretation when possible. ORD continues to pursue model development aligned with these principles, including use of computational approaches such that updated model training and test data, critical to defining the

[21] *Predictive Models and Tools for Assessing Chemicals under TSCA*

[22] Using the OECD principles for validating QSAR models (see OECD 2007)

applicability domain and predictivity, can be applied more rapidly and systematically. The ability to rapidly extend the model training and test data to include more TSCA-relevant chemicals may increase the applicability domain, thereby increasing the relevance of (Q)SAR predictions for new chemical evaluations. In addition to extending model training and test data, evaluating applicability domain, and updated reporting on model predictivity, ORD may also incorporate newer machine learning-based approaches in model development. Whenever practicable, these new (Q)SARs will be contextualized with existing models to build confidence and increased understanding across approaches. This will include working with other stakeholders and peer reviewers to build confidence that the models related to hazard or toxicokinetics meet the TSCA statutory requirement of Section 4(h)(1)(B)(i) to encourage and facilitate "the use of scientifically valid test methods and strategies that reduce or replace the use of vertebrate animals while providing information of equivalent or better scientific quality and relevance that will support regulatory decisions…."

The goal of this effort is to update the models to reflect the best available science, increase transparency, and establish a process for updating these models as science allows. This will enhance the capabilities of OPPT to perform risk assessments for new chemicals. In addition, refining and developing such tools will lead to their use by EPA, submitters, and other stakeholders in designing safer chemicals, and will build confidence in their use for regulatory purposes. This research builds upon ORD efforts to predict properties on the basis of structure for a variety of applications (Table 4).

Table 4. StRAP Outputs Relevant to (Q)SAR and Prediction

Properties predicted	*Relevant StRAP4 Output*	*Relevant StRAP4 Output Title*
Chemistry, hazard, and properties	CSS.407.3	Develop new and improve existing structure activity relationship models to support risk assessment
Functional use and exposure	CSS.402.3	Refine exposure models that enable high-throughput exposure predictions for chemicals
Functional use and exposure	CSS.408.4	Strengthening the science to support new chemicals evaluation
Repeat dose point-of-departure	CSS.408.4	

A. Informatics platform for (Q)SAR development, implementation, and data management

Work in StRAP4 will extend ongoing ORD research on (Q)SAR modeling to inform chemical evaluation, using a foundation of modeling best practices (*e.g.*, the OECD (Q)SAR framework), updated

methodologies, and input datasets. Where feasible, the predictive performance of resultant models will be compared with current, peer-reviewed models being used by OPPT (*e.g.*, EPI Suite, ECOSAR, OncoLogic, OECD (Q)SAR Toolbox structure-based profilers)[23] to ensure fit for purpose application and build confidence in newer (Q)SARs. In addition, there is a need to establish cheminformatics approaches for model management and versioning to enable real-time model predictions, data provenance, and long-term sustainability to include ability to update, manage versions, and reproduce (Q)SAR values. Such cheminformatics approaches also include development of automated workflows to transform raw experimental data to modeling data sets and then optimize and streamline (Q)SAR model development.

Currently, WebTEST1.0 is accessible through the CompTox Chemicals Dashboard for prediction[24] of a number of physicochemical and toxicological properties, such as oral rat acute 50% lethal dose. These predictions are accessible by locating a structure in the CompTox Chemicals Dashboard or by drawing it for real-time prediction. These models and their performance reports, including how the predicted chemical fits within the applicability domain (in line with OECD (Q)SAR validation principles), are available through the user interface. The cheminformatics approaches to be developed in StRAP4 represent ongoing and planned work to develop WebTEST2.0. This updated version of WebTEST will include access to raw experimental data, modeling datasets, molecular descriptor values, and (Q)SAR models. Each (Q)SAR model is associated with a versioned data set, (Q)SAR methodology, and molecular descriptor set so that the predictions are reproducible. The new WebTEST2.0 workflow can be used to develop models using Python-based machine learning methods such as random forest and support vector machines, all within the WebTEST platform.

The impact of the work is that users will have access to real-time predictions from a large array of (Q)SAR models, as well as other models run at defined intervals, in a single website. WebTEST2.0 predictions will be linked to an HTML report which indicates whether the chemical is within the applicability domain of the models and provides prediction results for structurally similar chemicals from the training and test sets. WebTEST 2.0 also provides extensive documentation on each (Q)SAR model. For each property (and associated dataset), the tool will link to a spreadsheet which provides the training and test set statistics, the training and test sets, and the test set prediction results for each model. For each model, the tool will provide a PDF document in the (Q)SAR model reporting format that outlines all the details of the model. Having a (Q)SAR model reporting format document is often a

[23] Predictive Models and Tools for Assessing Chemicals under TSCA
[24] https://comptox.epa.gov/dashboard/predictions

requirement for using models for regulatory applications. Models for physicochemical properties (*e.g.*, octanol water partition coefficient (logKow), vapor pressure, and Henry's law constant) are being developed using the WebTEST2.0 workflow. Revised toxicity models will be developed by expanding the toxicity datasets for WebTEST1.0 (*e.g.*, acute aquatic toxicity). In addition, models will be developed for additional toxicity endpoints (*e.g.*, carcinogenicity, repeat dose toxicity, skin sensitization) to support TSCA new chemical evaluations.

The architecture of WebTEST2.0 will also be extended to allow for incorporation of models which are developed external to WebTEST2.0 using a different workflow. For example, a specialized fish (Q)SAR model will be developed which is based on chemical classes obtained from the ClassyFire webservice (Djoumbou Feunang et al., 2016). This model is essentially a more advanced version of the fish toxicity model in EPI Suite. Models developed outside of the WebTEST platform will be implemented via Docker containers or via API calls to external webservices. OPERA, EPI Suite, and WebTEST1.0 models will be incorporated into WebTEST2.0 via webservices. Additionally, bioactivity-based models for estrogen receptor (Judson et al., 2017; Judson et al., 2015), androgen receptor (Judson et al., 2020; Kleinstreuer et al., 2017), steroidogenesis (Haggard et al., 2018; Haggard et al., 2019), and potentially other bioactivities based on *in vitro* NAM data, will be included in the WebTEST2.0 model registration platform. Registration of all models, regardless of their development within or outside of the WebTEST platform, will include meta-data on the input features used in the modeling, the model output, and version information about that model; this constitutes an important goal for WebTEST2.0 and for rapid integration of information from disparate sources for next generation risk assessment.

