DISCUSSION DOCUMENT

Possible Approaches and Tools for
Identifying Potential Candidate Chemicals for Prioritization

*This document was prepared for discussion purposes only, in advance of the public meeting on December 11, 2017 regarding approaches for identifying potential candidate chemical substances for prioritization.

Table of Contents

Introduction……3

Key Terms……..4

Background…….……..…….5

Goals and Guiding Principles ……9

Overview of EPA's TSCA Work Plan Chemicals Methods Document……...…12

Possible Approaches and Tools for Identifying Potential Candidates ……....15

1. The TSCA Work Plan as a Tool ……...15

2. Canada's Chemicals Management Plan…….26

3. Utilizing Safer Chemicals Ingredients List ……...33

4. Functional Category Approach, based on Use and Exposure Potential ……38

5. Functional Category Approach, based on Chemical Structure and Function ……...49

6. Integration of Traditional and New Approaches……..……....53

Introduction

This discussion document introduces a set of approaches that the Agency is considering to help guide the identification of potential candidates for prioritization, and is intended to be a starting point for a dialogue with stakeholders on best practices for EPA's activities during this phase. EPA plans to describe these possible approaches at the public meeting on December 11, 2017, and solicit feedback from interested stakeholders. This document, the associated public meeting, and the public comment period are the initial steps of a dialogue EPA expects will last approximately 6 months. EPA is asking for input on the approaches presented here, as well as any additional recommendations. EPA is not committing to any particular final "product," but rather will consider any input received and move forward as appropriate. EPA is open to considering these approaches, and any other approaches identified through the stakeholder process, as ones that could be used either singly or together as combinations of approaches to identify potential candidates for prioritization.

Key Terms

TSCA – the Toxic Substances Control Act, as amended in 2016 by the "Frank R. Lautenberg Chemical Safety for the 21st Century Act"
NAM – New Approach Methods
HTTK – High-Throughput Toxicokinetics
ER – estrogen receptor
AR – androgen receptor
PBT – Persistent, bioaccumulative, and toxic
QSAR – quantitative structure activity relationship
QBAR – quantitative bioactivity relationship
SHEDS-HT – High-Throughput Stochastic Human Exposure and Dose Simulation Model
SEEM – Systematic Empirical Evaluation of Models
BAF – bioaccumulation factor
H/BER - hazard/bioactivity exposure ratio
NOAEL – no observed adverse effect level
LOAEL – lowest observed adverse effect level
HTS – High-Throughput Screening

Background

Under the Toxic Substances Control Act (TSCA), as amended by the Frank R. Lautenberg Chemical Safety for the 21st Century Act, EPA has a mandate to prioritize and evaluate the risks of existing chemical substances[1]. The law imposes deadlines and minimum requirements for numbers of chemicals, and provides the general process and criteria by which prioritization and risk evaluation must be conducted (Figure 1).

Prioritization is a 9- to 12-month public process during which a chemical substance or category of chemicals is designated as high-priority for risk evaluation or designated as low-priority. A chemical designated as low-priority means an evaluation is not warranted at that time. A chemical designated as high-priority must move immediately into the Risk Evaluation phase. TSCA requires that High-Priority chemicals be evaluated for hazard and exposure to both humans, including potentially exposed or susceptible subpopulations[2], and the environment, under the conditions of use, culminating in determinations of risk – either no unreasonable risk, or unreasonable risk necessitating risk management to eliminate those identified risks. The risk evaluation must take no longer than 3 to 3.5 years to complete. If unreasonable risks are identified, EPA has 2 to 4 years to address those risks by regulation.

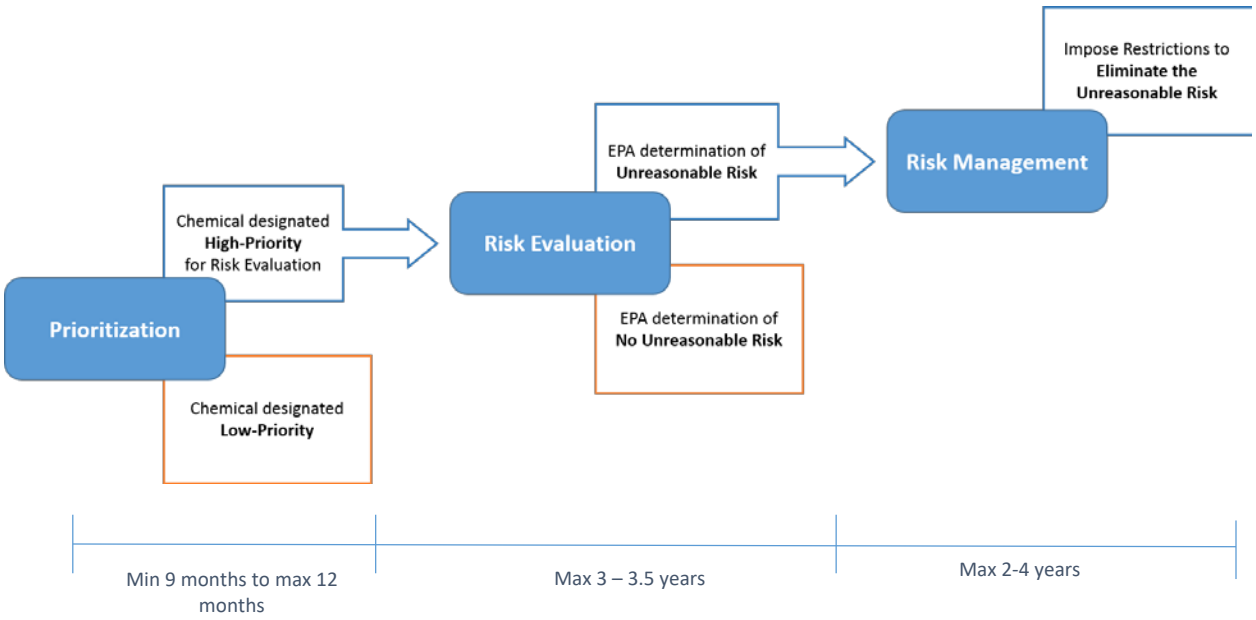


Figure 1. Process and timelines for prioritizing and evaluating the risks of existing chemicals

[1] Unless otherwise indicated, any references to "chemical" or "chemical substance" throughout this document means a "chemical substance" as defined in TSCA Section 3(2).

[2] "Potentially exposed or susceptible subpopulation," as defined in TSCA Section 3(12), means a group of individuals within the general population identified by the Administrator who, due to either greater susceptibility or greater exposure, may be at greater risk than the general population of adverse health effects from exposure to a chemical substance or mixture, such as infants, children, pregnant women, workers or the elderly.

The new requirements in TSCA ensure that EPA will make continued progress in reviewing existing chemicals, determining which of those chemicals merit further evaluation, and managing identified unreasonable risks. In accordance with those requirements in TSCA, EPA finalized rules for the Prioritization and Risk Evaluation processes[3]. The statute provides EPA with discretion to choose which chemical substances to put into the Prioritization process.

EPA included in the proposed Prioritization rule a 'Pre-Prioritization' process, which provided considerations for identifying potential high- and low-priority candidates, and general hazard and exposure considerations. Commenters expressed a strong desire to better understand these criteria and how they might be applied, and to increase public participation and opportunities for comment during the pre-prioritization phase. As such, in the final rule, EPA deferred final action on the pre-prioritization provisions. EPA also pledged to initiate a stakeholder process, to include an additional public comment opportunity to discuss how EPA could identify potential candidates for prioritization, and to carry out these activities in a transparent manner to help ensure successful implementation of EPA's TSCA existing chemicals program. EPA remains committed to gathering and incorporating additional feedback from interested members of the public on how to identify potential candidates for prioritization. This discussion document is intended to provide a set of possible approaches under consideration by the Agency to start a stakeholder dialogue at the public meeting. In addition to those comments received on the proposed Prioritization rule, EPA will consider feedback received at the December 11 public meeting and during the associated opportunity for comment as part of the ongoing dialogue. This discussion document is intended to provide a set of possible approaches under consideration by the Agency and to inform stakeholder dialogue at the public meeting.

Overarching Goal and Guiding Principles

EPA has identified an overarching goal, milestones, and guiding principles intended to shape initial thinking for identifying potential candidates for prioritization. TSCA requires that EPA prioritize at least 20 high- and at least 20 low-priority chemicals within 3.5 years of the law's enactment, or by approximately the end of December 2019. With stakeholder input and collaboration with other state and federal agencies, the overarching goal is to develop an

---

[3] Final Rule, "Procedures for Prioritization of Chemicals for Risk Evaluation Under the Toxic Substances Control Act," available at https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/federal-register-notice-procedures-prioritization
Final Rule, "Procedures for Chemical Risk Evaluation Under the Amended Toxic Substances Control Act," available at: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/procedures-chemical-risk-evaluation-under-amended-toxic

approach, or set of approaches, that will enable EPA to identify a sufficient number of potential candidates, initiate the prioritization process for those candidates, and finalize priority designations within the statutory deadline. The goal and guiding principles are described further in the following section.

Timeframe for Prioritization Activities

The statute requires EPA to have *completed* the prioritization process for 40 chemicals, with 20 designated as low-priority and 20 designated as high priority[4], by approximately the end of December 2019 (i.e., within 3.5 years of the law's enactment in June 2016). As such, EPA must identify at least 40 potential candidates and *begin* the 9 to 12-month prioritization process by no later than the end of March 2018. Additionally, with each risk evaluation completed, EPA must be prepared to initiate risk evaluation on another high-priority chemical as required under TSCA. EPA must therefore be prepared to begin the prioritization process on additional potential high-candidates 9 to 12 months before the expected completion date of a high-priority chemical risk evaluation. Finally, although TSCA does not require designation of more than 20 low-priority chemicals, EPA is committed to identifying more than this statutory minimum through the prioritization process.

Identifying Information Gaps

Prior to designating a chemical as a high-priority for risk evaluation, it is important for EPA to ensure the reasonably available information is sufficient to conduct a scientifically robust risk evaluation. In many cases, EPA believes it would be difficult to require the development of necessary chemical substance information, evaluate that information, and incorporate that information into analyses and decisions within the statutory timeframes associated with the prioritization and risk evaluation processes. Therefore, it will be useful for EPA to identify information needs and determine whether any of these needs should be addressed before initiating the prioritization process.

Information needs may be filled through a variety of methods. For example, high-throughput screening and computational modeling may fill important gaps in traditional animal toxicology data, as well as focus future data requests. Additionally, voluntary submissions, information sharing with other state and federal agencies, and utilization of authorities under TSCA sections 4, 8, and 11(c) for the development of necessary information for prioritization and risk evaluation will be important information gathering methods which EPA will utilize as appropriate.

[4] Note: the requirement to designate 20 high-priority chemicals by approximately the end of 2019 is in addition to the first 10 chemicals that EPA identified outside of the prioritization process and for which risk evaluations are already underway.

Additionally, the Interagency Testing Committee (ITC) – an independent advisory committee formed under TSCA - can make recommendations to the EPA Administrator on prioritizing and selecting chemicals for testing or information reporting to meet the coordinated data needs of its member U.S. Government organizations.

Possible Approaches and Tools

As an initial step and to assist in ongoing discussions, the Agency has developed a number of possible approaches and tools that could aid in this preliminary review and identification of possible candidates for prioritization. EPA does not expect that the path forward will necessarily entail choosing one single approach, but rather may include a number of differing approaches and tools, or components of differing approaches and tools, that could work in tandem. As approaches are evaluated, selected, implemented, and process adjustments made, the Agency is committed to being transparent and communicating with the public how it evaluates existing chemicals for Prioritization.

It is important to note that TSCA requires at least 50 percent of all chemical substances undergoing risk evaluation come from the EPA's 2014 update of the TSCA Work Plan for Chemical Assessments. Merely being on the 2014 Work Plan does not constitute a finding of risk, as the chemicals on this list will enter the Prioritization process as any other chemical, for determination of high- or low-priority for risk evaluation. However, to meet this statutory requirement, at least 10 of the first 20 high-priority candidates must be drawn from the 2014 TSCA Work Plan. EPA must continue to draw at least 50 percent of it's high-priority substances from the 2014 Work Plan until the Work Plan is exhausted. In addition to the approaches described below, EPA welcomes input on how to bring chemicals on the 2014 Work Plan into the prioritization process.

EPA has described six possible approaches and tools for identifying candidates for prioritization in subsequent sections of this document. EPA recognizes that there could be pros and cons to each approach/tool, and welcomes feedback on these, other possible approaches/tools, and opportunities for combining approaches.

1) Incorporating TSCA Work Plan Methodology – This approach would utilize lessons learned in EPA's previous efforts in developing chemical work plans, and apply an updated methodology to move beyond those chemicals listed on the 2014 Update to the TSCA Work Plan.
2) An approach modeled from Canada's Chemical Management Plan (CMP) – This approach would use lessons learned and tools developed under the CMP to identify chemicals that may be appropriate for high- and low-priority.
3) Safer Chemicals Ingredients List (SCIL) –A description of SCIL and how it may serve as a good starting point for identifying potential candidates for low-priority designations.

4) Functional Category Approach, based on Use and Exposure Potential – This approach would identify a group of chemicals with similar functional use in industrial applications or in commercial or consumer products.
5) Functional Category Approach, based on Chemical Structure and Function – This approach would group chemicals based on a chemical's structure and physicochemical properties to achieve a particular function at the chemical level.
6) Integration of Traditional and New Approach Methods (NAM) – This approach uses a software tool and databases containing information from traditional and NAM to efficiently and cost-effectively focus EPA's efforts to identify those chemicals that may or may not present hazard or exposure potential.

Overarching Goal, Milestones, Statutory Elements and Guiding Principles

As described in this section, EPA has identified an overarching goal, milestones, key statutory elements and guiding principles intended to shape initial thinking for identifying potential candidates for prioritization.

Goal
The overarching goal is to develop an approach, or set of approaches, with stakeholder input, that will enable EPA to identify a sufficient number of potential candidates for prioritization, initiate the prioritization process for those candidates, and finalize those priority designations within the statutory deadline.

TSCA requires that EPA prioritize at least 20 high- and at least 20 low-priority chemicals within 3.5 years of the law's enactment, or by approximately the end of December 2019. To meet this this statutory requirement, EPA intends to engage with stakeholders over the next six months, develop an approach or set of approaches to identify at least 40 potential candidates for prioritization and *begin* the prioritization process for these candidates by no later than the end of March 2019.

Milestones
November 2017 – Initiate stakeholder engagement; release meeting materials and discussion document for public comment
December 11, 2017 – Public Meeting
January 25, 2018 - Comment period closes; EPA to begin reviewing comments received
June 2018 – Conclude stakeholder engagement; identify approaches that will be used
June 2018 – March 2019 – Implement approach or set of approaches/tools

Key Statutory Elements

1. Of the chemicals designated as high-priority, at least 50% must come from the 2014 Update to the TSCA Work Plan, until that list has been exhausted. EPA must therefore be mindful to identify a sufficient number of its potential candidates for high-priority from the TSCA Work Plan

2. Risk-based criteria must be used for Prioritization.

3. Designation of a chemical as a high priority for risk evaluation begins the three-year statutory deadline for completing the risk evaluation.

4. For each risk evaluation completed on a high-priority chemical, EPA must designate another high-priority chemical and initiate risk evaluation.

Guiding Principles:

1. EPA's approach to identifying potential candidates for prioritization should be risk-based and supported by science, just as the processes for prioritization and risk evaluations.

2. EPA's approach to identifying potential candidates should be guided by input from stakeholders, including state and federal agencies.

3. EPA should factor in the need for analyses of candidate's readiness for both prioritization and risk evaluation in order to ensure responsible implementation of TSCA. EPA should identify data needs and actively address those needs before initiating prioritization. This could include voluntary collection of information, sharing information from state and federal partners, and/or utilizing the authorities provided in TSCA sections 4, 8, and 11(c). Once EPA has initiated the prioritization process for a chemical, EPA must issue a final designation as either a high- or low-priority within 12 months. Chemicals designated as high-priority move immediately into risk evaluation with an associated 3-year statutory deadline for completion. In many cases, it could be difficult to require the development of necessary chemical substance information, and to evaluate, and incorporate that information into analyses and decisions within the statutory timeframes of both the prioritization and risk Evaluation processes. Likewise, the scientific underpinnings of a risk evaluation must be strong enough to inform potential future risk management activities.

4. EPA should be mindful of its workload and resource constraints, given the requirement in TSCA that EPA designate a high-priority chemical and initiate risk evaluation upon completion of an EPA-initiated risk evaluation. Incorrectly identified potential low-priority candidates that are subsequently designated as high-priority, for example, have the potential to permanently increase the number of ongoing risk evaluations.

5. EPA should strive to identify more than the statutory-mandated minimum of 20 low-priority chemicals. Knowledge of what chemicals are low priorities for risk evaluation will be valuable for EPA and all stakeholders at many points throughout EPA's TSCA existing chemicals program, and for many types of decisions by stakeholders and EPA.

6. EPA should consider whether high-throughput approaches offer a rapid and cost-effective approach to conducting an initial screen for hazard and exposure.

7. EPA should focus its efforts to identify potential candidates for prioritization from the active inventory, once it has been updated, given that active chemicals may have a greater potential for exposure.

8. EPA should balance transparency and stakeholder concerns over the development of lists of candidate chemicals. The stakeholder feedback received on the proposed prioritization rule indicated concern for stigmatizing large numbers of chemicals, if for example, EPA created and published potential candidate 'lists' without actually putting them into the prioritization/risk evaluation process for some length of time. EPA should avoid approaches that further premature judgments on risks, while also striving for transparency on how potential candidates for both high- and low-priority chemicals are identified.

Overview of EPA's TSCA Work Plan Chemicals Methods Document

EPA's Work Plan Methods document[5] describes the approach that EPA took in 2012 and 2014 to identify priority chemicals for review. This approach consisted of a two-step prioritization process. In the first step, EPA selected an initial group of candidate chemicals meeting one or more of the following factors:

- Chemicals identified as potentially of concern for children's health (e.g., chemicals with reproductive or developmental effects).
- Chemicals identified as persistent, bioaccumulative, and toxic (PBT).
- Chemicals identified as probable or known carcinogens.
- Chemicals of concern for neurotoxicity.
- Chemicals used in children's products.
- Chemicals used in consumer products.
- Chemicals detected in biomonitoring programs.