B. Exposure predictions

Review of new chemical submissions under TSCA includes both engineering assessment and exposure assessment, including estimation of chemical releases to the workplace and environment based on the chemical "conditions of use," as well as estimation of resulting occupational, general population (ambient), and consumer exposures. In many cases, these estimates also include site-specific exposures. While the process and exposure models used in these evaluations are well-defined and peer-reviewed, there are several areas where ORD research efforts could improve or expand current evaluation workflows (Wambaugh et al., 2019). The research efforts include systematic approaches for identifying potential uses for submitted substances, estimating exposure via poorly characterized release scenarios (*e.g.*, down-the-drain consumer or industrial releases), characterizing variability and uncertainty (which could inform identification of highly exposed populations), and

addressing data-poor chemicals (*e.g.*, those for which no default information is available for parameterizing likely exposure scenarios). Methods for characterizing potential conditions of use may utilize expanded functional use databases (to better cover the use space for known chemical analogs) and refined (Quantitative) Structure Use Relationship ((Q)SUR) models that can consider a chemical functional role within a particular sector (*e.g.*, consumer versus industrial use) or specific industry or product category. In addition, ORD research efforts will improve the flexibility of currently available models to better characterize a wider range of exposure scenarios, including the consideration of spatial scale and environmental justice concerns. ORD is planning research to deliver new computational models, workflows, and datasets to support problem formulation (identification of exposure scenarios relevant to a particular chemical) and/or quantitative modeling related to exposure assessment in the PMN process. Where feasible, these newer research tools will be compared to the current approaches and models used by OPPT.

Though not currently used in assessment of new chemical submissions under TSCA, ORD is engaged in ongoing development of refined consensus exposure models that can use chemical structure representations to predict: 1) human daily chemical intake rates; 2) air concentrations in occupational settings; and 3) surface water concentrations. These models are examples of EPA's systematic empirical evaluation of models (SEEM) meta-model approach (Ring et al., 2019) and integrate available predictors from multiple exposure pathway models with available monitoring data to produce predictions for novel chemical structures. These models can provide screening-level exposure estimates (with associated uncertainty) for data-poor chemicals. When completed, SEEM exposure prediction models will be integrated with the WebTEST2.0 cheminformatics platform as models developed outside of the WebTEST platform.

C. Toxicokinetic predictions

ORD has developed a library of empirically measured HTTK parameters (that is, *in vitro* toxicokinetic measurements) for >1000 chemicals. However, with tens of thousands of chemicals that may be of interest, ORD has been engaged in development of a series of quantitative structure-property relationship models for predicting key toxicokinetic parameters, including fraction of chemical unbound to protein in plasma, and hepatic metabolic clearance (Dawson et al., 2021; Pradeep et al., 2020; Sipes et al., 2017). Work is underway to compare different *in silico* methods for toxicokinetic parameter prediction, and these approaches will be applied to TSCA-relevant chemicals to make IVIVE possible even without experimental data for toxicokinetic parameters to inform generic toxicokinetic models. Further, predicted toxicokinetic parameters inform models of bioavailability and dermal absorption and

may refine chemical categories and analog selection (*i.e.*, considering neighboring chemicals with similar toxicokinetic properties). The toxicokinetic parameter predictions, and the subsequent generic HTTK model predictions, will be integrated with the WebTEST2.0 cheminformatics platform as models developed outside of the WebTEST platform.

4. Explore Ways to Integrate and Apply *In Vitro* NAMs in New Chemical Assessments

Section 4(h) of TSCA promotes reducing testing on vertebrate animals and sets forth some requirements for NAMs. As recognized in OPPT's Strategic Plan,[25] leveraging *in vitro* NAMs to generate mechanistic, hazard, and toxicokinetic data may further inform data gap filling approaches for new chemicals. As required under TSCA 4(h), OPPT maintains a list of NAMs that are scientifically reliable, relevant, and capable of providing information of equivalent or better scientific reliability and quality to that which would be obtained from vertebrate testing.[26]

EPA and the broader scientific community have invested heavily in the development of *in vitro* NAMs. As part of the NCCRP, OPPT is proposing to take advantage of previous and ongoing research in ORD and by other partners that have identified important biological targets representing potential hazards, improved estimates of dose extrapolation from *in vitro* systems, incorporated routes that are key to highly exposed populations (*e.g.*, inhalation and dermal exposure), and continued to develop resource effective technologies that broadly characterize biological activities across pathways, processes, and different cell types. The *in vitro* NAMs applied within the NCCRP will be evaluated for reliability and relevance for new chemical evaluation. Fit-for-purpose application of NAMs will rely, to the extent possible, on the concepts of (1) adverse outcome pathways (AOPs) and the key events leading to toxicity; and (2) Integrated Approaches to Testing and Assessment (IATA) for weight of evidence evaluation and the use of Defined Approaches.[27] Although some informative NAMs may not be associated with an IATA or Defined Approach, and some health outcomes do not have established AOPs, this would not prevent OPPT from applying these methods if they represent the best available science.

The proposed effort is intended to develop a suite of accepted, fit-for-purpose NAMs that could be used by external stakeholders for testing and data submissions under TSCA as well as informing and expanding new chemical categories. In this Research Area, ORD will collect *in vitro* NAM data to

[25] See published plans by OPPT under TSCA (2018) and by EPA for the Agency (2021).
[26] See TSCA Section 4(h) NAM list
[27] AOP; see G Patlewicz et al. (2015). Proposing a scientific confidence framework to help support the application of adverse outcome pathways for regulatory purposes. Regul. Toxicol. Pharmacol. 71(3):463-77. doi: 10.1016/j.yrtph.2015.02.01. IATA - see IATA and Defined Approaches (OECD 2017).

demonstrate how NAMs for bioactivity and toxicokinetics can be used in a NAM-informed assessment of data-poor chemicals. The NCCRP presents the opportunity for ORD to make a leap in progress on prospective application of a screening strategy for hazard (Thomas et al. 2019) (see Figure 3 for outline of screening work planned in StRAP4). In a first step, ORD will focus on development of a dataset for 200-300 chemicals, including some reference chemicals as well as TSCA-relevant chemicals from the nonconfidential inventory, to increase scientific confidence in application of this suite of bioactivity NAMs for informing chemical safety. These data will be needed to evaluate performance of these NAMs for further application and may also inform evolving frameworks for using multiple data streams to inform bioactivity-based dose-response assessment and hazard identification. Pending additional infusion of resources, bioactivity screening could be extended to additional chemicals, which is a necessary component of using bioactivity for analog selection or informing putative chemical categories that cover a large TSCA-relevant chemical universe.

A cheminformatics step will identify candidate chemicals for screening from the TSCA active inventory, including examination of: chemical structural and physicochemical diversity to promote coverage of putative chemical categories; amenability to aqueous based-screening or potential volatility or aerosolize-ability; ability to procure the chemical in sufficient quantities for screening; and, chemicals of interest with respect to current gaps in bioactivity and/or (Q)SAR model data sets. Following chemical selection, a set of 200-300 chemicals amenable to aqueous screening will be screened in both broad and targeted biological screening technologies for human health relevant endpoints, and a subset of these chemicals will be tested in broad screens with ecological relevance. “Broad” screening refers to technologies such as high-throughput phenotypic profiling (HTPP) and high-throughput transcriptomics (HTTr) that characterize the biological activity of chemicals using highly-multiplexed measurements of many different cellular features or transcripts, respectively, thereby capturing chemical effects that may result from specific interactions with molecular initiating event (MIE) targets as indicated by fingerprints or signatures indicative of those MIEs, as well as effects that may result from generalized cellular responses to stress or activity at multiple MIEs (Harrill et al., 2021; Nyffeler et al., 2020). These broad screening modalities inform both estimates of *in vitro* bioactivity-based dose-response assessment and identify whether a chemical may act at specific MIEs or non-specifically at many targets (Thomas et al., 2019). Targeted screening complements broad screening by providing information about MIEs, key events, or other processes related to hazards of interest. Broad and targeted screens are combined in ORD work to support the NCCRP for human and ecological health as well as for acute portal of entry effects from potential inhalation exposures. A smaller subset of chemicals with inhalation exposure

potential (8-10 chemicals per year) will be screened in NAMs for inhalation exposure that include both broad and targeted measures (Figure 3 for outline).

Figure 3. Overview of an initial in vitro NAM screening strategy.

Cheminformatic approaches will be used to categorize the structural and physicochemical diversity of the TSCA active inventory and identify potential chemical screening candidates; these candidates will be further refined to those that are procurable and amenable to aqueous-based screening in cell-based and cell-free assays (including physicochemical property evaluation and non-volatility) or cell-based models of inhalation with an air:liquid interface. Candidate chemicals would also fill gaps in available *in vitro* or *in silico* screening information (*e.g.,* filling gaps in the applicability domain for ECOSAR). Broad and targeted screening covering human and ecological health and inhalation exposure will be applied. Broad screens are represented by HTPP also known as Cell painting for human health, tiered application of modified ecotoxicology studies and HTTr for ecotoxicology (EcoHTTr), and HTTr applied to multiple cell models of the human respiratory tract in an air:liquid interface system. Targeted screens are represented by a safety pharmacology panel (Safety Pharm), a DevTox Germ Layer Reporter assay, assays for genotoxicity, and assay data to inform high-throughput toxicokinetics (HTTK). Phenotypic measures in models of human respiratory tract also represented targeted screens conducted in parallel to HTTr in the inhalation exposure system. Inverted triangles represent "funnels" to indicate relative numbers of chemicals to be screened in the initial strategy out of the 200-300 chemicals selected.

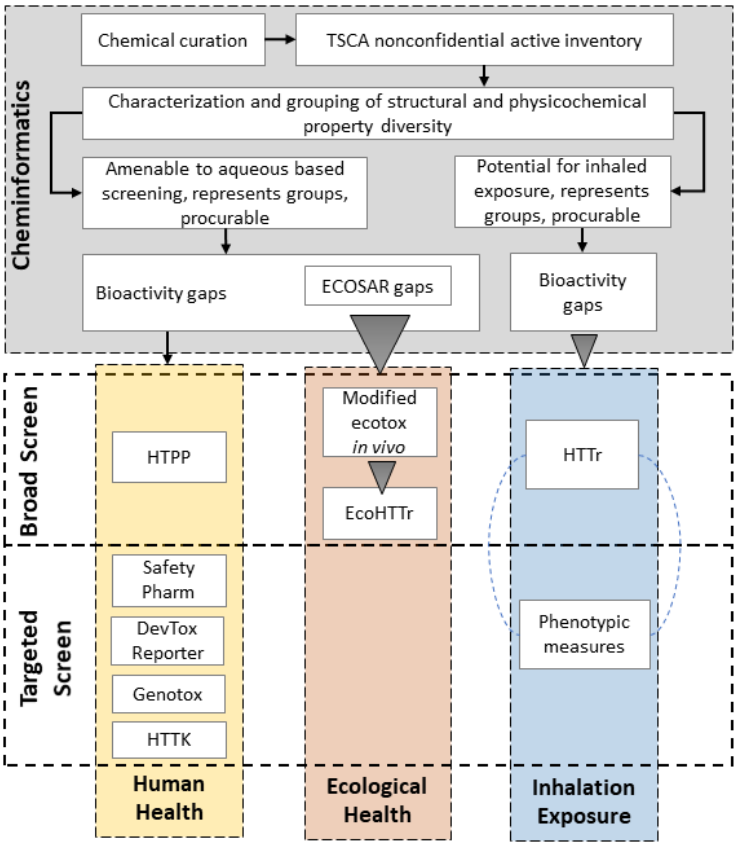


Table 5. StRAP Outputs Relevant to In Vitro NAM Data Generation for NCCRP

Bioactivity screening planned	*Relevant StRAP4 Output*	*Relevant StRAP4 Output Title*
Analytical quality control	CSS.408.4	Strengthening the science to support new chemicals evaluation
Developmental toxicity screening		
Acute portal of entry effects of inhaled exposures		
Toxicokinetic NAMs (hepatic clearance and fraction unbound)		
Broad screening (high-throughput phenotypic profiling; safety pharmacology)	CSS.408.4, *CSS.401.1*	*Additional reliance on: Advance a tiered, high-throughput toxicity testing strategy*
Ecological hazard screening	CSS.408.4, *CSS.406.5*	*Additional reliance on: Improve ecological methods and models for predicting exposure, accumulation, and effects of PFAS*

A. Analytical quality control of chemicals

For aqueous-based screening assays for human and ecological health, ORD will procure chemical samples, solubilize in dimethyl sulfoxide, and characterize the identity and purity of the samples using liquid chromatography or gas chromatography, as appropriate, with tandem mass spectrometry. This analytical quality control step will ensure that the learnings from the 200-300 chemical set screened in the cell-based and cell-free assays will be more interpretable with respect to the nominal concentration and identity of chemicals screened. For volatilized samples used in the inhalation exposure model, air samples are collected just before reaching the exposure chamber by syringe and directly transferred to a gas chromatograph for characterization of the chemical and its concentration.