In the second step, EPA screened the selected chemicals for hazard, exposure and persistence and bioaccumulation and binned them as high, medium or low.
To generate the Step 1 chemicals meeting the Agency's prioritization factor criteria as potential candidates for review and assessment, the following sources were used to identify chemicals:

- ***Carcinogenicity:***
  - IRIS: 1986 Class A, B1; 1996 Known or Probable; 1999 or 2005 Carcinogenic
  - IARC Carcinogens, Group 1, 2A
  - NTP Known Carcinogens

- ***PBT:***
  - TRI PBT Rule
  - Great Lakes Binational PBT
  - Canadian P, B, and T (all three criteria met)
  - LRTAP POPS
  - Stockholm POPs

- ***Children's Health:***
  - IRIS: Repro/Dev (RfD or RfC for repro or dev)
  - NTP CERHR: Infants Any Effect or Pregnant Women Any Effect
  - Cal Prop 65 Reproductive

- ***Neurotoxicity:*** IRIS

- ***Children's Product Use****:*

[5] USEPA TSCA Work Plan Chemicals: Methods Document (February 2012) https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/2012-tsca-work-plan-chemicals

    - Reported in products intended for use by children in 2006 IUR
    - Washington State Children's List

- ***Biomonitoring*** (both human and environmental indicative of potential human exposure):
    - NHANES
    - Drinking Water Contaminants
    - Fish Tissue Studies

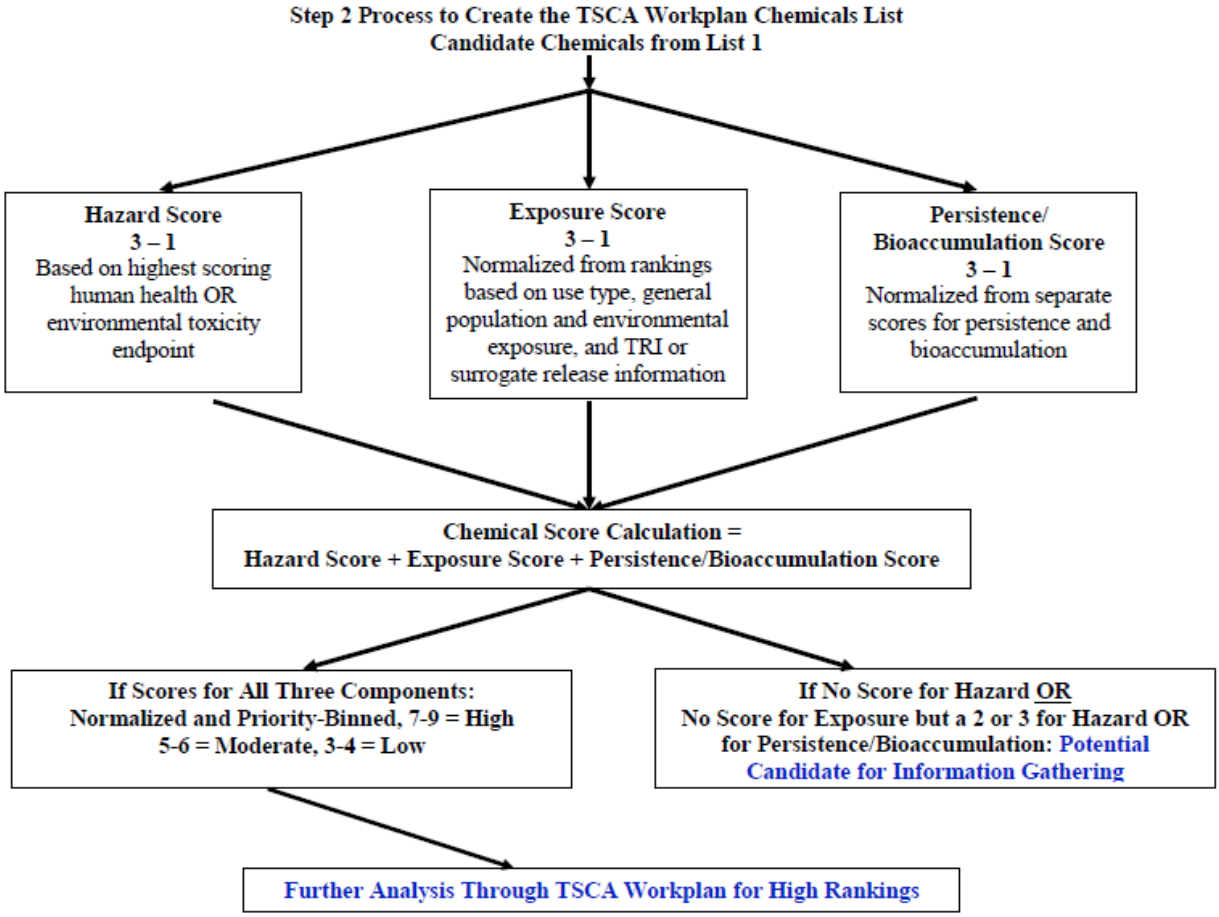
These sources produced a combined total of 1,235 chemicals, each of which matched at least one criterion. Pesticides, polymers, drugs, hormones, and pharmacological chemicals, certain radioactive materials, and certain other chemicals were excluded as potential candidates. (See the TSCA Work Plan Chemicals Methods Document at https://www.epa.gov/sites/production/files/2014-03/documents/work_plan_methods_document_web_final.pdf for the list of the types of substances excluded). 345 chemicals[6] remained as potential candidates and entered into Step 2.

The chemicals identified as potential candidates for review and assessment under TSCA based on the Step 1 prioritization factors were screened in Step 2 and received a score through the application of a numerical algorithm. This score was based on three characteristics: hazard, exposure, and potential for persistence and/or bioaccumulation. It should be noted that the hazard screening considered all toxicological endpoints, not only those considered in Step 1. Also note, the exposure ranking focused on exposure via consumer products and presence in biota and/or environmental media and considered Toxics Release Inventory data and TSCA Chemical Data Reporting data.

Using this system, chemicals were sorted into one of four bins – high, moderate, low and information gathering. Chemicals able to be scored on all three characteristics were scored as High, Moderate, or Low based on their available information. (See the figure below.) Chemicals with High or Moderate hazard or persistence/bioaccumulation scores that could not be scored for exposure because of an absence of data, together with chemicals that could not be scored for hazard, were identified separately as potential candidates for information gathering. 83 candidate chemicals from Step 1 that received scores on all three ranking factors and ranked High on the basis of their total score and were considered the 2012 Work Plan chemicals.

The TSCA Work Plan list was revised in 2014 using updated industry data submitted to EPA through the Toxics Release Inventory in 2011 and the TSCA Chemical Data Reporting requirements in 2012 on chemical releases and potential exposures. The final list of chemicals on the 2014 Work Plan list can be found at: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-chemical-assessments-2014-update.

[6] https://www.epa.gov/sites/production/files/2016-08/documents/2012_workplan_step_2_chemicals-for_web-final.pdf


Step 2 Process to Create the TSCA Workplan Chemicals List
Candidate Chemicals from List 1
Hazard Score
3 – 1
Based on highest scoring human health OR environmental toxicity endpoint
Exposure Score
3 – 1
Normalized from rankings based on use type, general population and environmental exposure, and TRI or surrogate release information
Persistence/ Bioaccumulation Score
3 – 1
Normalized from separate scores for persistence and bioaccumulation
Chemical Score Calculation =
Hazard Score + Exposure Score + Persistence/Bioaccumulation Score
If Scores for All Three Components:
Normalized and Priority-Binned, 7-9 = High
5-6 = Moderate, 3-4 = Low
If No Score for Hazard OR
No Score for Exposure but a 2 or 3 for Hazard OR
for Persistence/Bioaccumulation: Potential Candidate for Information Gathering
Further Analysis Through TSCA Workplan for High Rankings

Possible Approaches and Tools for Identifying Potential Candidates for Prioritization

1) The TSCA Work Plan as a Tool for Identifying Potential Candidates for Prioritization

Background

The 2014 Work Plan is a list of existing chemicals identified for near-term review and assessment under the Toxic Substances Control Act (TSCA). Agency use of the Work Plan is "to focus activities of the Existing Chemicals Program…so that existing chemicals having the highest potential for exposure and hazard are assessed…."[7]

EPA is required under the statute to prioritize the chemicals on the Work Plan as at least 50 percent of high-priority risk evaluations must come from the Work Plan. Therefore, the chemicals on the Work Plan must serve as initial candidates for prioritization. The Agency anticipates that identifying Work Plan chemicals for prioritization will entail considerations such as which chemicals first satisfy the preferences stated in the statute, including the consideration of potentially exposed and susceptible subpopulations, and determining if additional information is required prior to moving the chemical to Prioritization.

It is important to note that solely because a chemical is on the Work Plan does not constitute a finding of risk. These chemicals, just as any other chemical chosen, must be prioritized through the Prioritization process and if it is determined to be a high-priority chemical, will it be evaluated for risk.

For the remainder of this section, the 2014 Work Plan refers to the chemicals identified in the 2014 Update of the TSCA Work Plan[8] and the Work Plan Methodology refers to the screening methodology developed in the 2012 TSCA Work Plan Methods Document[9].

[7] https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-chemical-assessments-2014-update
[8] https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-chemical-assessments-2014-update
[9] TSCA Work Plan Methods Document: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-methods-document

What is the 2014 Work Plan? As described in the previous section, 2014 Work Plan is an update to the 2012 TSCA Work Plan for Chemical Assessment[10] with the removal and addition of substances due to the availability of new exposure data (reported under the Chemical Data Reporting Rule and to the Toxics Release Inventory). The Work Plan Methodology was developed to identify candidates in the 2012 TSCA Work Plan for Chemical Assessment and the 2014 Work Plan. In 2012, the Agency used several sources to identify chemicals meeting prioritization factor criteria (see Box 1) as potential candidates for review; a total of 1,235 chemicals were identified. These chemicals were screened to determine if any chemicals should be excluded because they are not subject to TSCA or there was already significant regulation under TSCA, or due to radioactivity, complex process streams, natural occurrence, or other properties. After screening, 345 chemicals[11] remained as potential candidates and entered the second stage of the Work Plan screening. In 2014, new data considerations were used to update the exposure ranking for the 345 existing chemicals that were generated under the two-step screening process of the Work Plan Methodology. The new data was submitted in 2012 under TSCA's Chemical Data Reporting or in 2011 as part of the EPA's Toxics Release Inventory reporting. Specifically, these data were used to update the exposure rankings for the chemicals initially screened as part of the Work Plan. These data were also used to screen ten Action Plan chemicals[12] and two additional chemicals identified by the Agency during EPA's assessment of flame retardants[13]. In total, 90 chemicals were identified in the 2014 Update of the TSCA Work Plan.

Box 1: Step 1 Factors considered in 2-step process:

- Chemicals identified as potentially of concern for children's health (e.g., chemicals with reproductive or developmental effects.
- Chemicals identified as persistent, bioaccumulative, and toxic (PBT).
- Chemicals identified as probable or known carcinogens.
- Chemicals used in children's products.
- Chemicals used in consumer products.
- Chemicals detected in biomonitoring programs.

Overview of the Work Plan Methodology

[10] 2012 TSCA Work Plan Chemicals: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/2012-tsca-work-plan-chemicals

[11] TSCA Work Plan: 2012 Scoring of Potential Candidate Chemicals Entering Step 2: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-2012-scoring-potential-candidate

[12] https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/current-chemical-risk-management-activities

[13] https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/fact-sheet-assessing-risks-flame-retardants

As described in the Work Plan Methodology document[14], EPA employed a two-step prioritization process that was intended to select an initial group of candidate chemicals for review. Based on comments received through the discussion forum, the webinar, and the stakeholder meeting[15], EPA made adjustments[16] both to the Step 1 factors and to the data sources utilized in both Step 1 and Step 2 to develop the 2014 workplan.

In Step 1, EPA used a specific set of data sources to identify chemicals meeting one or more of the Step 1 factors (see inset Box 1). This group of chemicals was further screened to determine if any chemicals should be excluded because they are not subject to TSCA or there was already significant regulation under TSCA, or due to radioactivity, complex process streams, natural occurrence, or other properties.

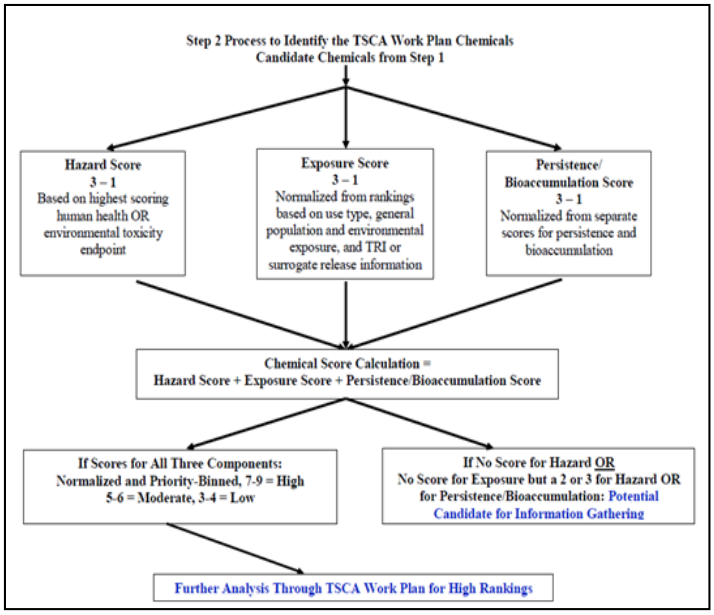
The chemicals identified as potential candidates for review and assessment under TSCA based on the Step 1 prioritization factors were screened in Step 2.[17]

[14] TSCA Work Plan Methods Document: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-methods-document

[15] In the Agency's August 2011 *Discussion Guide: Background and Discussion Questions for Identifying Priority Chemicals for Review and Assessment*, EPA described the two-step process the Agency intended to use to identify potential candidate chemicals for near-term review and assessment under the Toxic Substances Control Act (TSCA). The Agency intends to use these TSCA Work Plan Chemicals to help focus and direct the activities of the Existing Chemicals Program in the Office of Pollution Prevention and Toxics (OPPT). EPA invited public comment through an online discussion forum conducted from August 18 through September 21, 2011, as well as through a webinar and stakeholder meeting held on September 7, 2011. The meeting summaries and public comments are available for review in the docket for this activity, EPA-HQ-OPPT-2011-0516, which can be accessed online at http://www.regulations.gov.

[16] TSCA Work Plan Methods Document: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-methods-document

[17] List of chemicals can be found in these documents: 2014 Update of the TSCA Work Plan: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-chemicals#updates; 2012 TSCA Work Plan Chemicals: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/2012-tsca-work-plan-chemicals; TSCA Work Plan: 2012 Scoring of Potential Candidate Chemicals Entering Step 2: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-2012-scoring-potential-candidate



As shown in Box 2, chemicals were evaluated and received a score through the application of a numerical algorithm. This score was based on three characteristics: hazard, exposure, and potential for persistence and/or bioaccumulation.[18]

The Hazard Score encompasses both human health and environmental toxicity concerns. The specific hazard classification criteria are based on the Alternatives Assessment Criteria for Hazard Evaluation[19] developed by EPA's Design for the Environment Program (DfE)[20]. The DfE criteria for classifying the toxicity of specific chemicals were developed from authoritative sources including the United Nation's Globally Harmonized System (GHS) for Chemical Classification and Labeling[21] and other EPA programs.

The Exposure Score was based on a combination of chemical use, general population and environmental exposure, and release information. The Use Type score included consideration of consumer product applications as well as industrial and commercial uses that could result in widespread exposures. The General Population and Environmental Exposure score encompassed measured data on the presence of a chemical in biota and environmental media. The Release score was based on EPA's Toxics Release Inventory (TRI) data for chemicals subject to TRI reporting. For non-TRI chemicals, the Release score was calculated using a method involving Inventory Update Reporting data (IUR, now called Chemical Data Reporting, or CDR), including production volume, number of sites, and type of use.

[18] A quick overview of each score follows, but more details regarding the criteria for each score can be found in the TSCA Work Plan Methods Document: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-methods-document

[19] https://www.epa.gov/saferchoice/alternatives-assessment-criteria-hazard-evaluation

[20] https://www.epa.gov/saferchoice/design-environment-programs-initiatives-and-projects Design for the Environment was a US EPA program created in 1992 that developed several programs and tools to help its stakeholders evaluate human health and environmental attributes of chemicals in products. The program was renamed to Safer Choice in 2015 and maintains the Safer Choice Ingredient List (SCIL) which contains a list of chemical ingredients, arranged by functional-use class, that the Safer Choice Program has evaluated and determined to be safer than traditional chemical ingredients. The list is designed to help manufacturers find safer chemical alternatives that meet the criteria of the Safer Choice Program. See: https://www.epa.gov/saferchoice

[21] https://www.unece.org/trans/danger/publi/ghs/ghs_welcome_e.html

Persistence scoring consisted of the evaluation of the potential half-life in air, water, soil, and sediment while considering the expected partitioning characteristics of the chemicals and all potential removal pathways based on standard physical-chemical properties and environmental fate parameters. Bioaccumulation scoring consisted of evaluation of bioaccumulation/bioconcentration (measured or estimated BAF/BCF) data. When BAF data were not available, bioconcentration data (measured or estimated) were used to evaluate the potential for a chemical to bioaccumulate in organisms in the environment.

Using this system, chemicals were sorted into one of four bins (see Box 2). Chemicals able to be scored on all three characteristics were scored as High, Moderate, or Low based on their available information. Chemicals with High or Moderate hazard or persistence/bioaccumulation scores that could not be scored for exposure because of an absence of data, together with chemicals that could not be scored for hazard, were identified separately as "Potential Candidates for Information Gathering." These chemicals were identified as potential candidates for information-gathering activities focused on producing sufficient information to determine where they would rank in the prioritization process. Information-gathering activities that EPA would consider include both voluntary data submission and utilization of authorities under TSCA sections 4, 8, and 11(c). EPA created this separate category to ensure that chemicals with unknown toxicity or with known potential human health or environmental toxicity implications would not be removed from further investigation simply because there was a lack of exposure or hazard information, an issue stakeholders identified during the webinar and discussion forum as being of concern.[22]

The 2014 Work Plan Chemicals

The 2014 Work Plan Chemicals covers 90 chemicals that were identified in the 2014 Update of the TSCA Work Plan. The 2014 Work Plan Chemicals were identified after re-screening the 345 Step 2 chemicals[23] identified in 2012 with updated exposure rankings using new data (reported

[22] In the Agency's August 2011 *Discussion Guide: Background and Discussion Questions for Identifying Priority Chemicals for Review and Assessment*, EPA described the two-step process the Agency intended to use to identify potential candidate chemicals for near-term review and assessment under the Toxic Substances Control Act (TSCA). The Agency intends to use these TSCA Work Plan Chemicals to help focus and direct the activities of the Existing Chemicals Program in the Office of Pollution Prevention and Toxics (OPPT). EPA invited public comment through an online discussion forum conducted from August 18 through September 21, 2011, as well as through a webinar and stakeholder meeting held on September 7, 2011. The meeting summaries and public comments are available for review in the docket for this activity, EPA-HQ-OPPT-2011-0516, which can be accessed online at http://www.regulations.gov.