B. Screening for human health

Broad-based biological screens will be employed to provide insight into as many putative chemical-by-biological target interactions as possible, as suggested by the use of Tier 1 NAMs in the CompTox BluePrint (Thomas et al., 2019). These broad profiling assay data could potentially provide a basis for

derivation of a bioactivity-based point of departure (POD) (Baltazar et al., 2020; Nyffeler et al., 2020; Paul Friedman et al., 2020) as well as coverage of specific MIEs, indicated by the fingerprint or pattern of behavior in these assays, that may be of interest for ascertaining the need for additional hazard information. One such Tier 1 NAM, high-throughput phenotypic profiling (HTPP), also known as Cell Painting, measures more than 1000 cellular morphological features using high content imaging to inform a quantitative POD, putative molecular targets or molecular initiating events, and bioactivity fingerprints that could be used to evaluate similarity in biological effects measured by this assay (Nyffeler et al., 2021; Nyffeler et al., 2022; Nyffeler et al., 2020; Willis et al., 2020).

Targeted NAMs for hazard will also be important to informing human safety assessment gaps left by broad profiling NAMs. Safety pharmacology has been employed to detect off-target interactions in drug safety (Bowes et al., 2012; Papoian et al., 2017; Smit et al., 2021) and cosmetic products under a next generation risk assessment framework (Baltazar et al., 2020). There appears to be consensus that a broad and diverse panel of pharmacological targets can be useful information for drug safety assessment (Ietswaart et al., 2020; Valentin et al., 2018). A similar panel of receptor, ion channel, transporter, and enzyme assays has been employed previously within the larger set of assays included in the US EPA ToxCast program (Sipes et al., 2013). HTPP in multiple cell lines, each expressing a different suite of potential molecular targets, will be combined with a large panel of nuclear receptor, G-protein-coupled receptor, transporter, ion channel, and enzymatic target assays to ensure coverage of MIEs known to be of interest for toxicology.

Another important hazard gap is genotoxicity; commercially available Ames and *in vitro* micronucleus assays will be employed to evaluate genotoxicity potential of the 200-300 chemical set. Developmental toxicity potential is rarely evaluated in data submitted or associated with data-poor new chemical submissions to OPPT. To evaluate developmental toxicity potential, ORD will apply an assay adaptation of the human pluripotent stem cell test (Kameoka et al., 2014), which identifies potential developmental toxicants using a biomarker of early embryonic development. The assay adaptation employs the human RUES2-GLR stem cell line that has been engineered with a fluorescent reporter for the SOX17 biomarker to monitor differentiation of the endoderm germ layer. Recently, the assay protocol has been optimized to a 384-well plate format with shortened exposure duration to enable more rapid high-throughput screening (Gamble et al., 2022). The assay demonstrates the ability to distinguish assay negatives such as acetaminophen from known developmental toxicants such as thalidomide. Applying this cell-based developmental toxicity germ layer reporter assay (DevTox GLR-

Endo assay) to a 200-300 chemical set would provide the data necessary to demonstrate utility not only for the NCCRP and needs of TSCA but also for broader acceptance and validation. The *in vitro* developmental toxicity data, combined with (Q)SAR, read-across, and chemical categories, may provide information that is typically not available from *in vivo* studies for new chemical submissions under TSCA.

In vitro toxicokinetic assays for metabolic intrinsic clearance and plasma protein binding will also be conducted for the 200-300 chemical data set. As mentioned above in Research Area 3, empirical (and *in silico*) toxicokinetic parameters can be used to inform generic HTTK models for IVIVE, predictions of bioavailability and dermal absorption, and potentially, as inputs into characterization of chemical categories or analogs that might share similar toxicokinetic properties. Combining *in vitro* assay data from these and similar assay platforms with *in vitro* distribution and toxicokinetic modeling will increase the utility of these data by providing improved estimates of doses that correspond to *in vitro* bioactivity in exposure units that can be compared to *in vivo* dose estimates or exposure estimates (Honda et al., 2019; Klaren et al., 2019; Ring et al., 2021).

C. Screening for ecological health

Ecological hazard for new chemicals has focused on potential toxicity to three representative groups of aquatic organisms: fish, invertebrates, and plants/algae. Data for these three taxonomically diverse groups is often lacking for new chemical submissions, and as such, new chemical assessments have relied heavily on read-across to structural analogs with available toxicity estimates or (Q)SARs such as ECOSAR that rely largely on physicochemical properties of the new chemical such as its relative lipophilicity. ORD proposes to work with OPPT to identify a set of up to 60 chemicals that represent five chemical structural domains of interest for which ecological toxicity data and/or understanding of the applicability of currently (Q)SAR models is limited. Selected chemicals will be screened using higher-throughput adaptations of guideline *in vivo* ecotoxicity assays to estimate no and lowest observable effect concentrations for fish and invertebrates, 50% lethal concentration for fish and invertebrates, as well as a 50% effect concentration for algae for comparison with the values predicted by ECOSAR. For cases where current (Q)SAR approaches do not appear predictive, a smaller subset of the chemicals would then be selected for additional screening using ecological high throughput transcriptomics (Eco-HTTr) assays with fish, invertebrates, and/or plants as relevant. Data would be used to both derive a transcriptomics-based POD and support mechanistic inference for the tested chemicals.

D. Screening for inhalation exposures

For chemicals with potential inhalation exposures, specialized cellular architectures to more closely recapitulate human biology are needed, but the throughput on these systems currently prevents screening large numbers of chemicals. An engineering and exposure assessment for a data-poor substance under TSCA may suggest the potential for inhalation exposure, but frequently limited to no inhalation data may be available for assessing hazard of these exposures for the target chemical. Read-across, category-based approaches, and adaptation of the threshold of toxicological concern for inhalation toxicity can be leveraged to understand hazard based on inhalation, but all of these approaches rely on existing information on well-characterized chemicals. ORD plans to continue developing and applying *in vitro* methods for evaluating acute portal of entry effects of inhaled chemical exposures as part of the NCCRP to demonstrate the utility of this approach and its transferability (Zavala et al., 2017; Zavala et al., 2018). The prospective view is that ORD would screen 8-10 chemicals per year in multiple cell models of the human respiratory tract, pending sufficient resources, and that these data would inform read-across and other *in silico* approaches that would extend the impact of limited empirical screening. These studies combine ORD-developed *in vitro* exposure technology (the Cell Culture Exposure System), organotypic *in vitro* cell-based models of the human respiratory tract, and an assay battery that evaluates phenotypically relevant endpoints (i.e., targeted screens) such as cytotoxicity, cellular metabolism, epithelial barrier function, mucin production, ciliary beat frequency, and cytokine production. In addition, high-throughput transcriptomic (HTTr) evaluation in models of human respiratory tract represents a broad screening technology to identify biological pathway activity and a benchmark concentration, increasing the richness of the information and sensitivity for acute portal of entry effects (Speen et al., 2022). Importantly, generation of this dataset for additional TSCA-relevant chemicals, including PFAS that have not been tested in an *in vitro* air-liquid interface system previously, with potential inhaled exposures will inform guidance to support deployment of this *in vitro* NAM for chemical evaluations by external partners and stakeholders.