[23] TSCA Work Plan: 2012 Scoring of Potential Candidate Chemicals Entering Step 2: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-2012-scoring-potential-candidate

under the Chemical Data Reporting Rule and to the Toxics Release Inventory). These data were also used to screen ten Action Plan chemicals[24] and two additional chemicals identified by the Agency during EPA's assessment of flame retardants[25]. Chemicals added and removed in the 2014 Update are detailed in Table 1. Numbers of chemicals identified within each step of the Work Plan Methodology are given in Table 2 for both the 2012 and 2014 Work Plan lists.[26]

With the passage of the Lautenberg Act, EPA was required to select the first 10 chemicals to undergo risk evaluations from the 2014 Work Plan. These 10 chemicals were announced on December 16, 2016, reducing the TSCA Work Plan to 80 chemicals for consideration. Further, under the Toxic Substances Control Act (TSCA), as amended by the Frank R. Lautenberg Chemical Safety for the 21st Century Act, EPA has new authorities to regulate certain existing chemicals. Section 6(h) directs EPA to take action on certain persistent, bioaccumulative, and toxic (PBT) chemicals. Of the 80 remaining 2014 Work Plan Chemicals, five PBT chemicals were identified for action by EPA according to statutory criteria.[27] After these two actions, 75 chemicals from the 2014 Work Chemicals remain for potential candidate selection for TSCA prioritization.

Table 1. Chemicals Added and Removed in the 2014 Update

Chemicals added in the 2014 Update	Chemicals removed in the 2014 Update
5 of the 10 Action Plan chemicals - chemicals (or groups of chemicals) that scored 'high' under the 2012 methodology including Bisphenol A, group of phthalates, HBCD, etc.	13 chemicals removed because they are not currently in commerce based on data the Agency received under the CDR rule and as part

[24] https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/current-chemical-risk-management-activities

[25] https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/fact-sheet-assessing-risks-flame-retardants

[26] 2014 Update of the TSCA Work Plan: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-chemicals#updates; 2012 TSCA Work Plan Chemicals: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/2012-tsca-work-plan-chemicals

[27] https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/persistent-bioaccumulative-and-toxic-pbt-chemicals-under
Two additional PBT chemicals met the TSCA section 6(h) criteria; however, manufacturers for these substances submitted timely requests to EPA for risk evaluations pursuant to section 6(h)(5) and are therefore not subject to the rulemaking effort. As a result of the requests, two chemicals (Ethanone, 1-(1,2,3,4,5,6,7,8-octahydro-2,3,5,5-tetramethyl-2-naphthalenyl) and Ethanone, 1-(1,2,3,4,5,6,7,8-octahydro-2,3,8,8-tetramethyl-2-naphthalenyl)) are excluded from the expedited action requirements under TSCA section 6(h).

	of TRI reporting (e.g., no longer present exposure potential from current consumer or commercial use).
10 chemicals added from the 345 chemicals screened in 2012 due to submitted CDR and TRI data (in 2012 and 2011, respectively) indicating an increase in their exposure score and subsequent final score of 'high'	Mercury and mercury compounds removed because their hazards are already well characterized and the agency has an existing risk reduction effort in place (e.g., continued risk management measures, including efforts to implement the Minamata Convention)
3 flame retardants added (not among the chemicals screened in 2012) – TPP, isopropylated phenol and iPTPP	Quartz removed because it presents a hazard only in the context of silicosis from inhalation which might occur during occupational activities as sandblasting or stone cutting (e.g., potential exposure covered under regulations issued by OSHA).

Table 2. Numbers of Chemicals Identified Within TSCA Prioritization Step

TSCA Prioritization Step	# Chemicals
2012 Meeting 1 or more of the Step 1 Factors	1235
2012 Step 2	345
2012 Work Plan chemicals identified	83
2014 Work Plan chemicals updated	90

Potential Activities to Address and/or Update the 2014 Work Plan Chemicals[28]:

[28] It is important to note that this approach will not change or update the existing chemicals on the 2014 Work Plan list. This section describes an approach that could be used to evaluate

A. Use the 2014 Work Plan Chemicals

TSCA, as amended by the Frank R. Lautenberg Chemical Safety for the 21st Century Act, requires that at least 50% of all High-Priority designations be drawn from the 2014 Update of the TSCA Work Plan. As a simple approach, EPA could consider identifying 50-100% of potential high-priority candidates from the 2014 TSCA Work Plan - a static list of chemical substances. Such an approach could, for example, be used as an interim method while EPA continues to refine approaches for identifying potential candidates.

B. Integrate the 2014 Work Plan Methodology with high-throughput screening (HTS) & *in silico* data streams to gather information and identify data needs and assess the chemical landscape (information gathering) for prioritization.

Because the data required to arrive at a priority designation in the 2014 Update is largely a static list derived from data available at the time of the analysis, one approach under consideration is to integrate new hazard, exposure, and potential for persistence and/or bioaccumulation information from new data streams (*i.e.*, data which incorporates alternative testing strategies such as HTS data and *in silico* predictions) into the Work Plan Methodology and to re-screen the 345 Step 2 Work Plan Chemicals. The 345 chemicals in 2012 prioritization (see Table 2 above) would represent the minimal screening set of chemicals (i.e., consistent with the spirit of the 2014 Update) given the exclusion of the chemicals due to TSCA exemptions or other properties (i.e., polymers). However, it is worth noting that identified polymers in the 1,235 chemicals that met one or more of the 2012 factors would be potential candidates for a rapid screening process.

Some of these data streams, as described in Approach 6 within this document, would represent newer exposure modeled estimates, and *in silico* prediction/identification of human health and environmental hazard endpoints not available at the time of the 2014 Update. For instance, chemicals that could not be scored for exposure because of an absence of data (*i.e.*, "Potential Candidates for Information Gathering" in the Work Plan Methodology) would potentially benefit from the integration of New Approach Methodology information and would allow EPA to address additional chemicals that were not able to be screened in Step 2 due to data gaps in exposure or hazard information.

This process would expand the data landscape of the 345 chemicals with additional information not available at the time of the 2014 Update as well as updating scientific methodologies used in the development of these models and technologies for consideration in priority designation. This activity would identify data types (*e.g.*, *in silico*, HTS *in vitro* activity, traditional *in vivo*, *etc.*) as well as data gaps/errors and targeted opportunities to generate data (e.g., *in silico* predictions, *in vitro* and *in vivo* data), if necessary, for conducting risk evaluations. Identification of data types may also inform expert judgement in determining weighting factors for data used in weight-of-

additional existing chemicals that could get added to a new list to help identify chemicals for prioritization.

evidence approaches that may be employed within the prioritization process, as is used in streamlined Canadian approaches and tools described in Approach 2.

C. Update Data Streams and Criteria/Factors Used in the Work Plan Methodology

In the Work Plan Methodology, the Step 2 score was based on three characteristics: hazard, exposure, and potential for persistence and/or bioaccumulation. Another considered approach is to update data sources and to reevaluate criteria for each characteristic used in the screening methodology which would update the 2014 Work Plan Chemicals.

As mentioned, the hazard classification criteria in the Work Plan Methodology was based on the Alternatives Assessment Criteria for Hazard Evaluation[29] developed by EPA's Design for the Environment Program (DfE) which were developed from authoritative sources in 2012. Chemicals were scored on the basis of readily available data, and no judgment was made concerning gaps in or completeness of the available data set for a given chemical. For instance, chemicals that scored as high for hazard only on the basis of acute or chronic aquatic toxicity but that did not present human health concerns were grouped separately as being of potential concern for the environment.

One approach would be to reconsider the criteria and data used in hazard identification. For instance, the criteria for acute aquatic toxicity varies across different programs. For a hypothetical chemical with a fish LC_{50} value of 4 mg/L, the DfE criteria of > 1ppm to 10 ppm would result in a high hazard ranking. However, under the EPA Sustainable Futures/New Chemicals Program, the hazard ranking is moderate since the LC_{50} value meets the criteria of "any of the 3 acute values are between 1.0 mg/L and 100 mg/L" ("any 3 acute values" meaning any three LC50 values from fish, daphnid or algae toxicity studies).[30] Additionally, depending on the data source used, updates may not have been incorporated into the available data at the time the Work Plan Methodology was implemented. New CDR and TRI data was used to update the exposure score in the 2014 Update. However, acute aquatic toxicity studies based on the Aster model results were sourced from an older version of Canada's Ecological Categorization Results on the Canadian Domestic Substances List. During the 2014 Update, Canada updated their Ecological Categorization Results on the DSL for many chemicals, but EPA did not update the corresponding hazard ratings for the 2014 Work Plan update. Therefore, updates to more current data streams, as described in Approach 6, would help harmonize the use of specific sources within the Work Plan Methodology and reduce any potential for errors. In a similar way, hazard criteria for human health hazard identification may differ across programs and may require examination and reconsideration in this approach.

Further modification of the Step 1 factors may also be considered appropriate given the criteria specified in the Prioritization Rule 40 CFR section 702.9 and TSCA section 6(b)(1)(A). : (1) The chemical substance's hazard and exposure potential; (2) the chemical substance's persistence and

[29] https://www.epa.gov/saferchoice/alternatives-assessment-criteria-hazard-evaluation
[30] Sustainable futures link: https://www.epa.gov/sites/production/files/2015-05/documents/06.pdf

bioaccumulation; (3) potentially exposed or susceptible subpopulations; (4) storage of the chemical substance near significant sources of drinking water; (5) the chemical substance's conditions of use or significant changes in conditions of use; (6) the chemical substance's production volume or significant changes in production volume; and (7) other risk-based criteria that EPA determines to be relevant to the designation of the chemical substance's priority. For instance, storage near significant sources of drinking water may be considered as part of the selection criteria for high-priority candidates and may reduce the burden on the Prioritization process itself as a pre-filter to include or exclude candidates before Prioritization.

D. Integrate activities using a sector analysis approach and functional use.

Using a functional use approach, such as those described in Approaches 4 & 5 within this document, for identifying chemicals of interest within the 2014 Work Plan Chemicals would facilitate or inform the development of chemical categories or groupings. These chemical categories would provide a basis for similar data needs by class or analogues and would potentially inform streamlined assessments or rapid screening approaches for select categories of chemicals much like those developed by the Government of Canada (i.e., Polymer Rapid Screening I and II)[31].

E. Other Updates

EPA will consider other updates or changes through feedback received at the December 11, 2017 public meeting or during the associated opportunity for comment as part of the ongoing dialogue. Suggestions solicited through discussion of other changes that can be made may include new approaches and new data streams not previously considered in the development of the Work Plan Methodology or the presented approaches in this document.

Benefits

- TSCA, as amended by the Frank R. Lautenberg Chemical Safety for the 21st Century Act, requires that 50% of all High-Priority designations be drawn from the 2014 Update of the TSCA Work Plan. The 2014 Work Plan chemicals would satisfy this requirement and would represent a pragmatic consideration (time, effort and resources) to begin using these chemicals as a starting point given the efforts and resources already devoted to its development and update.
- The criteria used in the two-step prioritization process as described in the Work Plan Methodology satisfies many of the criteria specified in the Prioritization Rule 40 CFR section 702.9 and TSCA section 6(b)(1)(A) – hazard and exposure potential of the chemical substance, persistence and bioaccumulation, and conditions of use with high

[31] https://www.canada.ca/en/health-canada/services/chemical-substances/chemicals-management-plan/initiatives/polymer-rapid-screening-approach.html

exposure potential. The Work Plan Methodology also has specific criteria to account for potentially exposed and susceptible subpopulations as stated by amended TSCA.

- Updating this approach would be responsive to public comments EPA has received since 2014. In addition, advances in New Approach Methodologies (*e.g.*, data science, *in silico* models, HTS *in vitro* assays, *etc*) may provide mechanisms to fill data gaps that were identified in the original methodology.

Caveats

- The 2014 Work Plan Chemicals represent a static snapshot of the data and priorities (as specified by criteria and factors) at a given time. The data and criteria incorporated into the screening methodology may not be representative of the current state of science or information available at the time of the 2014 Update. Any future approach to update the methodology would need to verify updated data sources and models to ensure harmonization of data with external and internal data and model streams.
- Finally, the Work Plan Methodology is more aligned towards identifying candidates for High-Priority designation than identifying Low-Priority designations. This will require having an alternative mechanism to identify candidates for Low-Priority designation, such as Approach 3 described in this document.

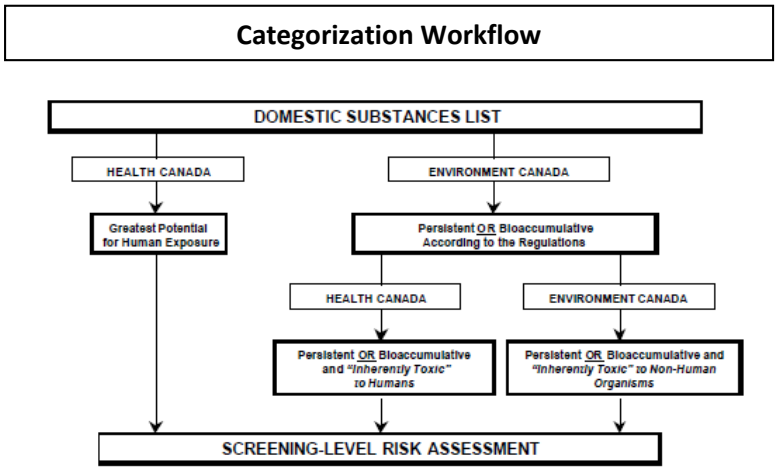
2) Canada's Chemical Management Plan (CMP)

Background

What is the Chemical Management Plan? Canada's Chemical Management Plan[32] (CMP) was designed to help Canada meet goals set by the World Summit on Sustainable Development for the sound management of chemicals by 2020. The CMP integrates Canadian federal programs into a single strategy to ensure that chemicals are managed appropriately in order to prevent harm to Canadians and their environment. The CMP provides a plan for the assessment and management of approximately 4300 existing substances that were identified as priorities from the categorization exercise conducted from 1999 – 2006. The principal legislation behind the CMP is the Canadian Environmental Protection Act, 1999 (CEPA 1999).

Categorization of the Domestic Substances List Overview

Categorization was a prioritization exercise and was required by the CEPA 1999. Once a substance met the categorization criteria, a screening assessment was required to determine if the substance posed a risk to the environment or to human health. CEPA 1999 set a goal for the Government of Canada to sort through or "categorize" all 23,000 chemical substances on their Domestic Substances List (DSL) which included substances in Canadian commerce, used for commercial manufacturing purposes, or manufactured in or imported into Canada in a quantity of 100 kg or more between 1984 and 1986. This task was completed by September 2006, as required by the Act.



[32] Government of Canada website on Chemical Substances. https://www.canada.ca/en/health-canada/services/chemical-substances.html

Categorization prioritized substances on the DSL using criteria for persistence (P), bioaccumulation (B) and inherent toxicity (iT) to humans and non-human organisms, or greatest potential for human exposure (GPE). (See "Categorization Workflow" figure.) Substances that met the criteria for P or B and iT or GPE were then considered for further screening assessment. In order to facilitate the process, tools were developed for identification and screening of substances considered to be priorities for either human health or environmental perspectives.

Box 1: Categorization lessons learned:

- Limitations to conducting a priority-setting exercise based on dated inventory data.
- Need for consideration of exposure to ecological receptors and use of weight of evidence approach during prioritization; so far less than 10% of substances identified as priorities based only on ecological hazard categorization criteria were found to pose a risk following assessment.
- Streamlined approaches like rapid screening for low volume chemicals were necessary.
- Direct exposures (i.e., consumer and children's products) typically key driver in human health assessment outcome, but it is often difficult to identify uses of substances in these products.
- Unable to model Persistence and Bioaccumulation properties for substances with challenging chemistries – *e.g.*, ionizing chemicals, persistent/mobile/ toxic chemicals, Unknown or Variable Composition, Complex reaction products, and Biologicals (UVCBs).
- Categorization of inorganics was able to use a metal moiety approach.
- Strong stakeholder engagement was important.
- Need to invest in development of efficient/novel ways to facilitate prioritization and assessment – *i.e.*, QSAR tools at the time of prioritization /categorization.

Categorization utilized both experimental and modelled values. For example, in both human exposure and hazard, simple discriminating tools were applied initially to focus on highest priorities. For exposure, three criteria were used to estimate the human exposure potential of a substance: the quantity reported to be in commerce between 1984 and 1986, the sum of expert ranked use codes, and number of submitters. More complex tools were subsequently applied to additionally refine both estimates of exposure and identification of hazard through various sources of information. Complex hazard tools were developed utilizing a hierarchical approach for consideration of multiple endpoints to human health from various sources of relevant information – including both experimental and modelled data. The approach was protective, with conservative choices being made in the absence of data. It identified priorities for assessment and appropriately weighted persistence and bioaccumulation in the context of their potential to contribute to human exposure. In the ecological arena, 90-95% of data used to compare with threshold criteria were modelled values due to the lack of existing empirical data.

The outcome of Categorization was the identification of approximately 4300 substances requiring further consideration which led to the creation of the CMP, under which the majority of Canada's risk assessment work is now focused. Several lessons were learned from this exercise and have influenced the implementation of the CMP. One of the key lessons learned was the need to invest in the development of efficient/novel ways to facilitate prioritization and assessment – *i.e.*, Quantitative structure–activity relationship (QSAR) tools at the time of

prioritization /categorization, as well as using up-to-date inventory information and application of streamlined rapid screening assessment tools (See Box 1 for "Categorization lessons learned").

Evolution of the Chemical Management Plan

To date, the CMP has been rolled out in 3 phases[33] with each phase building on lessons learned in the previous phase. From 2006 to 2011, CMP Phase 1 addressed 1064 substances using the best available traditional toxicity data and QSAR modeling including substance by substance assessment of the 200 substances identified by Categorization as being of highest priority. This phase also included the development of a rapid screening approach to assess substances expected to be of low concern due to low volume of use. The rapid screening approach for substances of low concern made use of both qualitative and quantitative steps to efficiently evaluate the likelihood that a substance may cause harm, given conservative estimates of exposure. At each step in the rapid screening process, any substance that appeared to present a potential for harm was identified as requiring further assessment.

From 2011 to 2016, CMP Phase 2 expanded the use of alternative approaches such as *in silico* modeling and read-across on 1700 substances, but also developed grouping initiatives such as, aromatic azo- and benzidine-based substances, substituted diphenylamines and phthalates, as well as moiety based approaches for metals. Rapid screening approaches were expanded and developed for polymers as well.

Since 2016, CMP Phase 3 addresses the remaining priorities identified from Categorization. The integration of emerging data (i.e., NAM) and novel approaches to address substances with limited data sets are being explored. Other streamlined approaches that have been developed and are being extended in application include: the use of the Ecological Risk Classification (ERC) platform which uses a weighted multi-descriptor profile approach to classify ecological risk, the use of Toxicological Thresholds of Concern (TTC) approach to establish a human exposure threshold value for chemicals with a low probability of risk to human health, as well as other rapid screening approaches and computationally derived margin of exposures - *i.e.*, bioactivity-exposure ratios (BER).

On-going Prioritization: Approach for Identification of Risk Assessment Priorities (IRAP)

Since 2006, priorities for risk assessment of chemicals and other substances under the CEPA have largely been based on the results of Categorization and New Substances Notifications. In 2014, an approach was published outlining the systematic collection, consolidation and analysis

[33] Government of Canada Chemical Management Plan website: https://www.canada.ca/en/health-canada/services/chemical-substances/chemicals-management-plan.html

of new information, in order to determine appropriate action, including possible risk assessment, for substances with new information. The approach describes the ongoing prioritization activity that contributes to the identification of risk assessment priorities (IRAP) for chemicals and polymers.

There are three steps involved in the identification of risk assessment priorities: **acquisition** of information relevant to the potential health and ecological risks of substances, **evaluation** of the information available for each substance, and identification of appropriate **action** for each substance. The process is different from Categorization, where each substance on the DSL was categorized based on prescribed criteria. With the IRAP process, new information is evaluated against numerous guiding principles and considerations to determine appropriate action for implicated substances. Generally, for a substance to be identified as a priority for risk assessment, the process identifies information for a potential risk – that is, the presence of both a hazard and a significant potential for exposure in Canada. The acquisition of new information occurs on an ongoing basis, while the other two steps are generally performed at regular intervals.

Snapshot of the CMP Chemicals

The total number of chemicals screened and prioritized according to Categorization and the CMP Phases is given in Table 1. Broad chemical groupings of the approximately 4300 categorized substances are displayed in Figure 1. Notably, ~26% of the categorized substances are polymers, and petroleum substances which may be amenable to sector approaches, such as Approach 5 described in this document.

Table 1. Total Number of Chemicals Screened in Categorization and CMP Phase 1 – 3

Snapshot of the CMP Chemicals	# Chemicals
Domestic Substances List	~23000
Categorized Substances/Chemical Management Plan (CMP)	~4300
CMP Phase 1	1064
CMP Phase 2	1500
CMP Phase 3	1700

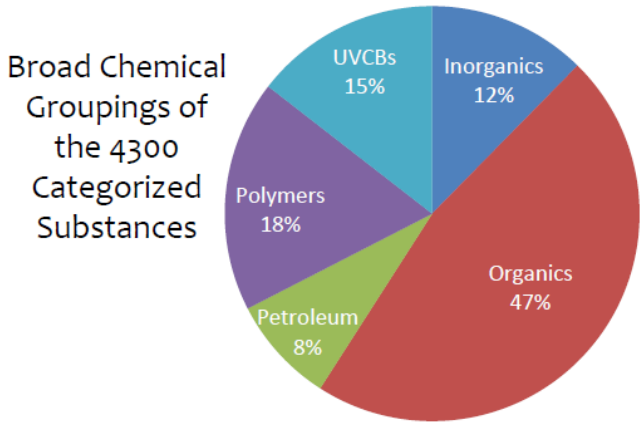


Figure 1. Broad chemical groupings of the Categorized Substances

Proposed activities to enhance the identification of candidates for high priority chemicals:

Consultation with both Environment & Climate Change and Health Canada on the Categorization/CMP process to share "lessons learned" and use tools developed by Canada.

EPA has consulted with the Government of Canada to share relevant information on the prioritization processes adopted by each jurisdiction. Discussions have been held and information was exchanged to better understand the process of Categorization within the CMP, with the focus on the inclusion of *in silico* and analog data in a way that would be consistent with the data landscape (*i.e.*, data need) of each country's high priority substances. One approach to consider would focus on how Canada's approach could be more tailored towards EPA's current needs under TSCA as well as current efforts to integrate modelled data and more complex modeling approaches through new data streams, such as EPA's Chemistry Dashboard/RapidTox initiatives, as a possible mechanism to address data gaps and to assess the data quality issue when used in a prioritization and risk evaluation workflows.

General consideration for use of models and data within the Canadian approach would require additional work to ensure that TSCA statutory requirements are met.

For instance, work would be required to investigate the use of streamlined approaches such as the Ecological Risk Classification approach developed by Canada in the context of TSCA and the Prioritization Rule. While many similar criteria specified in the Prioritization Rule 40 CFR section 702.9 and TSCA section 6(b)(1)(A) overlap with regulations and approaches used by Canada – e.g., hazard and exposure potential of the chemical substance, persistence and bioaccumulation, and conditions of use with high exposure potential – there are significant differences based on jurisdiction and provincial legal requirements in Canada. For instance, the CMP Categorization process does not consider worker exposures under conditions of use as Canadian federal regulations defer worker exposure assessments to the provincial or regional jurisdiction. Proximity to significant sources of drinking water would need to be considered as well in any adaptation of the Canadian approaches for use in TSCA. Another approach to consider is the use of the Categorized substances and their data as an inventory to screen. In

consideration of this approach, all statutory requirements of TSCA would again need to be considered before adapting this inventory of chemicals.

Proposed activities to enhance the identification of candidates for low priority chemicals:

Share/crosswalk of chemical inventories between EPA's Safer Chemical Ingredients List and Canada's substances of low concern [ongoing with EPA]

A crosswalk of inventories between the SCIL and the Canadian substances of low concern that were rapidly screened in CMP Phase 1 could inform the selection of low priority candidates. This effort would help identify potential candidates for low priority designations as well as additional data requirements through a coordinated effort to share available public information.

Benefits

- The Canadian approach incorporates external peer-review of approaches developed and used within Categorization and the CMP. Further, open public and stakeholder comment periods have also been taken into account in the development and use of these tools.
- The Canadian approaches[34] are receptive to integrating modelled values within the prioritization steps (i.e., Categorization) as well as using rapid screening approaches. Tools such as the complex Hazard screening tool integrate QSAR estimates through a hierarchical approach for consideration of multiple endpoints to human health including developmental and reproductive toxicity. Integrated approaches such as those proposed by the EPA's National Center for Computational Toxicology (NCCT) data streams within the TSCA Work Plan 2-step prioritization process would potentially incorporate updated state-of-the-science that are potentially equivalent to (or better than) those used with Categorization and the CMP phases.
- The use of the Categorization substances classified as low concern would also augment the current approaches to identify low-priority candidates.

[34] Guidance Manual for the Categorization of Organic and Inorganic Substances on Canada's Domestic Substances List Existing Substances Branch, Environment Canada (June 2003)

Final Integrated Framework for the Health-Related Components of Categorization of the Domestic Substances List Under CEPA 1999, Health Canada (2009)

Approach for identification of chemicals and polymers as risk assessment priorities under Part 5 of the Canadian Environmental Protection Act, 1999 (CEPA 1999), Health Canada and Environment Canada (2014)

Science Approach Document for the Threshold of Toxicological Concern (TTC)-based Approach for Certain Substances. Health Canada (Sept 2016)

Science Approach Document for the Ecological Risk Classification of Organic Substances. Environment Canada and Climate Change Canada (July 2016)

Caveats

- As mentioned previously, additional work would be required to verify that the Canadian approaches (models and data used) are consistent with our statutory requirements of TSCA and the Prioritization rule.
- CMP Categorization process does not consider worker exposures under conditions of use as Canadian federal regulations defer worker exposure assessments to the provincial or regional jurisdiction. This would require additional work to validate assumptions used in the exposure criteria.

3) The Safer Chemical Ingredients List (SCIL)

The EPA's Safer Choice program's Safer Chemical Ingredients List[35] is a list of low hazard chemicals that could be a complement to any other organizing tool to identify chemicals for prioritization under TSCA. While this section references chemicals listed on SCIL, and is derived largely from the ingredients in cleaning and related products, other low hazard and/or exposure candidate chemicals used in other sectors and not currently listed on SCIL may also be candidates for prioritization.

Background on SCIL

The Safer Chemical Ingredients List (SCIL) is the EPA's Safer Choice program's listing of chemical ingredients that meet its low-concern criteria and are generally acceptable for use in Safer Choice-labeled products. The low hazard thresholds used as criteria for chemicals on SCIL can be found at https://www.epa.gov/saferchoice/standard. SCIL is organized by functional-use classes. Chemicals on SCIL are assigned geocodes reflecting their hazard profile and available data. The SCIL geocodes include:

Green circle: Chemicals marked with full green circles are considered low hazard based on experimental or modeled data. There are 605 green circle chemicals on SCIL.

Green half-circle: Chemicals that are marked with green half-circles are expected to be of low hazard based on experimental or modeled data. Additional data would strengthen confidence in the chemical's status. Currently, there are 102 green half-circle chemicals.

Yellow triangle: Chemicals with the yellow triangle geocode meet Safer Choice Criteria for their functional ingredient class, but there are some hazard profile issues. There are 210 chemicals listed as yellow triangles on SCIL.

Why was SCIL developed?

SCIL was developed by EPA in 2012 as a list of safer chemicals available to product manufacturers, including those who now use the list to formulate Safer Choice-labeled products. SCIL was developed to meet stakeholder demand, especially from product manufacturers, for transparency and information on chemicals that can be used in Safer Choice-labeled products. The first 435 SCIL chemicals listed were from Safer Choice-labeled products. Based on stakeholder feedback, EPA believes the SCIL is a useful resource for chemical and product manufacturers that encourages green chemistry and innovation. SCIL geocodes signal an opportunity for innovation in functional categories lacking a robust, highly functioning set of full-green-circle chemicals.

[35] The Safer Choice program allows product manufacturers to use the Safer Choice label on products made up only of safer chemicals. The SCIL contains chemicals that the Safer Choice program has evaluated and determined to meet the program's safer chemical criteria.

Since 2012, the Safer Choice program has expanded SCIL by adding chemicals from newly certified products, close structural analogs, chemicals based on research into potential new products, and through chemical manufacturers submitting their chemicals for listing on SCIL. As of October 2017, there are 917[36] chemicals and 987 listings on SCIL.

Table 1. The chemical listings by SCIL functional-use class.

Antimicrobial Actives (7)	Polymers (59)
Chelating Agents (22)	Preservatives & Antioxidants (34)
Colorants (44)	Processing Aids & Additives (149)
Defoamers (12)	Skin Conditioning Agents (46)
Emollients (26)	Solvents (67)
Enzymes & Enzyme Stabilizers (30)	Specialized Industrial Chemicals (14)
Fragrances (152)	Surfactants (282)
Oxidant & Oxidant Stabilizers (19)	Uncategorized (24)

Why could SCIL be a good source of candidate low-priority substances?

As shown above, SCIL chemicals represent and cover a range of functional classes. Their use in Safer Choice-labeled products allows Safer Choice to have a deep understanding of these chemicals, the products they are used in, and a meaningful subset of conditions of use. The low toxicity thresholds from Safer Choice criteria used for SCIL may provide support for low-priority proposals. These criteria cover many of the considerations specified in the Prioritization Rule (40 CFR section 702.9) – hazard and exposure potential of the chemical substance, persistence and bioaccumulation, and a subset of conditions of use with high consumer and worker exposure potential.

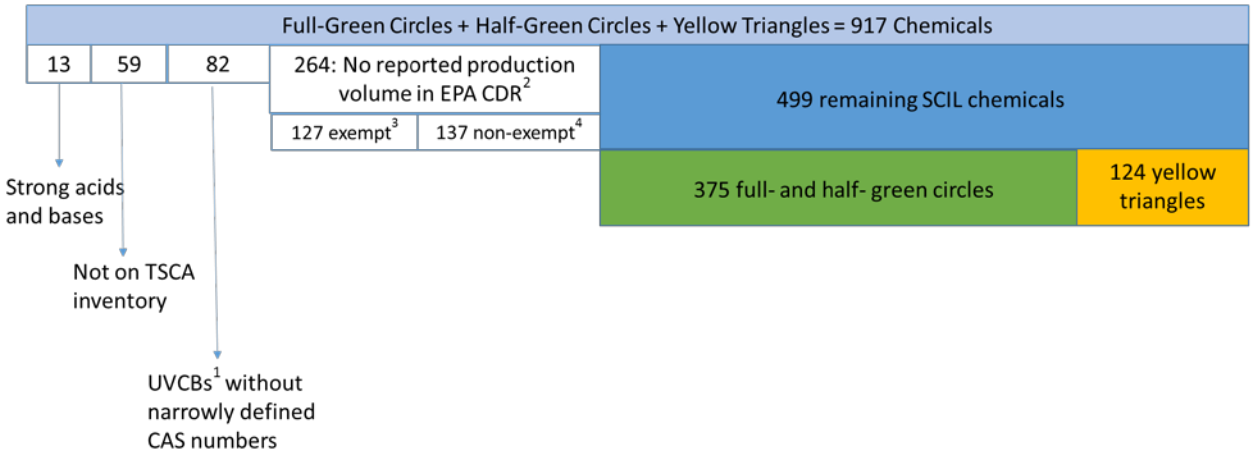
SCIL could also serve as a complement to other proposed approaches that integrate exposure data.

Chemicals on SCIL could serve as the starting point for identifying potential low-priority candidates. The full- and half-green circles on SCIL denote chemicals with toxicological properties that are very low concern for hazard. Chemicals marked with yellow triangles represent the best in class for their functional uses, but are not low concern for all hazard endpoints. Some chemicals on SCIL may be removed as potential candidates because of their toxicological and chemical properties. Examples include: strong acids and bases that may have high acute hazard when assessed under all conditions of use (13 chemicals qualify for use in Safer Choice-labeled products, but with use restrictions); chemicals not on the TSCA Inventory (59 chemicals); and UVCB chemicals[37] with CAS numbers that include a range of chain lengths

[36] Includes chemicals marked with full green circles, green half-circles, and yellow triangles.
[37] The 82 UVCB chemicals identified in the diagram are chemicals with CAS numbers that have greater than 4 carbons or ethoxylation and/or propoxylation (82 chemicals qualify for Safer Choice with use restrictions to qualify for Safer Choice). Further, a subset of SCIL chemicals do

or ethoxylation and/or propoxylation that present challenges in interpretation of toxicity data. Designating chemicals with high production volumes may maximize the benefits of chemical prioritization. With those caveats, the remaining 499 chemicals could be proposed as candidates. These chemicals are on the TSCA inventory, are actively produced, have relatively low hazard profiles, and cover a range of functional uses.

Figure 1. One proposed approach to organize SCIL chemicals.



[1]UVCB: Unknown of Variable compositions, Complex reaction products and Biological materials.
[2]Environmental Protection Agency Chemical Data Reporting Rule under TSCA
[3] Under 40 CFR, Part 711, TSCA data reporting requirements provide chemical classes that exempt or non-exempt for chemical data reporting. Chemicals exempt under this data reporting rule may still be candidates.
[4]Non-exempt chemicals are either not in production or have production volumes less than 25,000 lbs.

not have reported aggregate volumes in the 2016 CDR (264 chemicals). Some of these chemicals are exempt from reporting based on 40 CFR 711.6 exemptions (e.g., polymers, naturally occurring substances, polysaccharides, etc.), while other chemicals may not be in production or have production volumes less than 25,000 pounds.

Table 2. This table illustrates that a number of SCIL chemicals have high production volumes. Production volumes for chemicals marked with full- and half-green circles are shown below.

Production Volume (lbs.)	# SCIL Candidates
>10,000,000	140
1,000,001-10,000,000	100
500,001-1,000,000	25
100,000-500,000	50
<100,000	60
Total	*375*

Table 3. The remaining chemicals marked with yellow triangles on SCIL are best in class for a given functional use. An approach to include them proposes 41 yellow triangle chemicals with the production volumes shown below.

Production Volume	# SCIL Candidates
>10,000,000	23
1,000,001-10,000,000	6
500,001-1,000,000	1
100,000-500,000	5
<100,000	6
Total	*41*

Table 4. Potential Candidates by SCIL functional-use class. (In the parentheses, the first number represents the total in that functional class[38], the second, the number in that class with production volumes over 1 million lbs.).

Antimicrobial Actives (6, 6)	Polymers[39] (7, 6)
Chelating Agents (14, 13)	Preservatives and Antioxidants (13, 13)
Colorants (11, 9)	Processing Aids and Additives (93, 79)
Defoamers (5, 4)	Skin Conditioning Agents (13, 10)
Emollients (13, 10)	Solvents (44, 37)
Enzyme Stabilizers (4, 4)	Surfactants (141, 107)
Fragrances[40] (35, 20)	Other (12, 8)
Oxidant and Oxidant Stabilizers (2, 2)	

Benefits

SCIL is an available EPA resource made up of lower-hazard chemicals. The listings are supported by toxicological data and used in many consumer and institutional and industrial products. Further, SCIL is supported, understood, and used by many stakeholders. As a source of low-hazard chemicals, SCIL could complement other approaches discussed in this document by contributing candidate low hazard chemicals with functional uses similar to those for chemicals that may be identified as high hazard chemicals.

Caveats

To satisfy statutory criteria in the rule for SCIL-listed chemicals, during prioritization, EPA would have to further investigate storage near significant sources of drinking water, the possibility of impacts on potentially exposed susceptible sub-populations, and understand all conditions of use for a given chemical. Further, the low hazard data supporting the SCIL geocodes may need to be updated to account for any new information or New Approach Methods (NAMs) to reach a threshold of data sufficiency for prioritization.

[38] Several SCIL chemicals fall under two or more functional classes and are listed in all appropriate functional classes.
[39] Polymers included in this list did not contain "*polym," "*alkyd," or "*oxylated" in the Chemical Abstract Index Name in the Master Inventory File, and thus were not considered exempt from CDR reporting (40 CFR 711.6), but have been classified as a polymer under SCIL.
[40] Included fragrances are listed as full- or half-green circles on SCIL.

4) Functional Category Approach, based on Use and Exposure Potential

This section describes how a functional use category approach could be used to identify groups of candidate chemicals with similar functional uses for prioritization, and the benefits of such an approach to EPA, industry, states, NGOs and others after final priority designations have been made.

The functional use categories proposed in this approach are based on the EPA's Chemical Data Reporting (CDR)[41] and the Organization for Economic Co-Operation and Development (OECD) use categories found in the document "Internationally Harmonized Functional, Product and Article Use Categories."[42] The CDR functional use categories represent the manner in which the chemical substance is used. They are listed at 40 CFR 711.15(b)(4)(i)(C) and 40 CFR 711.15(b)(4)(ii)(A). A copy of the list of categories is included in Appendix 1 to this approach.

Background

A functional use category approach could be applied to identify a group of chemicals that share a particular function in industrial processes, chemical formulation (e.g., for preservation or blending), or end-use product level (e.g., surfactants[43] for use in cleaning products for industrial or consumer use). The benefits of this approach include:

- Chemical and product manufacturers regularly apply a functional use category approach to design and formulate products. CDR information includes the functional use category for the chemical and the broader stakeholder community is familiar with the concept of functional use.
- The statute requires consideration of conditions of use as part of the prioritization process. By grouping functionally related chemicals, this provides a better understanding of the uses within the cluster. This process could identify a more robust description of conditions of use across the cluster, and could later be useful during risk evaluation.

Overview – How functional use categories can be an organizational tool

This approach proposes a four-step process that focuses first on the exposure potential based on the functional use of the chemical. Functional use categories could be divided into tiers based upon the exposure potential, using methodology developed for the TSCA Work Plan and further refinements to the process based on lessons learned since 2012, the assessments of the first 10 priority chemicals, and using information from other tools (e.g. High-Throughput Screening and Computational Modeling (Approach 6)). The second step would rank functional use categories within a tier using additional exposure potential factors. The third step involves a data adequacy and risk screen on the chemicals within a functional use category. This step can help identify

[41] Additional information regarding the Chemical Data Reporting can be found at: https://www.epa.gov/sites/production/files/2016-05/documents/instructions_for_reporting_2016_tsca_cdr_13may2016.pdf.