E. Additional bioactivity data

Additional bioactivity data may be generated in other endeavors within the StRAP, including HTTr screening work in other ORD CSS research under High-throughput Toxicity Testing (CSS 401.1 and CSS 401.2, see Figure 2), for the 200-300 chemical case study, depending on resources allocated. When additional bioactivity data are available for these 200-300 chemicals from other assays, work to

demonstrate the translation of these data to meaningful biological or quantitative POD estimates may be informative for future NCCRP work products.

5. Develop a TSCA New Chemicals Decision Support Tool to Modernize the Process

Within OPPT, searching, collating, and integrating data on new chemicals is inefficient and hinders the timeliness of decision-making. The international regulatory community has been moving towards using IUCLID to capture, store, maintain, and exchange data on intrinsic and hazard properties of chemical substances. Data in IUCLID require standardized reporting templates; for many data types, these reporting templates are consistent with internationally accepted test guidelines. ORD is proposing to use IUCLID to capture, store, and maintain publicly available data on intrinsic and hazard properties and exposure-related data. These efforts will promote data interoperability between OPPT, ORD, and other stakeholders.

Available digitized data for TSCA chemicals is important for delivery of a decision support tool that integrates all the data streams (*e.g.*, chemistry, fate, exposure, hazard) for risk assessment and transparently documents the decisions and assumptions made by expert users based on available information. This will facilitate NCD tracking decisions over time and evaluating consistency within and across chemistries. OPPT and ORD propose to collaborate on identifying the appropriate content and workflow for such a decision support tool. For example, the proposed decision support tool may allow expert chemists to examine the types of data available for analogs for a target. Information on chemical categories and/or analogs, and estimates of physicochemical properties, environmental fate, hazard, and toxicokinetics generated from predictive and *in vitro* models, will be included in the decision support tool, thereby limiting time spent on manual searching, compiling, and contextualizing available information and enabling more rapid and reproducible decisions over time.

The work to support Research Area 5 has three main components: (1) OPPT and ORD collaboratively working to increase the amount of computationally accessible CBI data previously submitted to OPPT for use within a CBI-protected environment; (2) ORD bringing computationally accessible and public data, some of which is already in ORD databases, into IUCLID-compatible formats that enable collation of these data with other data in IUCLID format; and, (3) development of a proof-of-concept decision support tool that consolidates as much traditional and NAM data as possible into a single workflow to inform assessments and documentation of selections and assumptions, such as data for decisions. This work, spanning multiple StRAP Outputs (Table 6), is intended to begin addressing the

challenge OPPT NCD faces in searching for, collating, and integrating data for a new chemical assessment.

Table 6. StRAP Outputs Relevant to IUCLID and a NCCRP decision support tool

Relevant StRAP4 Output	***Relevant StRAP4 Output Title***
CSS.408.1	Integrating data systems to enable knowledge delivery
CSS.408.3	Cross-disciplinary integration and applied case studies to support chemical safety decision making
CSS.408.4	Strengthening the science to support new chemicals evaluation

A. Implementing the International Uniform Chemical Information Database (IUCLID) in ORD

Regulatory authorities including the European Chemicals Agency, the European Food Safety Agency, and Health Canada employ IUCLID database formatting, which uses standardized reporting templates to manage chemical data (*i.e.*, Organisation for Economic Co-operation and Development [OECD] Harmonized Templates, known as OHTs) (OECD, 2021). IUCLID is an international effort with dedicated resources to manage, update, and develop tools around IUCLID, such as the Data Uploader tool to convert data to IUCLID format. ORD has extensive ongoing data curation efforts in customized database schemas to capture research-oriented levels of detail. In support of NCCRP as well as ongoing internal and external collaboration, wherever possible based on existing OHTs, these databases will be mapped to OHTs for conversion to IUCLID format, with initial priorities on physicochemical properties and human and ecological health data. A stable computing environment for IUCLID will be established in ORD. Development of automated processing of data in the ORD research environment to IUCLID will enhance ORD's ability to use these public IUCLID data in a decision support tool for new chemicals as well as other modeling applications with internal and external stakeholders. Further, ORD would be able to transfer IUCLID formatted public data to operate within the TSCA CBI instance of IUCLID, such that a combined IUCLID dataset could supply one of the data streams for a proof-of-concept decision support tool.

B. Collaboration between ORD and OPPT on IUCLID data

Though OPPT is able to receive data in IUCLID, historical data and the majority of incoming data are not reported in OHT formats and may exist as documents or digitized data stored in disparate locations. Modernizing the new chemical assessment process involves being able to rapidly exchange data, which

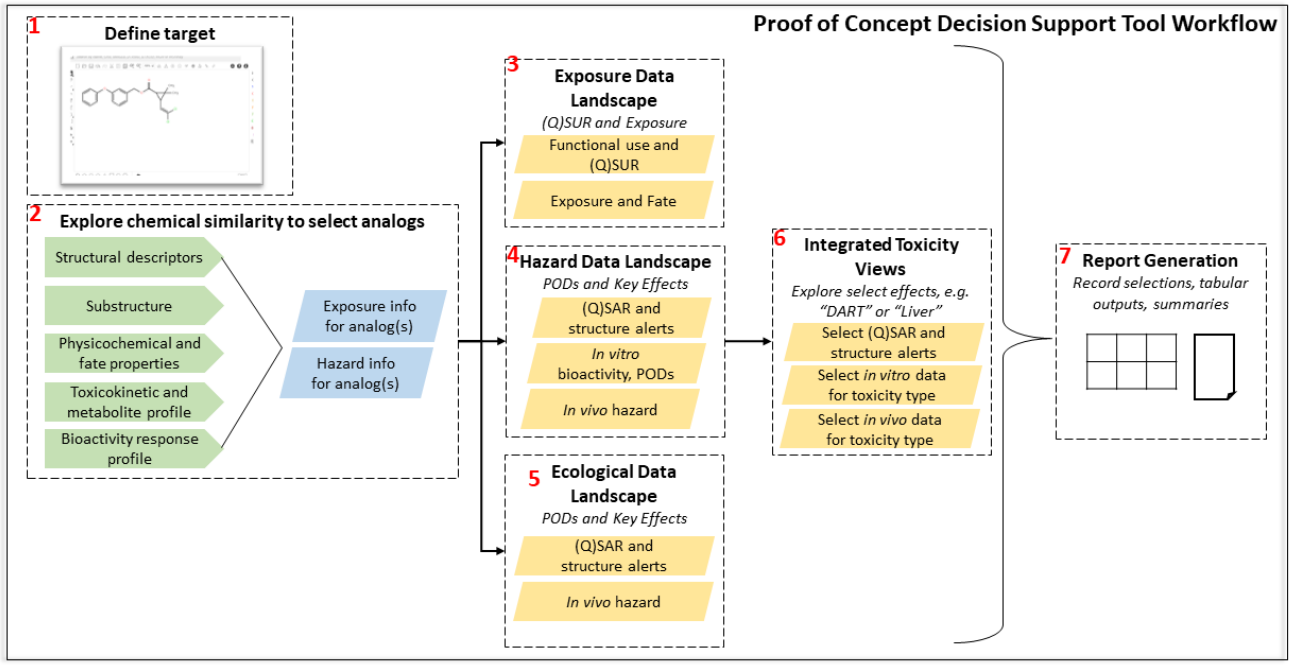
is supported by migrating as much of these data as possible via OHT formats to IUCLID. In work complementary to the NCCRP, OPPT is working to receive new chemical data submissions in IUCLID format or convert submitted study information to OHT-compliant formats and relay to IUCLID. Ongoing efforts within OPPT aim to reformat and migrate the data submitted under TSCA from current databases to an instance of IUCLID in OPPT's protected CBI environment. ORD is supporting this effort via pilot work to digitize some OPPT data, including publicly available TSCA Section 8(e) reports, as well as new chemical assessment reports and other summary hazard documents within OPPT's protected CBI environment. Overall, this collaboration to support digitization, integration, and ultimate conversion to IUCLID-compatible formats will support collation of these data for utilization by applications, such as a decision support tool.