[42] OECD use categories can be found at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2017)14&doclanguage=en.

[43] Also known as "surface active agents" per the CDR

those functional use categories with sufficient data for the chemicals to be prioritized as well as providing an opportunity to screen chemicals within a functional category that could be of high or low risk. The final step presents options on how chemicals from a functional use category could enter prioritization, particularly because many functional use categories could have large numbers of chemicals.

Step 1: Classification of functional use categories. The first step in this approach would consist of classifying functional use categories based on their exposure potential during manufacturing, processing and use in a final product. An example of how the tiers could be determined are:

- Tier 1 functional use categories include consumer (e.g. children) products widely used and with a high likelihood of exposure.
- Tier 2 functional use categories include other consumer, commercial, and industrial uses with a high likelihood of exposure.
- Tier 3 are the remaining functional use categories.

The Tiers are intended to reflect different types or levels of exposure potential, the assumption is that many industrial and commercial operations will have overarching health and safety procedures in place to minimize exposures, while consumers usually assume that products that they buy are "safe to use." There can be substantial overlap between consumer and commercial products; therefore, a consumer product widely used and with a high likelihood of exposure might also have commercial applications. Finally, children's exposures are accounted for under exposures to consumer products.

The classification of functional use categories into Tier 1, 2, and 3 could be made without a specific high-priority chemical in mind, and could be based on information reported under the 2016 CDR to account for how often a functional use category is reported as used in consumer products, commercial products, or during industrial processes. Other sources can also be used, e.g. Toxics Release Inventory or TRI or Consumer Product Safety Commission Product Evaluations. Additional consideration could be given to the experience in developing "Preliminary Information on Manufacturing, Processing, Distribution, Use, and Disposal"[44] documents for the first 10 priority chemicals and the persistent bioaccumulative toxic (PBT) chemicals subject to TSCA section 6(h). (e.g. some flame retardant chemicals are used in consumer and children's products, even though CDR reporting treats flame retardants as industrial use only). The tiers are intended to reflect different types or levels of potential exposure, recognizing that actual exposures depend on the specific conditions of use, the properties of the chemicals, and additional exposure controls in place.

Step 2. Further ranking of functional use categories. This tiering would be further refined by considering other exposure-related factors. Factors could include: chemicals associated with the

[44] https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/public-meeting-risk-evaluation-scoping-efforts-under

functional use are used by more than 1,000 workers and total volume of chemical used; uses with likely water releases or with potential to contaminate drinking water sources; uses that will result in discharge down the drain through residential or commercial use, or in commercial settings; functional uses involving spray application or emissive uses associated with volatile chemicals in indoor settings where susceptible subpopulation, such as children, could be exposed; other considerations such as continuous use of a chemical or the chemical is actively transported; etc. Two or more factors could be used at this step to rank the categories.

Step 3. Initial screening of the functional use categories to select the functional use categories with adequate data and a suitable number of chemicals for prioritization. Before a functional use category could be considered for prioritization, this step would serve as an initial screening to identify if the chemicals within a given functional use category have sufficient data and if the functional category has a suitable number of chemicals with possible high and low risk. In this third step, the chemicals within the functional use categories identified in steps 1 and 2 (e.g. the first 5, 10, 15, or 20 functional use categories) could be screened using the High-Throughput Screening and Computational Model (Approach 6). The screening methodology could identify if the chemicals within a functional use category have adequate data for prioritization. And, based on the screening methodology, identify if the chemicals within a functional use category have a potential for high or low risk.

If more than 50% of the chemicals within a functional use category have sufficient data, then the functional use categories associated with those chemicals would be further considered; while those functional use categories consisting of chemicals lacking sufficient data would be put in a queue for data gathering actions under TSCA and via other means. This step is intended to ensure that chemicals that enter prioritization have enough data to make a determination based on the prioritization criteria, and it is not intended to focus only on data rich chemicals. Data gathering actions could include a broad range of activities, such as modeling, the High-throughput Screening and Computational Model (Approach 6), additional testing, additional data collection such as through TSCA section 8(d), etc. Identification of groups of chemicals lacking sufficient data would allow industry to coordinate data gathering efforts.

Step 4. Scheduling of prioritization. Since the statute requires a minimum of 20 ongoing risk evaluations of high priority chemicals and an initial requirement of identifying 20 low priority chemicals, and each functional use category might contain a much larger number of chemicals, there is a need to schedule how the chemicals within a functional use category would be considered for prioritization in order to manage the number that could subsequently enter risk evaluation. Two example options for scheduling chemicals within a category for prioritization are presented below.

Option 1. Multifunctional chemicals. Considering a number of chemicals will fall under multiple functional use categories, the functional use categories could be compared to determine if there is significant overlap of chemicals. The chemicals that are used by two or more functional use categories (which have been selected based on the above tiering) could be considered for prioritization first. Since once a chemical enters risk evaluation all uses of the chemical are

considered, this would potentially allow for multiple functional use categories to be addressed at the same time.

Option 2. A second option could be to stagger the prioritization of chemicals from several functional use categories (based on the above tiering). This approach could be used where there is little overlap in chemicals among functional use categories.

Figure 1. Ranking Functional Use Categories to Identify Chemicals for Prioritization

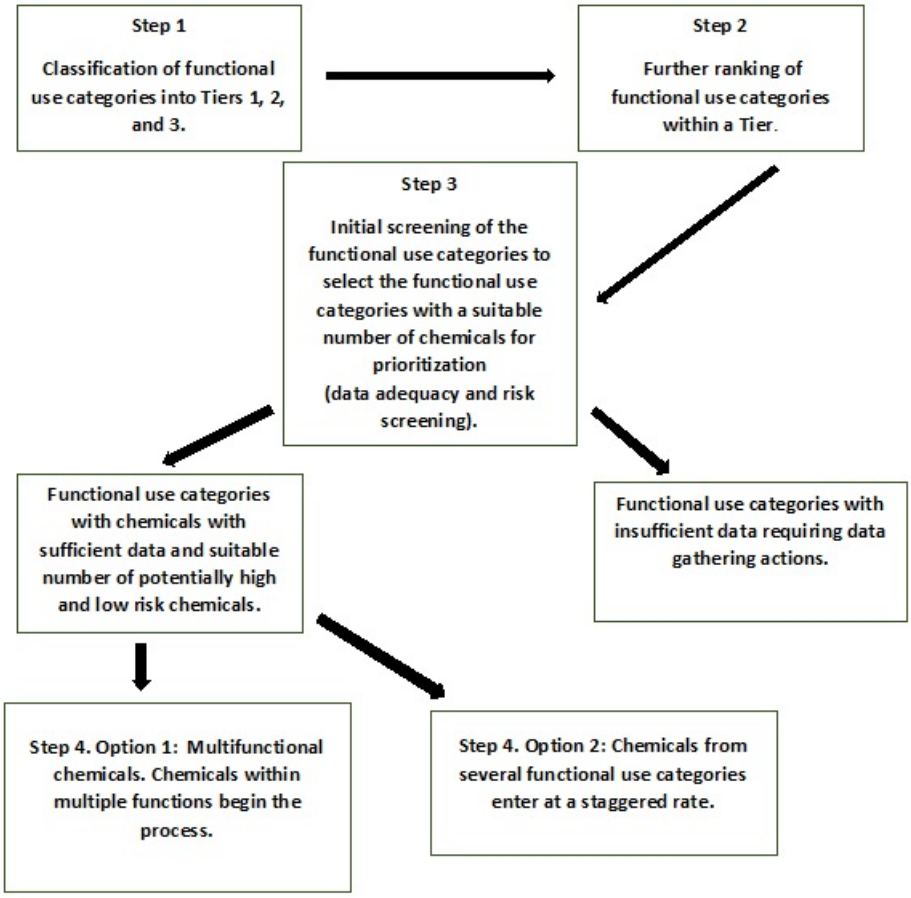


Figure 2. Step 4. Option 1. Multifunctional chemicals. Chemicals with multiple functional use categories begin the process. Chemicals that are common to two functional use categories with high tiering (the overlapped area in Categories A and B below), would enter prioritization first (first 1-3 years) and those chemicals that are only in one functional use category (only Category A or Category B), would enter prioritization later (years 4- 7).

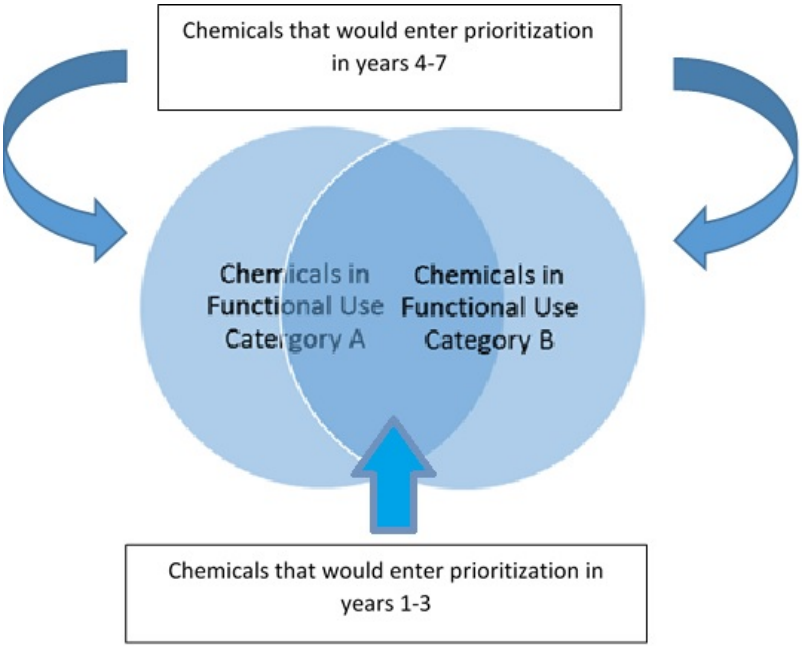
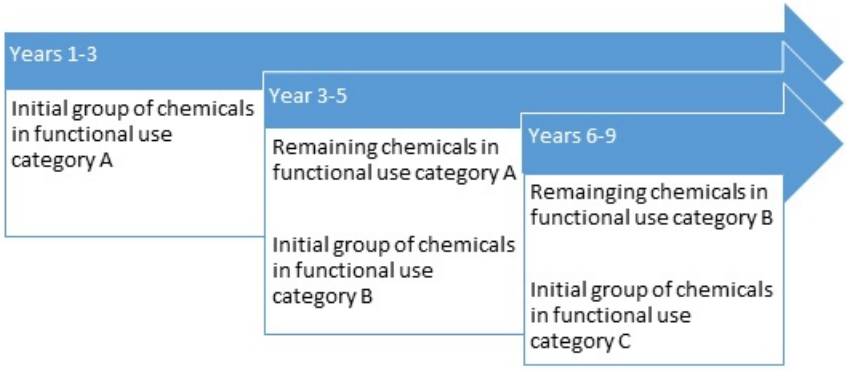


Figure 3. Step 4. Option 2. Chemicals from several functional use categories enter at a staggered rate.



Benefits

- Grouping chemicals with similar functional uses can lead to:
    - Efficiencies in chemical assessment where chemicals have similar use and exposure patterns.
    - A smoother substitutes transition for industry given that EPA would be assessing all the chemicals within the same functional use.

    - Identifying low-priority designations for a given functional use category to help ensure the availability of alternative chemicals, prevent unfortunate substitution and address uncertainty in the marketplace.
- For downstream risk management: Section 6 of TSCA requires the Agency to consider, to the extent practicable, whether technically and economically feasible alternatives for a specific condition of use will be reasonably available as a substitute when the proposed prohibition or other restriction takes effect. By considering functional use categories, EPA will have more complete information on which to base eventual risk management decisions.

Caveats

- The process outlined focuses on exposure potential related to the functional use category. Therefore, there is a possibility that a high hazard chemical that does not have a functional use category with high exposure potential might not be selected early for the prioritization process.
- Selection of a chemical based on consumer exposures or a chemical within the Tier 1 functional use category will not preclude the chemical from being evaluated under all conditions of use. The risk evaluation considers all conditions of use, consumer, industrial and commercial uses.

Stakeholder Involvement

Maximizing stakeholder involvement will be important, especially to:

- Ensure a solid understanding of the functional use categories and the use patterns of chemicals within the category;
- Focus opportunities to gather additional information to classify a particular functional use category as Tier 1, 2, or 3;
- Identify additional data/criteria to further group the functional use categories within each Tier; and gather additional information to identify a broad range of chemicals for each functional use category.

Appendix 1

EPA's Chemical Data Reporting: 40 CFR 711.15(b)(4)(i)(C): "…codes from Table 8 of this paragraph must be selected to designate the industrial function category(ies) that best represents the specific manner in which the chemical substance is used."

Table 8. Codes for Reporting Industrial Function Categories

Code – Category
U001 Abrasives
U002 Adhesives and sealant chemicals
U003 Adsorbents and absorbents
U004 Agricultural chemicals (non-pesticidal)
U005 Anti-adhesive agents
U006 Bleaching agents
U007 Corrosion inhibitors and anti-scaling agents
U008 Dyes
U009 Fillers
U010 Finishing agents
U011 Flame retardants
U012 Fuels and fuel additives
U013 Functional fluids (closed systems)
U014 Functional fluids (open systems)
U015 Intermediates
U016 Ion exchange agents
U017 Lubricants and lubricant additives
U018 Odor agents
U019 Oxidizing/reducing agents
U020 Photosensitive chemicals
U021 Pigments
U022 Plasticizers
U023 Plating agents and surface treating agents
U024 Process regulators
U025 Processing aids, specific to petroleum production
U026 Processing aids, not otherwise listed
U027 Propellants and blowing agents
U028 Solids separation agents
U029 Solvents (for cleaning and degreasing)
U030 Solvents (which become part of product formulation or mixture)
U031 Surface active agents
U032 Viscosity adjustors
U033 Laboratory chemicals
U034 Paint additives and coating additives not described by other categories
U999 Other (specify)

EPA's Chemical Data Reporting: 40 CFR 711.15(b)(4)(ii)(A): "using the codes listed in Table 10 of this paragraph, submitters must designate the consumer and commercial product category or categories that best describe the consumer and commercial products in which each reportable chemical substance is used…"

Table 10. Codes for Reporting Consumer and Commercial Product Categories

Code – Category
Chemical Substances in Furnishing, Cleaning, Treatment Care Products
C101 Floor coverings
C102 Foam seating and bedding products
C103 Furniture and furnishings not covered elsewhere
C104 Fabric, textile, and leather products not covered elsewhere
C105 Cleaning and furnishing care products
C106 Laundry and dishwashing products
C107 Water treatment products
C108 Personal care products
C109 Air care products
C110 Apparel and footwear care products Chemical
Substances in Construction, Paint, Electrical, and Metal Products
C201 Adhesives and sealants
C202 Paints and coatings
C203 Building/construction materials - wood and engineered wood products
C204 Building/construction materials not covered elsewhere
C205 Electrical and electronic products
C206 Metal products not covered elsewhere
C207 Batteries
Chemical Substances in Packaging, Paper, Plastic, Hobby Products
C301 Food packaging
C302 Paper products
C303 Plastic and rubber products not covered elsewhere
C304 Toys, playground, and sporting equipment
C305 Arts, crafts, and hobby materials
C306 Ink, toner, and colorant products
C307 Photographic supplies, film, and photochemicals
Chemical Substances in Automotive, Fuel, Agriculture, Outdoor Use Products
C401 Automotive care products Code Description
C402 Lubricants and greases
C403 Anti-freeze and de-icing products
C404 Fuels and related products
C405 Explosive materials
C406 Agricultural products (non-pesticidal)
C407 Lawn and garden care products
Chemical Substances in Products not Described by Other Codes
C980 Non-TSCA use
C909 Other (specify)

OECD – Internationally Harmonized Functional and Product Categories:

Functional Use Categories

Category
Abrasive
Absorbent
Adhesion/cohesion promoter
Adsorbent
Aerating and deaerating agents
Alloying element
Anti-adhesive/cohesive
Anti-caking agent
Anti-condensation agent
Anti-freeze agent
Antioxidant
Anti-redeposition agent
Anti-scaling agent
Anti-slip agent
Anti-stain agent
Anti-static agent
Anti-streaking agent
Binder
Biocide
Bleaching agent
Brightener
Catalyst
Chain transfer agent
Chelating agent
Chemical reaction regulator
Cleaning agent
Cloud-point depressant
Coalescing agent
Conductive agent
Corrosion inhibitor
Crystal growth modifiers (nucleating agents)
Deflocculant
Defoamer
Dehydrating agent (desiccant)
Demulsifier
Density modifier
Deodorizer
Diluent
Dispersing agent
Drier
Dust Suppressant
Dusting agent
Dye
Elasticizer
Embalming agent
Emulsifier
Energy releasers (explosives, motive propellant)
Etching agent
Explosion inhibitor
Filler
Film former
Fire extinguishing agent
Fixing agent (mordant)
Flame retardant
Flavouring and nutrient
Flocculating agent
Flotation agent
Flow promoter
Flux agent
Foamant
Fragrance
Freeze-thaw additive

Category
Fuel
Fuel agents
Hardener
Heat stabilizer
Heat transferring agent
Humectant
Hydraulic fluids
Impregnation agent
Incandescent agent
Insulators
Intermediate
Ion exchange agent
Leaching agent
Lubricating agent
Magnetic element
Monomers
Opacifer
Oxidizing agent
pH regulating agent
Photosensitive agent
Photosensitizers
Pigment
Plasticizer
Plating agent
Polymerization promoter
Preservative
Processing aids not otherwise specified
Propellants, non-motive (blowing agents)
Reducing agent
Refrigerants
Sealant (barrier)
Semiconductor and photovoltaic agent
Sizing agent
Softener and conditioner
Soil amendments
Solids separation (precipitating) agent, not otherwise specified
Solubility enhancer
Solvent
Stabilizing agent
Surface modifier
Surfactant (surface active agent)
Swelling agent
Tanning agents not otherwise specified
Terminator/blocker
Thickening agent
Tracer
UV stabilizer
Vapor pressure modifiers
Viscosity modifier
Waterproofing agent
Wetting agent (non-aqueous)
Wrinkle resisting agent
X-Ray absorber
No specific technical function
Other

Product Use Categories

Categories
Adhesives and sealants
Agricultural products
Air care products
Arts, crafts and hobby materials
Anti-freeze and de-icing products
Apparel and footwear care products
Fabric, textile and leather products not covered elsewhere
Automotive care products
Cleaning and furnishing care products
Explosive materials
Fuels and related products
Ink, toner and colorant products
Laundry and dishwashing products
Lubricants and greases
Other Use
Personal care products
Paints and coatings
Photographic supplies, film and photochemicals
Water treatment products

5) Functional Category Approach, based on Chemical Structure and Function

This section describes how a functional category approach could be used and also discusses potential benefits for EPA and stakeholders. This functional category approach could group chemicals based on a chemical's structure and physicochemical properties to achieve a particular function at the chemical level.

Goal for an approach oriented to functional use:

Identify clusters of structurally and functionally related chemicals that might ultimately provide a spectrum of functional options, across a range of toxicity, that would help both in identifying potential high priority chemicals and also in broadening the pool of chemicals considered to include potential alternatives that are low priority (and high-performing). This goal could be achieved by developing hybrid approaches for grouping chemicals based on structural-level similarities (e.g., TSCA New Chemicals Program's chemical categories) and functional use (e.g., in Safer Choice SCIL). Using chemicals from SCIL with their uses in consumer and institutional products and known toxicity profiles would help build-out the listing of low-priority candidate chemicals.

Overlap of New Chemicals Program and SCIL:

The TSCA New Chemicals Program (NCP) groups chemicals with shared chemical and toxicological properties into categories as it makes determinations for new chemical Pre-Manufacture Notices (PMNs). These NCP groups could be integrated with SCIL functional-use categories. Clusters could be limited to a particular structural class, or used to broaden to structural classes that might comprise a given functional use. A good starting point might be to use EPA's Office of Research and Development (ORD)'s quantitative structure-use relationship models (QSUR) and their analyses of chemicals by functional use[45].

Examples of Functional Use Categories:

Phillips et al. (2017)[46] collected publicly available information on the function of chemicals in consumer products or industrial processes to create the functional use database (FUse)[47]. This information was used to construct 41 validated quantitative structure-use relationship (QSUR)

[45] Phillips, K. A., Wambaugh, J. F., Grulke, C. M., Dionisio, K. L., & Isaacs, K. K. 2017. High-throughput screening of chemicals as functional substitutes using structure-based classification models. *Green Chemistry*. 19(4): 1063-1074.
[46] The Phillips et al. (2017) paper is available online at http://pubs.rsc.org/-/content/articlelanding/2016/gc/c6gc02744j#!divAbstract
[47] Isaacs, K.K., Goldsmith, M.R., Egeghy, P., Phillips, K., Brooks, R., Hong, T., & J.F. Wambaugh. 2016. Characterization and prediction of chemical functions and weight fractions in consumer products. *Toxicology Reports.* 3: 723-732. http://www.sciencedirect.com/science/article/pii/S2214750016300671

models. Using these models, the likelihood of a chemical being able to serve a functional use can be predicted using only the structure and physicochemical properties of a chemical. Below in Table 1 are some examples of the functional uses identified in FUse. This list is not exhaustive and could be further refined through EPA review and based on stakeholder comment. It currently represents a subset of preliminary clusters of TSCA chemicals.

Table 1[48]. The numbers in parentheses represent the number of TSCA chemicals identified as containing a functional use in the given category by the FUse database (first number), followed by the number of SCIL chemicals that were identified by the FUse database to have that functional use (second number).The FUse database predicts that many chemicals on SCIL fall into functional categories not currently identified on SCIL, indicating that SCIL may have value as a potential source of low hazard chemicals in functional classes beyond those categories currently listed. These categories are in non-bolded, non-underlined text.

- Adhesion promoter (129, 5)
- **Antioxidant (221, 46)**
- Antistatic agent (409, 10)
- ***Catalyst (171, 4)***
- **Chelator (167, 60)**
- **Colorant (657, 98)**
- Crosslinker (491, 15)
- **Emollient (467, 72)**
- ***Emulsifier (495, 110)***
- Emulsion stabilizer (154, 54)
- ***Film forming agent (290, 46)***
- **Fragrance (2707, 311)**
- ***Humectant (130, 46)***
- ***Lubricating agent (88, 30)***
- **Oxidizer (38, 11)**
- ***pH stabilizer (106, 30)***
- **Preservative (181, 62)**
- ***Rheology modifier (87, 27)***
- **Skin conditioner (848, 103)**
- **Solvent (372, 131)**
- **Surfactant (855, 384)**
- UV absorber (133, 6)
- ***Viscosity control agent (561, 70)***
- ***Whitener (14,1)***

Refinements

The above list might be further refined. Refinements could include:

1) Re-categorize existing chemicals on the TSCA inventory to seek harmonization with OECD industrial use category definitions, which may vary from chemical function in some cases. This would involve re-building the QSUR models based on the OECD functional use harmonization; and/or
2) Within a functional use category, perform clustering based on chemical structures to refine sub-categories of a functional use (e.g., polar vs. non-polar). This would involve

[48] **Bold, underlined** text represents current SCIL functional use categories. ***Bold, italicized*** text represents functional use classes that the Safer Choice program may identify as SCIL functional use sub-categories in the future.

re-building the QSUR models within the functional categories to strengthen the predictive power of the models.

Functional Uses as an Organizing Tool
Step 1: Selection of a functional use category

Functional use categories could be selected based on any of the following approaches:

1) ORD's high-throughput prioritization predictions;
2) SCIL profiles for potential low priority substance candidates; or
3) Potential for exposure, which could be determined by reported 2016 CDR Aggregate Production Volume[49].

Step 2: For a given functional use category, the chemicals included could be refined and filtered by the amount of data available for a chemical's hazard and exposure potential. The criteria used to determine the hazard and exposure potential could take advantage of the TSCA Work Plan Methodology or ORD's high-throughput prioritization prediction models. Once the criteria have been applied (necessary for #1 and #3), the list of chemicals identified could be compared to the SCIL to ensure the process is not skewed towards identifying only high-priority chemicals.

Example:
We propose to develop QSUR models to be able to predict the alternatives that may fall within a functional category. To illustrate, suppose the FUse database identified a list of approximately 1000 chemicals on the TSCA inventory for a functional use, call it "Functional Category A." To further refine the chemicals included in this category, chemical structure clustering, through development of a QSUR model, could be performed on Functional Category A subcategories (e.g. polar and non-polar chemicals within Functional Category A) as discussed in the refinements at Step 2. The list could then be organized by hazard and/or exposure data availability using either ORD's high-throughput prioritization model or the TSCA Work Plan Methodology to identify chemicals with the most hazard and exposure data available. Chemicals with the most data could progress through the process. Suppose the QSUR model identified 100 SCIL chemicals as Functional Category A; then it is likely that low hazard alternatives may exist in this space. The list of potential alternative chemicals in this space could be expanded using updated QSUR modeling techniques (as described in #1 and #2 refinement options). This method shows promise, but to be successful, it would have to be complemented with stakeholder input to verify that identified potential alternatives have practical value.

Benefits of a functional use approach

This approach could provide a resource for chemical manufacturers and product formulators by increasing the likelihood of the availability of alternative chemicals and helping to address uncertainty in the marketplace. Defining functional use categories at the chemical structure-level

[49] Download the public version of the initial 2016 CDR databased into an Access data file at https://www.epa.gov/chemical-data-reporting/2016-chemical-data-reporting-results

could allow for identification of low-hazard chemicals with similar functionality. In many cases, more than one group of chemicals will provide similar functionality, allowing broader opportunity for identification of low-hazard chemicals, with consideration for potentially exposed sensitive subpopulations. This approach could be used alone to explore potential high and low risk chemicals in a given functional category. It could also be valuable as a complement to other approaches that are likely to identify only high risk chemicals.

By predicting functional uses based on chemical structure, the QSUR models developed using this approach could broaden the pool of alternatives for a given functional use. Model predictions for low-hazard alternatives would serve as a starting point. Stakeholder comment could confirm usefulness of the approach and that the chemicals selected through the approach are actually viable alternatives.

Caveats

Functional uses in this approach are based on publicly available information, including EPA's Safer Chemical Ingredients List[50]. Chemicals may have additional functional uses that are not captured in these public sources.

The high-throughput methods in this approach could be a starting point for identifying low-hazard chemicals with structural and functional similarity to higher-hazard chemicals that may be candidates for prioritization. The high-throughput approach described here may not be useful for functional categories with unique chemistries. Such chemistries may be better suited to manual, situation-specific identification of functional replacements.

This functional category approach may be most effective as a complement to other organizing approaches (as described in Step 1).

Stakeholder input would be important to successful implementation.

[50] http://www.epa.gov/saferchoice/safer-ingredients

6) Integration of Traditional and New Approach Methods

Introduction

This section describes a prototype, web-based tool and the integration of new approach methods (NAM) and traditional data for identifying potential candidates for prioritization. The data cover hazard, exposure, persistence, and bioaccumulation. The preliminary application of the tool is illustrated using five distinct methods that range from a method that mirrors the scoring approach outlined in the Toxic Substance Control Act (TSCA) 2012 Work Plan Methods Document, several methods that incorporate NAM into the TSCA 2012 scoring approach in different ways, and a method that uses an estimated margin-of-exposure. The chemicals prioritized in the example methods were from Step 2 in the TSCA 2012 Work Plan process (EPA, 2016) and the Safer Chemicals Ingredients List (SCIL)(EPA, 2017). The TSCA Step 2 chemicals were originally identified by OCSPP using a specific set of data sources meeting one or more factors including chemicals identified as potentially of concern for children's health (e.g., chemicals with reproductive or developmental effects, chemicals identified as persistent, bioaccumulative, and toxic (PBT), chemicals identified as probable or known carcinogens, used in children's products, and used in consumer products). The SCIL chemicals are managed by the EPA's Safer Choice Program to reflect chemical ingredients that meet specific low-concern criteria and are generally acceptable for use in Safer Choice-labeled products. The selection of these two lists is for illustration purposes only and intended to demonstrate applicability of the tool by comparing the relative scoring for a list that was more likely to contain a higher percentage of high priority chemicals (i.e., TSCA Step 2) with a list that was likely to contain a higher percentage of low priority chemicals (i.e., SCIL). It should be noted that the scoring criteria used in the tool differ from the criteria associated with the Safer Choice program and may not agree with scores made public through the SCIL.

The application of the tool illustrates how one could carry out candidate selection in TSCA. The tool is intended to provide greater transparency on the prioritization process and allow users to flexibly explore the relative impact of different approaches. Future plans could include expanding the tool to provide a flexible, user-defined prioritization process that can be used in many contexts and the incorporation of other NAM related to endpoints and measurements of regulatory significance.

Chemical Lists

Input into the candidate selection tool begins with a chemical list. Chemicals in the lists are defined by matching CASRN and chemical name pairs, which are curated and mapped to entries in the DSSTox database. Additional steps are taken to ensure the chemical and corresponding data are correctly paired. For the examples, we are using 3 lists of chemicals defined in Table 1.

Table 1. Chemical lists used in the examples

List	Definition	Number of chemicals
TSCA2	The TSCA Step 2 chemicals	344*
SCIL	Safer Chemicals Ingredients List	867
TSCA2 / SCIL	Unique set of chemicals from the merged TSCA Step 2 and SCIL lists	1184
*The TSCA Step 2 chemicals number 344 instead of the original 345 due to consolidation with another category (EPA, 2016).		

New Approach Methods (NAM)

The term new approach methods (NAM) was recently introduced to cover any *in vitro, in silico, or in chemico* technique used to provide data or information for regulatory decision making (ECHA, 2016). The data used by the tool are organized into domains, subdomains, and individual components. The domains include hazard, exposure, and persistence/bioaccumulation. The subdomains include human and ecological. The NAM data under this broad definition fall into all of the various domains (Table 2).

Table 2: List of NAM data currently used in the tool

Domain	Subdomain	Brief Description	Reference
Hazard	Human	Computational model predicting quantitative *in vivo* effect levels for rodent repeat dose toxicity studies based on chemical structure, high-throughput *in vitro* assay data from ToxCast, and high-throughput toxicokinetic data.	(Truong et al., 2017)
Hazard	Human	Potency value from the most sensitive, high quality ToxCast data is converted into a quantitative *in vivo* effect level using high-throughput toxicokinetic (HTTK) model. The value represents a conservative estimate of an *in vivo* effect level across a broad range of study types.	(Wetmore et al., 2013), (HC, 2016)
Hazard	Human	Computational model to predict estrogen receptor (ER) agonist activity based on ToxCast assays.	(Judson et al., 2015)
Hazard	Human	Computational model to predict ER antagonist activity based on ToxCast assays.	(Judson et al., 2015)
Hazard	Human	Computational model to predict androgen receptor (AR) agonist activity based on ToxCast assays.	(Kleinstreuer et al., 2016)
Hazard	Human	Computational model to predict AR antagonist activity based on ToxCast assays.	(Kleinstreuer et al., 2016)
Hazard	Human	Consensus QSAR model to predict ER agonist activity.	(Mansouri et al., 2016)
Hazard	Human	Consensus QSAR model to predict ER antagonist activity.	(Mansouri et al., 2016)

Hazard	Human	Generic read across (GenRA) approach that uses chemical structure and *in vitro* bioactivity to identify chemical analogs to determine/ predict qualitative *in vivo* responses.	(Shah, Liu, Judson, Thomas, & Patlewicz, 2016)
Exposure	Human	Machine learning models to predict probability of human exposure for one of four exposure pathways: 1) far field pesticide use; 2) non-pesticide dietary exposure; 3) far field industrial exposure; and 4) near field exposure.	(Ring et al., In preparation)
Exposure	Human	The Systematic Empirical Evaluation of Models (SEEM) produces a calibrated consensus prediction of human exposure using minimal input data.	(Wambaugh et al., 2014)
Exposure	Human	The High-Throughput Stochastic Human Exposure and Dose Simulation Model (SHEDS-HT) is a Monte Carol based population model of human exposure to chemicals in consumer products.	(Isaacs et al., 2014)
Persistence/ Bioaccumulation	NA	QSAR model to predict the potential half-life in air, water, soil, and sediment.	(Zang et al., 2017)
Persistence/ Bioaccumulation	NA	QSAR model to predict the bioaccumulation factor (BAF), a variable that represents the potential for a chemical to bioaccumulate in organisms in the environment.	(Zang et al., 2017)

Candidate Scoring Methods

Method 1 – TSCA 2012

Generally follows the TSCA 2012 Work Plan method (EPA, 2012), with no NAM data incorporated into the human hazard domain.

- Hazard Domain:
  - Select highest scoring component in the human health OR ecological subdomain: 1, 2, or 3 (Table 3)
- Exposure Domain
  - Select highest scoring component in the human OR ecological subdomain: 1, 2, or 3 (Table 3)
- Persistence/Bioaccumulation Domain:
  - Select highest scoring component: 1, 2, or 3 (Table 3)
- Calculate Total Score: Hazard Domain Score + Exposure Domain Score + Persistence/Bioaccumulation Domain Score
- Assign to High, Medium, or Low bin
  - High: 7-9
  - Medium: 5-6
  - Low: 3-4

Advantages of Method 1

- Established prioritization workflow
- Integrates hazard, exposure, persistence, and bioaccumulation

Disadvantages of Method 1

- Limited human and ecological hazard data for many chemicals
- Selects highest score in each domain which may inflate overall priority scores
- Does not penalize for lack of data

Method 2 – TSCA 2012 + NAM Equal Weight

Follows the TSCA 2012 Work Plan method except NAM were given equal weight in all domains.

- Hazard Domain:
  - Select highest scoring component in the human health OR ecological subdomain: 1, 2, or 3 (Table 3)
  - NAM components given equal weight as traditional *in vivo* toxicity studies
- Exposure Domain
  - Select highest scoring component in the human OR ecological subdomain: 1, 2, or 3 (Table 3)
- Persistence/Bioaccumulation Domain:
  - Select highest scoring component: 1, 2, or 3 (Table 3)
- Calculate Total Score: Hazard Domain Score + Exposure Domain Score + Persistence/Bioaccumulation Domain Score
- Assign to High, Medium, or Low bin
  - High: 7-9
  - Medium: 5-6
  - Low: 3-4

Advantages of Method 2

- Builds upon established prioritization workflow
- Integrates hazard, exposure, persistence, and bioaccumulation
- Expands chemicals with available data through application of NAM

Disadvantages of Method 2

- Selects highest score in each domain which may inflate overall priority scores
- Some NAMs are inherently conservative and using equal weighting may further bias priority scores
- Does not penalize for lack of data

Method 3 – TSCA 2102 + NAM Deferential

Follows the TSCA 2012 Work Plan method except for human hazard NAM is incorporated only in the absence of traditional *in vivo* studies. In other domains, NAM is given equal weight.

- Hazard Domain:
  - Select highest scoring component in the human health OR ecological subdomain: 1, 2, or 3 (Table 3)
  - NAM components are deferential to *in vivo* component if available (Table 3)
- Exposure Domain

    - Select highest scoring component in the human OR ecological subdomain: 1, 2, or 3 (Table 3)
- Persistence/Bioaccumulation Domain:
    - Select highest scoring component: 1, 2, or 3 (Table 3)
- Calculate Total Score: Hazard Domain Score + Exposure Domain Score + Persistence/Bioaccumulation Domain Score
- Assign to High, Medium, or Low bin
    - High: 7-9
    - Medium: 5-6
    - Low: 3-4

Advantages of Method 3

- Builds upon established prioritization workflow
- Integrates hazard, exposure, persistence, and bioaccumulation
- Expands chemicals with available data through application of NAM

Disadvantages of Method 3

- Selects highest score in each domain which may inflate overall priority scores
- Does not penalize for lack of data

Method 4 – Sum of Scores

A variation of the TSCA 2012 Work Plan method except instead of deriving a single value for hazard, exposure, and persistence/bioaccumulation, all of the individual component scores across the domains are summed. NAM data are incorporated as a component within the domains and given equal weight.

- Hazard Domain:
    - Sum all components in the human health and ecological subdomains (Table 3)
- Exposure Domain
    - Sum all components in the human and ecological subdomains (Table 3)
- Persistence/Bioaccumulation Domain:
    - Sum both components (Table 3)
- Calculate Total Score: Hazard Domain Sum Score + Exposure Domain Sum Score + Persistence/Bioaccumulation Domain Sum Score
- Assign to High, Medium, or Low bin
    - High: >30
    - Medium: 10-30
    - Low: ≤10

Advantages of Method 4

- Integrates hazard, exposure, persistence, and bioaccumulation
- Expands chemicals with available data through application of NAM

Disadvantages of Method 4

- Utilizes new prioritization methodology

- Does not penalize for lack of data

Method 5 – Hazard/Bioactivity Exposure Ratio (H/BER)

Prioritization is based on a binned estimate of the hazard/bioactivity-to-exposure ratio (oral only). NAM is incorporated for human hazard with equal weight to traditional *in vivo* studies. This is similar to the proposed approach for the Health Canada Chemicals Management Plan (CMP) (HC, 2016).

- Ratio of the minimum effect level from *in vivo* toxicity studies (all study types) and the quantitative human hazard NAM data divided by the maximum oral exposure.
- Assign to High, Medium, or Low bin
    - High: $\leq 10^4$
    - Medium: $10^4 - 10^6$
    - Low: $\geq 10^6$

Advantages of Method 5

- Risk-based
- Expands chemicals with available data through application of NAM

Disadvantages of Method 5

- Integrates only human hazard and exposure
- Currently only considers oral route (may add inhalation/dermal route in future)
- Does not penalize for lack of data

We illustrate application of the tool and the integration of NAM using 5 example methodologies. Each of the methods uses data compiled from a broad range of sources (Table 3). The data are organized by domain and include human hazard, ecological hazard, human exposure, ecological exposure, and persistence/bioaccumulation. Chemicals with data missing in each domain are flagged for potential follow-up and may provide an initial indication for readiness for risk evaluation. Each domain consists of multiple components (i.e., data sources). The different domains, components, and their integration in scoring functions are described in detail in Appendix I. Other methods may be added in the future as more analyses are completed and additional data becomes available.