C. Developing proof-of-concept decision support tool for new chemicals

Next generation risk assessment workflows that bring together chemistry, hazard, toxicokinetics, and exposure demonstrate early efforts in the open literature to create reproducible analyses and consolidate disparate data for regulatory toxicology (Baltazar et al., 2020; Beal et al., 2022; Dent et al., 2021; Ouedraogo et al., 2022; Paul Friedman et al., 2020; Rajagopal et al., 2022; van Tongeren et al., 2021). In ORD, CCTE is building an ecosystem of decision support tools to meet stakeholder needs and advance toward next generation risk assessments that utilize as much computationally accessible information as possible. To enable OPPT to rapidly review relevant information from both traditional and NAM sources and make reproducible and documented decisions using many types of information, a proof-of-concept decision support tool will be developed as part of this ecosystem. This decision support tool will integrate data domains such as chemistry, hazard (including data from IUCLID), bioactivity (including IVIVE of dose), environmental fate, functional use, and exposure. Development of this tool requires expertise and teamwork from regulatory experts in OPPT NCD, ORD technical experts in the data domains, and experts in software development, data engineering, systems administration, among other information technology domains. The intention within ORD is to develop this tool in a modular, rapid, and innovative way with feedback from OPPT. Continual assessment of how to best mature the application within CCTE's ecosystem of decision support tools and data architecture will also be needed. A draft overview of a possible proof-of-concept decision support tool is illustrated in Figure 4.

Figure 4. Draft overview of proof of concept for NCCRP decision support tool.

Expert users would (1) define a target structure; (2) explore chemical similarity among possible analogs using structural descriptors, substructure, physicochemical and fate properties, metabolite profile, and/or bioactivity response profile and select analogs based on available exposure and/or hazard information for these similar analogs; (3) explore the exposure data landscape for selected analog(s), including functional use and exposure/fate data and predictions; (4) explore the hazard data landscape for selected analog(s), including (Q)SAR and structure alerts, *in vitro* bioactivity, and *in vivo* hazard data, with the aim of identifying key toxicity types and/or a POD; (5) explore the hazard landscape for ecological health, including (Q)SARs and *in vivo* data relevant to a POD; (6) explore integrated toxicity views for select effects such as developmental and reproductive toxicity (DART) or specific organ toxicity to build a weight of evidence or fill data gaps for these toxicity types; and (7) generate reports that include selected data used, narrative justifications for decisions, and automated summaries of the data.



In support of these goals for the NCCRP, ORD plans to extend ongoing work on the Cheminformatics Analysis Modules, a proof-of-concept decision support tool that currently integrates data streams including curated *in vivo* hazard data, structural alerts, predicted and experimental physicochemical and environmental fate and transport properties, as well as (Q)SAR-predicted toxicity endpoints, and ToxPrint chemotypes. Important needs in the NCCRP are: (1) to connect experts in NCD with analog-related data as chemical submissions are mostly data poor; (2) organize/output data for easy interpretation and incorporation into new chemical evaluations; (3) add functionality to connect structures and/or their analogs to functional uses, and environmental fate and exposure data; (4) add functionality to evaluate bioactivity and integrate targeted bioactivity (from Research Area 4) with

structural alerts, (Q)SARs or predictive models for hazard, and/or any available hazard data for specific toxicity types (*e.g.*, developmental or reproductive toxicity and carcinogenicity), and (5) add functionality to connect analog information to any available (Q)SAR, structure alerts, or relevant *in vivo* hazard information relevant to ecological health. Each proof-of-concept module will need to export and/or capture selection of data for an assessment.

A draft overview of a possible proof-of-concept decision support tool, based on extension of the Cheminformatics Analysis Modules, is illustrated in Figure 4. Additional features for this proof-of-concept decision support tool could include a structure searching and chemical profiling module, which would provide various alerting rules (*e.g.*, membership in lists such as IARC collections, Ashby carcinogenicity alerts, Threshold of Toxicological Concern (TTC) alerts, and other user-definable approaches). Additionally, physicochemical and metabolite predictions may be generated. Read-across based on similarity metrics informed by structure, properties, and/or available data may be included as a module. Another module may include data and/or predictions for exposure, fate, and functional use; a set of modules could review the hazard data landscape for human and ecological health, including (Q)SAR predictions and category or analog approaches to identifying relevant *in vivo* and *in vitro* hazard data; and another set of modules could include integrated toxicity views, or views of structure alerts, (Q)SARs, applicable *in vitro* data and *in vivo* data for specific toxicity types, such as developmental and reproductive toxicity. Importantly, this workflow would provide reporting capabilities within each module to support OPPT in developing their assessments. Users will be able to save selections of information considered important for the new chemical assessment along with any narrative justifications. Developing this tool will allow ORD and OPPT to gain experience working together as they iteratively refine design requirements. As a long-term goal, ORD and OPPT will be creating a software tool with key functionality that can be populated with public and/or CBI information in a CBI-protected environment to improve the overall workflow and decision-making process for NCD chemical assessments.