Table 3: Mapping of domain and component scores to prioritization methods

Domain	Subdomain	Component Score	Method 1 (TSCA 2012) Scores	Method 2 (NAM Equal Weight) Scores	Method 3 (NAM Deferential) Scores	Method 4 (Sum of Scores) Scores	Method 5 (H/BER) Score
Hazard	Human	acute.invivo	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Human	subchronic.invivo	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Human	chronic.invivo	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Human	reprotox.invivo	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Human	devtox.invivo	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Human	neurotox.invivo	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Human	immunotoxicity.invivo	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Human	cancer.invivo	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Human	model.systemic.mm	NA	1,2, or 3	1,2, or 3 only if no chronic.invivo or subchronic.invivo data are available	1,2, or 3	NA
Hazard	Human	model.systemic.ivive	NA	1,2, or 3	1,2, or 3 only if no chronic.invivo or subchronic.invivo data are available	1,2, or 3	NA
Hazard	Human	genetox	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Human	er.invitro	NA	1,2, or 3	1,2, or 3 only if no devtox.invivo or reprotox.invivo data are available	1,2, or 3	NA
Hazard	Human	erant.invitro	NA	1,2, or 3	1,2, or 3 only if no devtox.invivo or reprotox.invivo data are available	1,2, or 3	NA
Hazard	Human	ar.invitro	NA	1,2, or 3	1,2, or 3 only if no devtox.invivo or reprotox.invivo data are available	1,2, or 3	NA

Hazard	Human	arant.invitro	NA	1,2, or 3	1,2, or 3 only if no devtox.invivo or reprotox.invivo data are available	1,2, or 3	NA
Hazard	Human	er.qsar	NA	1,2, or 3	1,2, or 3 only if no devtox.invivo or reprotox.invivo data are available	1,2, or 3	NA
Hazard	Human	erant.qsar	NA	1,2, or 3	1,2, or 3 only if no devtox.invivo or reprotox.invivo data are available	1,2, or 3	NA
Hazard	Human	genra	NA	1,2, or 3	1,2, or 3 only if no chronic.invivo or subchronic.invivo data are available	1,2, or 3	NA
Hazard	Ecological	eco.acute.aquatic	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Ecological	eco.chronic.aquatic	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	cdr.score.pv	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	cdr.score.sites	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	cdr.score.industrial	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	cdr.score.commercial	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	cdr.score.consumer	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	cdr.score.child	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	score.biomonitoring	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	score.residential	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	score.consumer	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	score.diet	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	score.children	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	score.exposure.quantile	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	score.exposure.quantitative	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Human	score.tri	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Exposure	Ecological	score.eco	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA

Persistence/ Bioaccumulation		physchem.persistence	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Persistence/ Bioaccumulation		physchem.bioaccumulation	1,2, or 3	1,2, or 3	1,2, or 3	1,2, or 3	NA
Hazard	Human	minimum oral effect level	NA	NA	NA	NA	Minimum effect level from *in vivo* toxicity studies (all study types), QSAR/QBAR, and *in vitro* bioactivity.
Exposure	Human	maximum quantitative oral exposure	NA	NA	NA	NA	Maximum oral exposure

Preliminary Results

To illustrate application of the tool and NAM data to the process, chemicals from Step 2 in the TSCA 2012 Work Plan and the SCIL were scored using the five prioritization methods. The results for Methods 1 – 4 are summarized in Figures 1 and 2. Across all four Methods, the SCIL chemicals have a lower proportion of High bin chemicals and a larger proportion of Low bin chemicals. For these chemical lists, there are some differences between Methods 1 – 3, with the addition of NAM data pushing more chemicals into the High bin. This is primarily because of filling data gaps in Method 1. Method 2 (NAM and traditional *in vivo* data are treated equally) has the most High bin chemicals, because the NAM data tends to be more conservative (lower points of departure) than the traditional *in vivo* data. Another important finding was that many chemicals have at least some missing data (discussed more below – See Figure 4), although this does not always lead to a low score. Recall that if all data for a domain (i.e., hazard, exposure, persistence/bioaccumulation) is missing, that domain is given a value of 1 in the final score for Methods 1-3. The histograms in Figure 2 show in more detail the actual scores, and one can see the minor changes in scores (especially in SCIL) going from Method 1 to 2 (e.g., in the score= 5,6,7 bars). The histogram plots also show the trend to lower scores for SCIL as compared with TSCA Step 2.

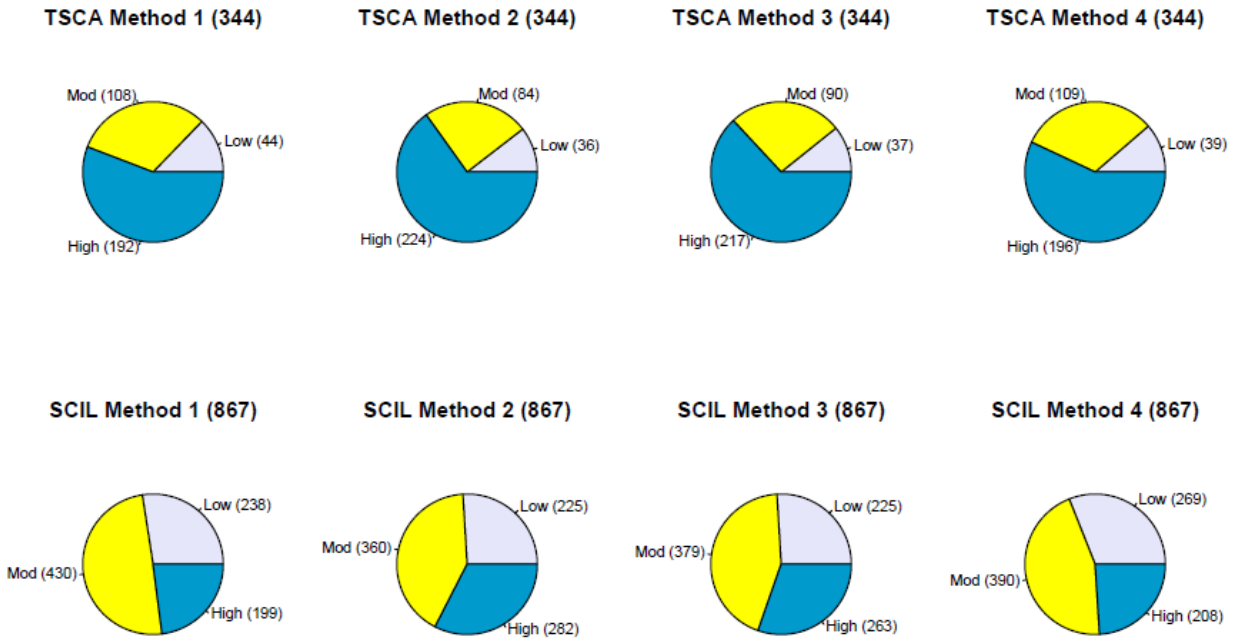


Figure 1: Bins of the TSCA Step 2 and SCIL chemical using Methods 1 – 4. Numbers of chemicals in the database in each bin are given in parentheses.

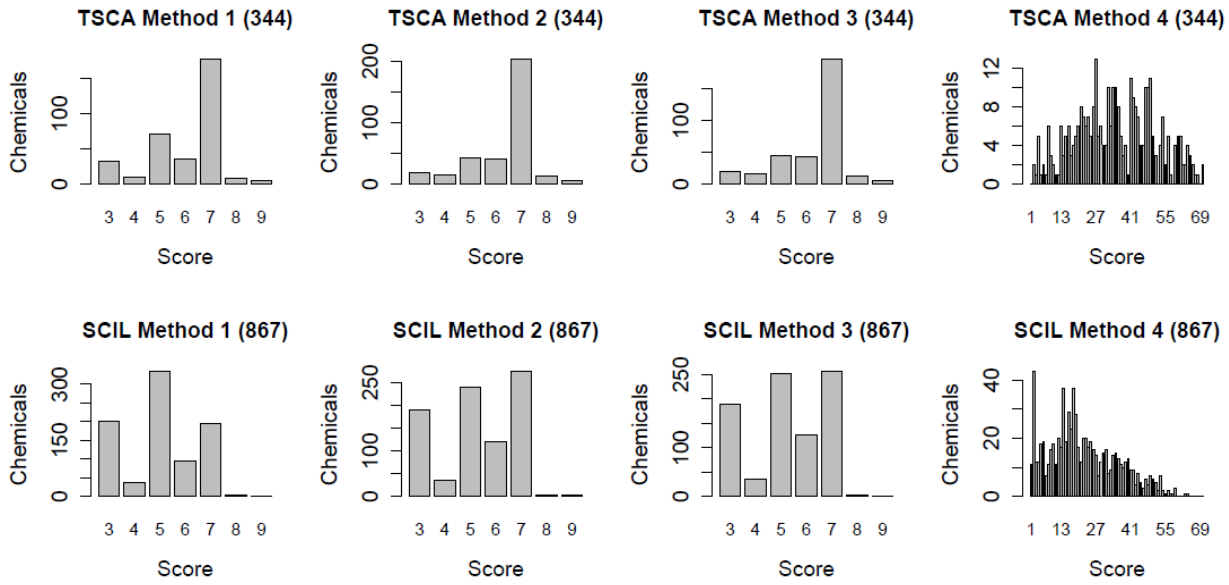


Figure 2: Histograms of the TSCA Step 2 and SCIL chemical using Methods 1 – 4. The total number of chemicals in the chemical set are given in parentheses in the title.

Method 5 is based on the Hazard/Bioactivity Exposure Ratio (H/BER), which is the ratio of the minimum effect level (*in vivo* or NAM) to the maximum estimated human exposure. Figure 3 shows the distributions of H/BER values for the two chemical lists. There are some chemicals with an expected H/BER<1, but the majority of chemicals are above 10,000.

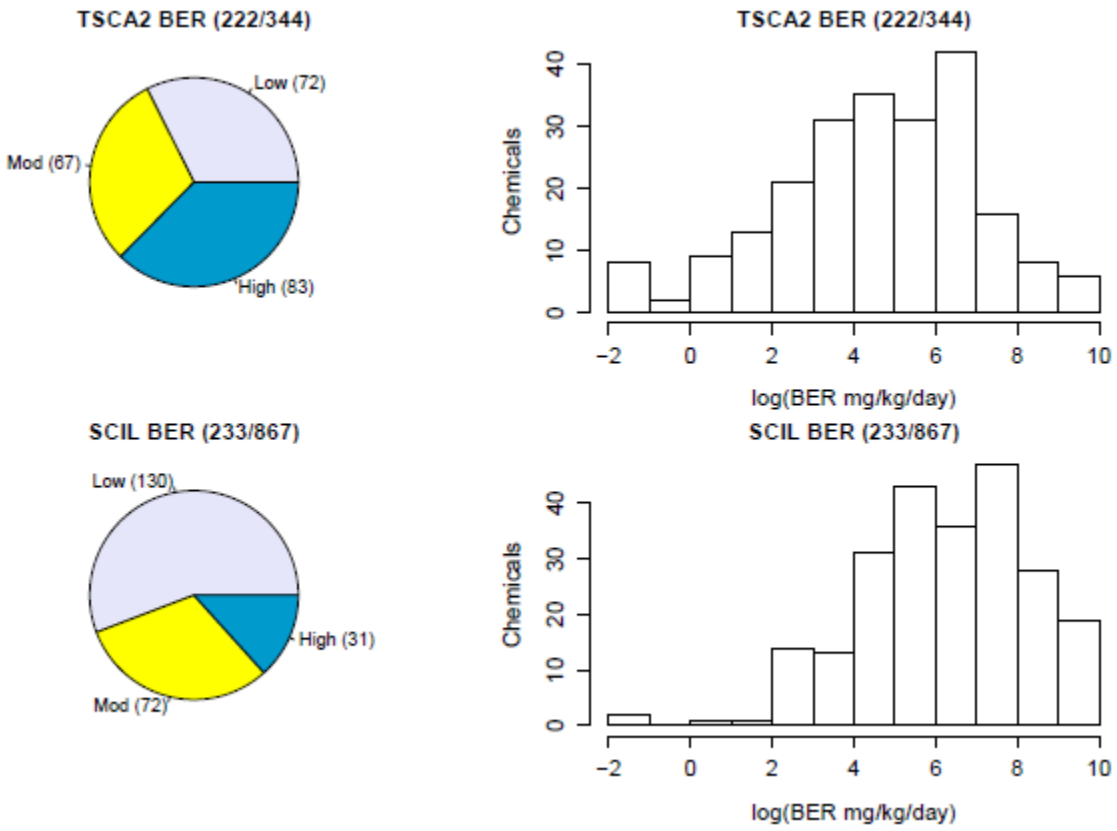


Figure 3: Bins of $\log_{10}$(H/BER) values for the TSCA Step 2 and SCIL chemicals in Method 5. As described in the Candidate Score Methods Section, chemicals with an H/BER ratio of $\leq 10^4$ ($\log_{10} \leq 4$) are binned as High, chemicals with a H/BER ratio of $10^4 – 10^6$ - ($\log_{10}$ 4 - 6) are binned as Moderate, and chemicals with an H/BER ratio of $\geq 10^6$ ($\log_{10} \geq 6$) are binned as Low. The values in the title are the numbers of chemicals with a H/BER calculated over the number in the chemical set (e.g. 222 of the 344 TSCA2 chemicals had a H/BER calculated, while the remaining 122 did not have the required information for either hazard or exposure). Distributions of the $\log_{10}$(H/BER) values are provided on the right hand side.

Another key issue is missing data. Figures 4 and 5 are heat maps of the maximum score across all of the domain and subdomains when both the traditional and NAM data are included. In the current implementation, missing data in each domain and subdomain are flagged for potential follow-up. A white cell indicates that no data are available from any source. It is clear that ecological data are sparse (hazard and exposure) for both of the TSCA Step 2 and SCIL lists, although more so for the SCIL list. One possible solution is to use more than just acute and chronic aquatic toxicity data in the ECOTOX database to augment the ecological hazard data. Recall that in the final scoring, if there was not any data for a given domain/subdomain pair, the score for that domain/subdomain was set to a value of 1. The heat maps include the NAM as well as traditional human hazard data.

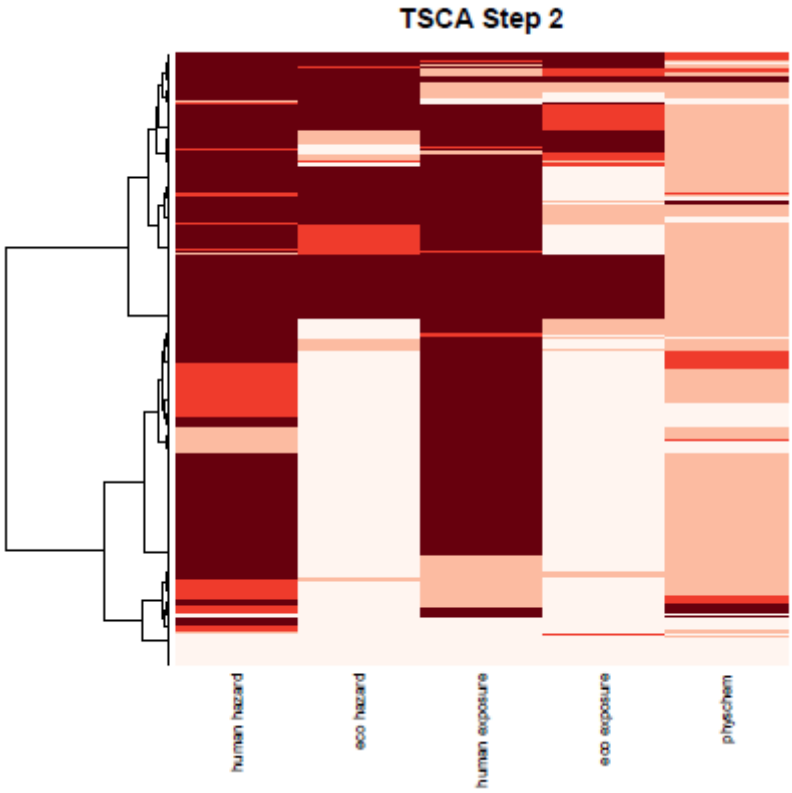


Figure 4: Heatmap of maximum scores for each domain and subdomain across the TSCA Step 2 chemicals. The y-axis are individual chemicals in the TSCA Step 2 list. The x-axis are the different domains and subdomains used in the scoring process for Methods 1 – 4. Dark red indicates a high score (i.e., 3) in that domain when both traditional and NAM data are included. Light red and pink indicate medium (2) and low (1) scores, respectively. White indicates no data for that chemical in that domain.

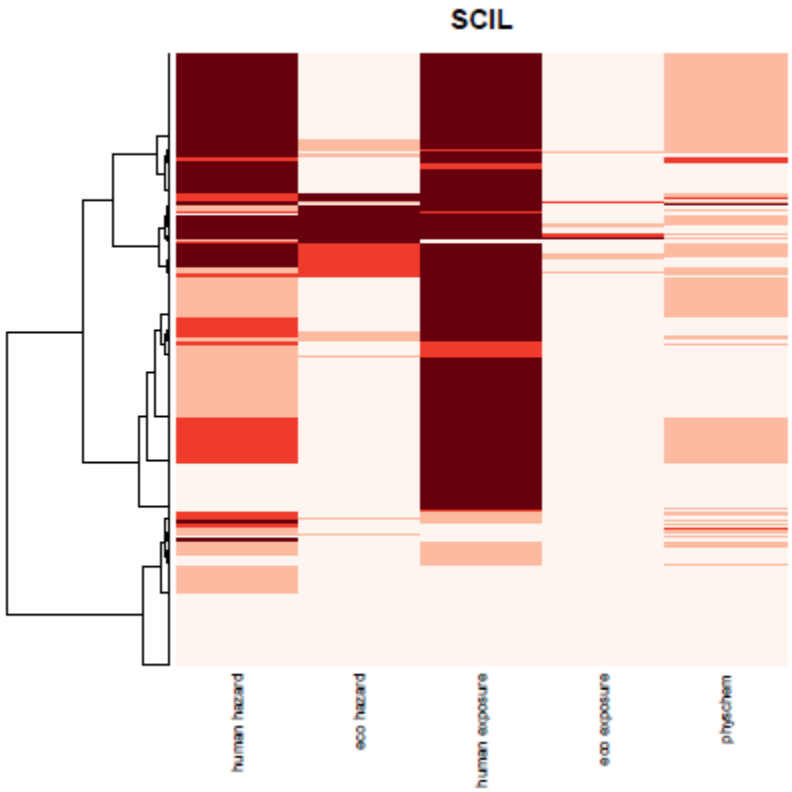


Figure 5: Heatmap of maximum scores for each domain and subdomain across the SCIL chemicals. The y-axis are individual chemicals in the SCIL list. The x-axis are the different domains and subdomains used in the scoring process for Methods 1 – 4. Dark red indicates a high score (i.e., 3) in that domain when both traditional and NAM data are included. Light red and pink indicate medium (2) and low (1) scores, respectively. White indicates no data for that chemical in that domain.