Conclusion

The proposed research relevant to the NCCRP in the FY23-26 StRAP is extensive, connecting with many of the goals in ORD to support next generation risk assessment through the development and implementation of NAMs and decision support tools. Collaboration with OPPT will ensure that this research leads to fit-for-purpose translation and implementation, and that the needs of regulatory decision-making influence the research in ORD. Though the ultimate success of the proposed research in

ORD is resource dependent, the vision laid out in this research program will not only create proof-of-concept next generation risk assessment tools for OPPT, but will also demonstrate progress toward accomplishing key goals in the EPA NAM Work Plan and the CompTox BluePrint.

References

Baltazar, M. T., et al., 2020. A Next-Generation Risk Assessment Case Study for Coumarin in Cosmetic Products. Toxicol Sci. 176**,** 236-252.

Beal, M. A., et al., 2022. Implementing in vitro bioactivity data to modernize priority setting of chemical inventories. ALTEX. 39**,** 123-139.

Bowes, J., et al., 2012. Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. Nat Rev Drug Discov. 11**,** 909-22.

Dawson, D. E., et al., 2021. Designing QSARs for Parameters of High-Throughput Toxicokinetic Models Using Open-Source Descriptors. Environ Sci Technol. 55**,** 6505-6517.

Dent, M. P., et al., 2021. Paving the way for application of next generation risk assessment to safety decision-making for cosmetic ingredients. Regul Toxicol Pharmacol. 125**,** 105026.

Dionisio, K. L., et al., 2018. The Chemical and Products Database, a resource for exposure-relevant data on chemicals in consumer products. Sci Data. 5**,** 180125.

Djoumbou Feunang, Y., et al., 2016. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. J Cheminform. 8**,** 61.

Gamble, J. T., et al., 2022. The DevTox Germ Layer Reporter Platform: An Assay Adaptation of the Human Pluripotent Stem Cell Test. Toxics. 10.

Grulke, C. M., et al., 2019. EPA's DSSTox database: History of development of a curated chemistry resource supporting computational toxicology research. Comput Toxicol. 12.

Haggard, D. E., et al., 2018. High-Throughput H295R Steroidogenesis Assay: Utility as an Alternative and a Statistical Approach to Characterize Effects on Steroidogenesis. Toxicol Sci. 162**,** 509-534.

Haggard, D. E., et al., 2019. Development of a prioritization method for chemical-mediated effects on steroidogenesis using an integrated statistical analysis of high-throughput H295R data. Regul Toxicol Pharmacol. 109**,** 104510.

Harrill, J. A., et al., 2021. High-Throughput Transcriptomics Platform for Screening Environmental Chemicals. Toxicol Sci. 181**,** 68-89.

Helman, G., et al., 2019a. Quantitative prediction of repeat dose toxicity values using GenRA. Regul Toxicol Pharmacol. 109**,** 104480.

Helman, G., et al., 2018. Extending the Generalised Read-Across approach (GenRA): A systematic analysis of the impact of physicochemical property information on read-across performance. Comput Toxicol. 8**,** 34-50.

Helman, G., et al., 2019b. Generalized Read-Across (GenRA): A workflow implemented into the EPA CompTox Chemicals Dashboard. ALTEX. 36**,** 462-465.

Honda, G. S., et al., 2019. Using the concordance of in vitro and in vivo data to evaluate extrapolation assumptions. PLoS One. 14**,** e0217564.

Ietswaart, R., et al., 2020. Machine learning guided association of adverse drug reactions with in vitro target-based pharmacology. EBioMedicine. 57**,** 102837.

Isaacs, K. K., et al., 2022. A harmonized chemical monitoring database for support of exposure assessments. Sci Data. 9**,** 314.

Judson, R., et al., 2020. Selecting a minimal set of androgen receptor assays for screening chemicals. Regul Toxicol Pharmacol. 117**,** 104764.

Judson, R. S., et al., 2017. On selecting a minimal set of in vitro assays to reliably determine estrogen agonist activity. Regul Toxicol Pharmacol. 91**,** 39-49.

Judson, R. S., et al., 2015. Integrated Model of Chemical Perturbations of a Biological Pathway Using 18 In Vitro High-Throughput Screening Assays for the Estrogen Receptor. Toxicol Sci. 148**,** 137-54.

Kameoka, S., et al., 2014. A high-throughput screen for teratogens using human pluripotent stem cells. Toxicol Sci. 137**,** 76-90.

Klaren, W. D., et al., 2019. Identifying Attributes That Influence In Vitro-to-In Vivo Concordance by Comparing In Vitro Tox21 Bioactivity Versus In Vivo DrugMatrix Transcriptomic Responses Across 130 Chemicals. Toxicol Sci. 167**,** 157-171.

Kleinstreuer, N. C., et al., 2017. Development and Validation of a Computational Model for Androgen Receptor Activity. Chem Res Toxicol. 30**,** 946-964.

Mansouri, K., et al., 2018. OPERA models for predicting physicochemical properties and environmental fate endpoints. J Cheminform. 10**,** 10.

Mansouri, K., et al., 2016. An automated curation procedure for addressing chemical errors and inconsistencies in public datasets used in QSAR modelling. SAR QSAR Environ Res. 27**,** 939-965.

Nyffeler, J., et al., 2021. Comparison of Approaches for Determining Bioactivity Hits from High-Dimensional Profiling Data. SLAS Discov. 26**,** 292-308.

Nyffeler, J., et al., 2022. Combining phenotypic profiling and targeted RNA-Seq reveals linkages between transcriptional perturbations and chemical effects on cell morphology: Retinoic acid as an example. Toxicol Appl Pharmacol. 444**,** 116032.

Nyffeler, J., et al., 2020. Bioactivity screening of environmental chemicals using imaging-based high-throughput phenotypic profiling. Toxicol Appl Pharmacol. 389**,** 114876.

OECD, Guidance Document on the Validation of (Quantitative)Structure-Activity Relationships (QSAR) Models. ENV/JM/MONO(2007)2. Environment Directorate, Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides, and Biotechnology, https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=env/jm/mono(2007)2, 2007.

OECD, Guidance on the Grouping of Chemicals, Second Edition. Vol. No. 194. Environment Directorate, Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides, and Biotechnology, ENV/JM/MONO(2014)4, 2014.

OECD, Customisation Opportunities of IUCLID for the Management of Chemical Data - 2nd Edition, OECD Series on Testing and Assessment. https://doi.org/10.1787/20777876. OECD Publishing, Paris, 2021.

Olker, J. H., et al., 2022. The ECOTOXicology Knowledgebase: A Curated Database of Ecologically Relevant Toxicity Tests to Support Environmental Research and Risk Assessment. Environ Toxicol Chem. 41**,** 1520-1539.

Ouedraogo, G., et al., 2022. Read-across and new approach methodologies applied in a 10-step framework for cosmetics safety assessment - A case study with parabens. Regul Toxicol Pharmacol. 132**,** 105161.

Papoian, T., et al., 2017. Regulatory Forum Review*: Utility of in Vitro Secondary Pharmacology Data to Assess Risk of Drug-induced Valvular Heart Disease in Humans: Regulatory Considerations. Toxicol Pathol. 45**,** 381-388.

Paul Friedman, K., et al., 2020. Utility of In Vitro Bioactivity as a Lower Bound Estimate of In Vivo Adverse Effect Levels and in Risk-Based Prioritization. Toxicol Sci. 173**,** 202-225.

Phillips, K. A., et al., 2017. High-throughput screening of chemicals as functional substitutes using structure-based classification models. Green Chem. 19**,** 1063-1074.

Pradeep, P., et al., 2020. Using Chemical Structure Information to Develop Predictive Models for In Vitro Toxicokinetic Parameters to Inform High-throughput Risk-assessment. Comput Toxicol. 16.

Rajagopal, R., et al., 2022. Beyond AOPs: A Mechanistic Evaluation of NAMs in DART Testing. Front Toxicol. 4**,** 838466.

Ring, C., et al., 2021. Predictive modeling of biological responses in the rat liver using in vitro Tox21 bioactivity: Benefits from high-throughput toxicokinetics. Comput Toxicol. 18.

Ring, C. L., et al., 2019. Consensus Modeling of Median Chemical Intake for the U.S. Population Based on Predictions of Exposure Pathways. Environ Sci Technol. 53**,** 719-732.

Rogers, D., Hahn, M., 2010. Extended-connectivity fingerprints. J Chem Inf Model. 50**,** 742-54.

Sayre, R. R., et al., 2020. Database of pharmacokinetic time-series data and parameters for 144 environmental chemicals. Sci Data. 7**,** 122.

Shah, I., et al., 2016. Systematically evaluating read-across prediction and performance using a local validity approach characterized by chemical structure and bioactivity information. Regul Toxicol Pharmacol. 79**,** 12-24.

Shah, I., et al., 2021. Generalised Read-Across prediction using genra-py. Bioinformatics.

Sipes, N. S., et al., 2013. Profiling 976 ToxCast chemicals across 331 enzymatic and receptor signaling assays. Chem Res Toxicol. 26**,** 878-95.

Sipes, N. S., et al., 2017. An Intuitive Approach for Predicting Potential Human Health Risk with the Tox21 10k Library. Environ Sci Technol. 51**,** 10786-10796.

Smit, I. A., et al., 2021. Systematic Analysis of Protein Targets Associated with Adverse Events of Drugs from Clinical Trials and Postmarketing Reports. Chem Res Toxicol. 34**,** 365-384.

Speen, A. M., et al., 2022. Benchmark Dose Modeling Approaches for Volatile Organic Chemicals Using a Novel Air-Liquid Interface In Vitro Exposure System. Toxicol Sci. 188**,** 88-107.

Thomas, R. S., et al., 2019. The Next Generation Blueprint of Computational Toxicology at the U.S. Environmental Protection Agency. Toxicol Sci. 169**,** 317-332.

USEPA, 2020a. Chemical and Products Database (CpDat) MySQL Data. https://doi.org/10.23645/epacomptox.5352997.

USEPA, User's Guide for T.E.S.T version 5.1: Toxicity Estimation Software Tool: A Program to Estimate Toxicity from Molecular Structure. In: Martin, T. M., (Ed.). Office of Research and Development, Center for Computational Toxicology and Exposure, Cincinnati, OH, 2020b.

USEPA, National PFAS Testing Strategy: Identification of Candidate Per- and Poly-fluoroalkyl Substances (PFAS) for Testing. Office of Pollution Prevention and Toxics (OPPT) and Office of Research and Development (ORD) https://www.epa.gov/system/files/documents/2021-10/pfas-natl-test-strategy.pdf, Washington, DC 20460, 2021a.

USEPA, New Approach Methods Work Plan: Reducing use of animals in chemical testing. Office of Research and Development and Office of Chemical Safety and Pollution Prevention, https://www.epa.gov/system/files/documents/2021-11/nams-work-plan_11_15_21_508-tagged.pdf, 2021b.

USEPA, A Proof-of-Concept Case Study Integrating Publicly Available Information to Screen Candidates for Chemical Prioritization under TSCA. https://doi.org/10.23645/epacomptox.14878125, 2021c.

USEPA, Modernizing the Process and Bringing Innovative Science to Evaluate New Chemicals Under TSCA. Office of Pollution Prevention and Toxics; Office of Research and Development, https://www.regulations.gov/docket/EPA-HQ-OPPT-2022-0218/document, 2022a.

USEPA, Multimedia Monitoring Database (MMDB). Office of Research and Development, Center for Computational Toxicology and Exposure. https://doi.org/10.23645/epacomptox.17065024.v1, 2022b.

Valentin, J. P., et al., 2018. In vitro secondary pharmacological profiling: An IQ-DruSafe industry survey on current practices. J Pharmacol Toxicol Methods. 93**,** 7-14.

van Tongeren, T. C. A., et al., 2021. Next generation risk assessment of human exposure to anti-androgens using newly defined comparator compound values. Toxicol In Vitro. 73**,** 105132.

Wambaugh, J. F., et al., 2019. New approach methodologies for exposure science. Current Opinion in Toxicology. 15**,** 76-92.

Watford, S., et al., 2019. ToxRefDB version 2.0: Improved utility for predictive and retrospective toxicology analyses. Reprod Toxicol. 89**,** 145-158.

Williams, A. J., et al., 2017. The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. J Cheminform. 9**,** 61.

Willis, C., et al., 2020. Phenotypic Profiling of Reference Chemicals across Biologically Diverse Cell Types Using the Cell Painting Assay. SLAS Discov. 25**,** 755-769.

Yang, C., et al., 2015. New publicly available chemical query language, CSRML, to support chemotype representations for application to data mining and modeling. J Chem Inf Model. 55**,** 510-28.

Zavala, J., et al., 2017. Regulating temperature and relative humidity in air-liquid interface in vitro systems eliminates cytotoxicity resulting from control air exposures. Toxicol Res (Camb). 6**,** 448-459.

Zavala, J., et al., 2018. A new cell culture exposure system for studying the toxicity of volatile chemicals at the air-liquid interface. Inhal Toxicol. 30**,** 169-177.