A final issue is the amount of information that the NAM adds. For human hazard, the NAM data changed 74 out of 344 (22%) of the TSCA Step 2 chemicals, either by adding data when there was none available from traditional methods, or by changing the overall bin (Low, Moderate, High). The score could increase if the predicted effect levels were more potent than the *in vivo* toxicity studies (Method 2) or if either the endocrine or GenRA data increased the human hazard score (Method 2 or 3). For the SCIL list, 281 out of 867 (32%) chemicals, the NAM data filled missing human hazard data or caused the overall bin to change.

Benefits

A prototype tool and a range of preliminary Methods were illustrated. The prototype tool, Methods, and results are a work-in-progress and results should be considered as examples. The benefits of the tool and the incorporation of NAM data include:

- Transparency and reproducibility for the candidate selection process

- Systematic examination of domains and components that contribute most to candidate selection
- Utilizes large collections of existing traditional and NAM data for hazard, exposure, persistence, and bioaccumulation
- Accommodates new methods and data when available
- Incorporates cost-effective NAM for collecting data on thousands of chemicals to fill gaps in traditional data
- Enables focused data requests to stakeholders

Caveats

The caveats of the current tool and integration of the NAM data include:

- Additional data cleaning and curation: Some of the data sets used in the tool require further cleaning and curation. This is an ongoing effort. For example, there are acute studies in the database that are not yet included in this analysis.
- Expansion of ecological hazard endpoints: The current endpoints used to define the ecological hazard scores were limited to acute and chronic aquatic toxicity studies resulting in many chemicals with missing data. An expansion of the species used to assess ecological toxicity and the development of corresponding scoring cutoffs may help fill in some data gaps.
- Expansion of quantitative exposure estimates: The current quantitative exposure estimates are limited to the oral route. The addition of high-throughput exposure estimates for the inhalation route is a work in progress.
- Lack of respiratory sensitizers: Respiratory sensitization was one of the endpoints used in the TSCA 2012 Work Plan Methods, but this endpoint has not been incorporated into the current tool or database. Future efforts may add this information.
- Experimentally measured persistence and bioaccumulation data: The current data on the persistence and bioaccumulation potential for the chemicals was calculated using OPERA QSAR models. Future efforts will integrate QSAR model and/or experimentally measured data from EPI Suite.

Future Directions

The prototype, web-based tool outlined in this section enables the transparent and systematic selection of candidate chemicals using a broad range of experimental and computational data related to hazard, exposure, persistence, and bioaccumulation. The data are derived from both traditional and NAM sources with a focus on endpoints and measurements of regulatory significance. In the current implementation, candidate identification is illustrated using five different pre-defined methods. In future iterations, the tool can be upgraded to allow flexible selection of individual components and domains, multiple data transformation or binning options, user defined scoring and weighting methods, options to deal with missing data, and flexible integration of NAM. The long-term goal is to allow customization of workflows that accommodate a range of decision contexts.

Apart from increased flexibility, future development of the tool may allow systematic evaluation of the relative sensitivity associated with the different components, binning selections, and scoring cutoffs. In the current Methods, it was noted that the hazard component scores were very sensitive to the example cutoffs and number of bins used. Additional capabilities for systematic sensitivity analysis will allow identification of potentially unforeseen areas of sensitivity in the prioritization decisions and the focus of scientific resources to ensure sensitive areas are robustly defined.

Appendix I

Hazard Domain

Data components in the human hazard subdomain consist of quantitative effect levels from traditional animal studies, genotoxicity, cancer indications from animal bioassays, estimates of effect levels using two different NAMs, and qualitative hazard predictions using automated read-across. Data components in the ecological hazard subdomain consist of experimental measurements of acute and chronic aquatic toxicity.

Quantitative Effect Levels from *In Vivo* Toxicity Studies

Quantitative effect levels are obtained from the EPA ToxValDB database, described in Appendix I. The database is composed of effect levels from *in vivo* toxicity studies of the following types: BMD/BMDL, LD50/LC50, NOEL/LOEL, NOEC/LOEC, NOAEL/LOAEL, NOAEC/LOAEC. Values must be in units of mg/kg or mg/kg-day for oral or mg/m3 for inhalation studies.

The following study types were included:

- Acute
- Subchronic (defined for this analysis as ≤90 days)
- Chronic (defined for this analysis as >90 days)
- Reproductive
- Developmental
- Neurotoxicity

For each chemical and study type, we find the lowest oral and lowest inhalation effect levels (if they exist). For Methods 1 – 4, the toxicity values are turned into component scores using the cutoffs given in Tables A1.1 and A1.2. In this example, missing values are assigned a score of 0, which flags them for future data needs.

Table A1.1: Example component scoring cutoffs in mg/kg/day for human hazard by study type (oral exposure), based on the TSCA 2012 Work Plan Methods Document.

Score	Acute	Subchronic	Chronic	Reproductive	Developmental	Neurotoxicity
1	>2000	>300	>100	>250	>250	>300
2	300-2000	30-300	10-100	50-250	50-250	30-300
3	< 300	<30	<10	<50	<50	<30

Based on the component scoring cutoffs in Table A1.1, many tested chemicals are in the High hazard bin for the acute toxicity studies while the chemicals are more equally distributed for the other study types (Fig. A1.1). This demonstrates that the choice of cutoffs in Table A1.1 have a significant effect on the overall human hazard domain score. Cutoffs provided for the different domains are only used as an example to illustrate the approach.

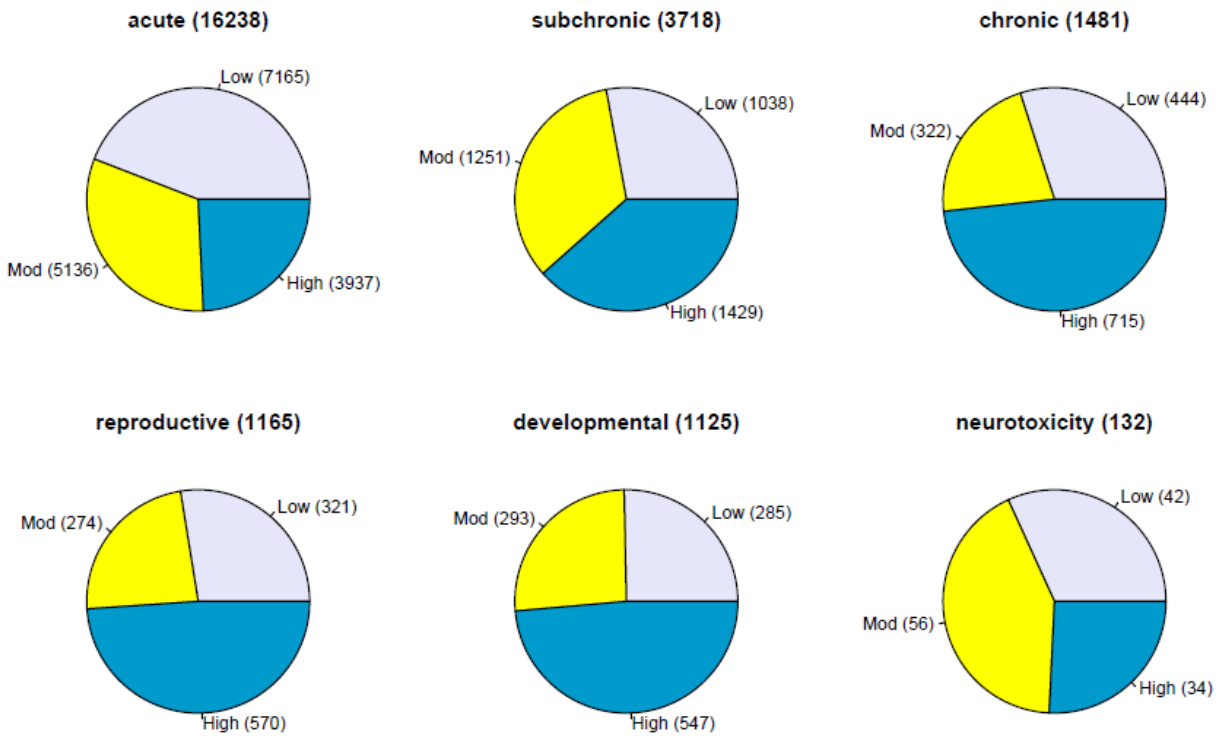


Figure A1.1a: Distribution of component scores based on quantitative effect levels from oral *in vivo* toxicity studies. The values in the title are the total number of chemicals with data in the ToxValDB database for the specified data domain. Data is shown for the full database which includes TSCA Step 2 chemicals, SCIL chemicals, and other chemicals not on these lists.

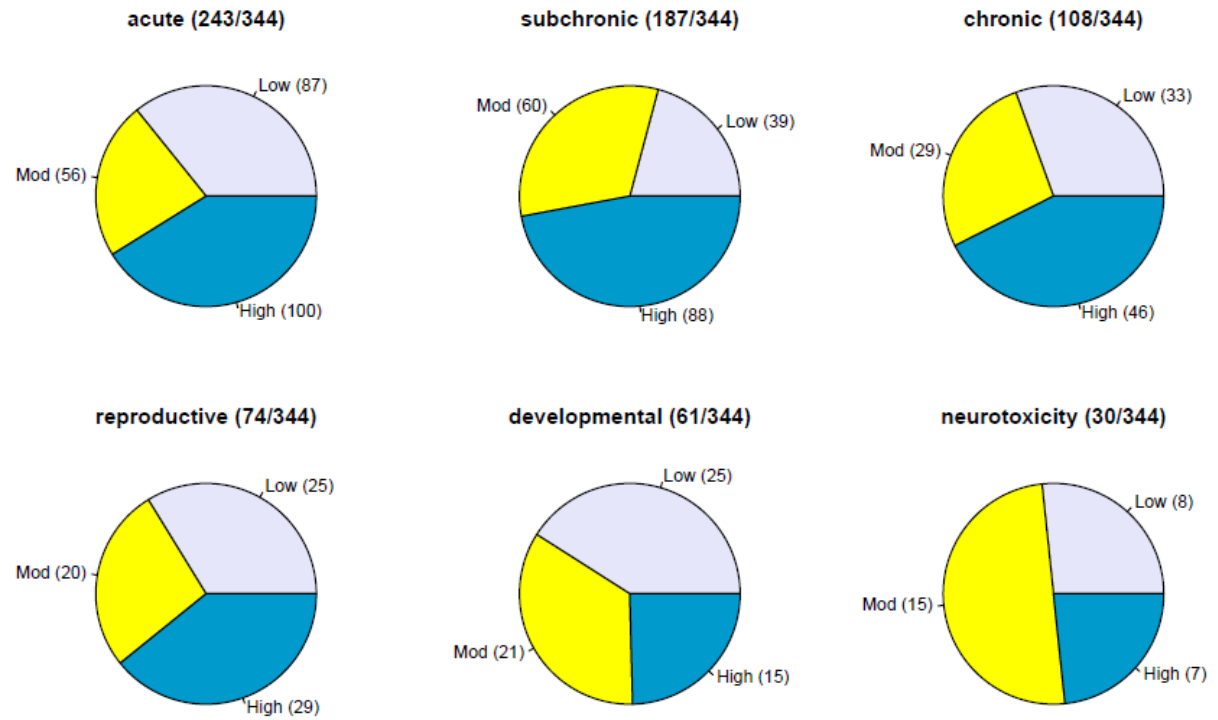


Figure A1.1b: Distribution of component scores based on quantitative effect levels from oral *in vivo* toxicity studies. Data is shown for the TSCA Step 2 chemicals only. The values in parentheses in the pie chart are the number of chemicals in the corresponding bin in the chemical collection and data domain. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2or SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

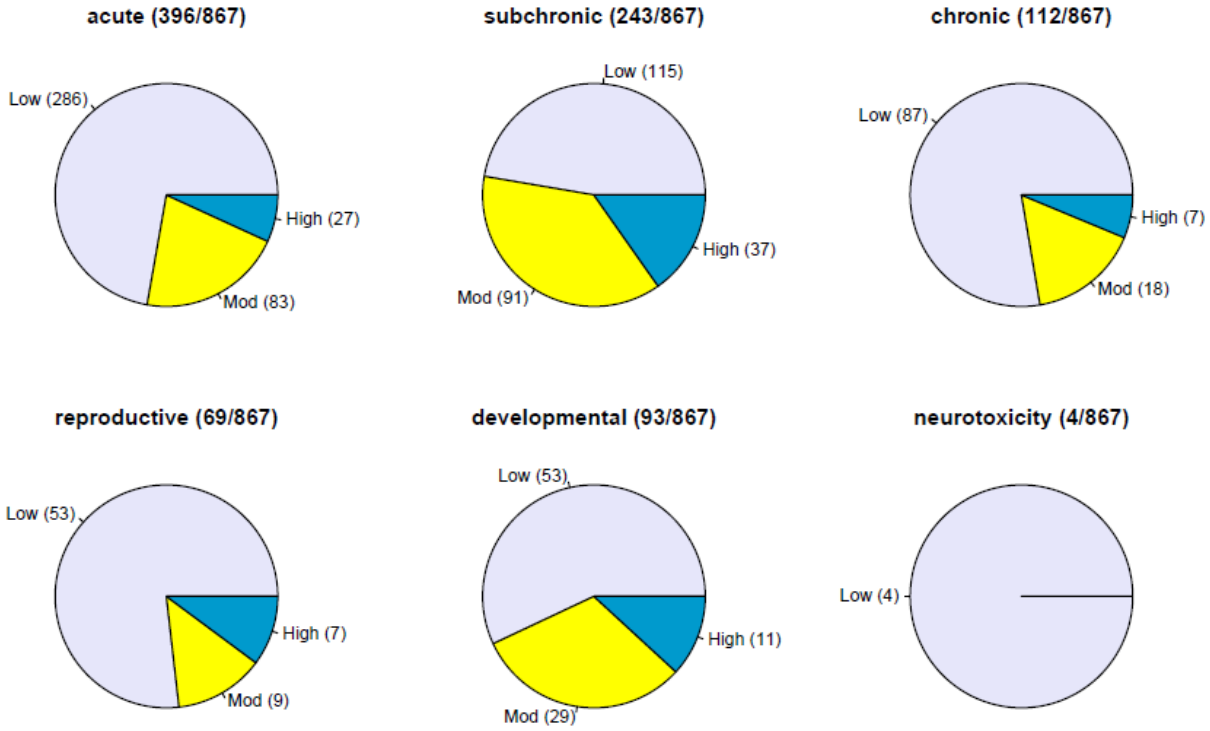


Figure A1.1c: Distribution of component scores based on quantitative effect levels from oral *in vivo* toxicity studies. Data is shown for the SCIL chemicals only. In the title, the first value in parentheses is the number of chemicals in the chemical collection (SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Table A1.2: Example component scoring cutoffs in mg/m3 for human hazard by study type (inhalation exposure), based on the TSCA 2012 Work Plan Methods Document*.

Score	Acute	Subchronic	Chronic	Reproductive	Developmental	Neurotoxicity
1	>20000	>2500	>2500	>2500	>2500	>2500
2	10000-20000	1000-2500	1000-2500	1000-2500	1000-2500	1000-2500
3	< 10000	<1000	<1000	<1000	<1000	<1000

*Based on the TSCA 2012 Work Plan Methods, only acute, reproductive, and developmental study types had defined ranges for inhalation effect levels. Ranges for subchronic, chronic, and neurotoxicity were added for this illustration. In addition, the ranges for acute, reproductive, and developmental study types listed in this table were for gas and vapor exposures. Different ranges were defined for mist and dust exposures. For the purposes, of this illustration, the same ranges were used regardless of whether it was a gas, vapor, mist, or dust.

Based on the component scoring cutoffs in Table A1.2, the distribution of chemicals across the High, Medium, and Low bins is significantly different than that observed for oral studies (Fig. A1.2). A greater proportion of chemicals are in the High hazard bin for acute, chronic, reproductive, and neurotoxicity studies.

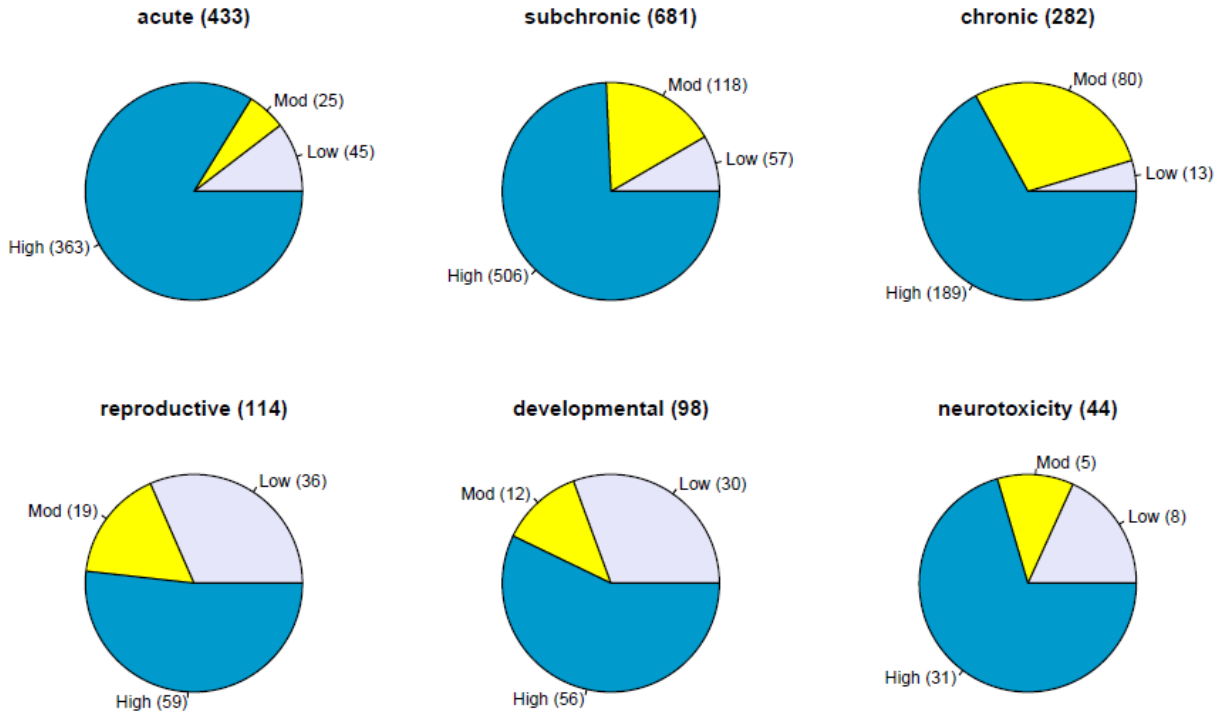


Figure A1.2a: Distribution of component scores based on quantitative effect levels from inhalation *in vivo* toxicity studies. The values in the title are the total number of chemicals with data in the ToxValDB database (which includes TSCA Step 2 chemicals, SCIL chemicals, and other chemicals not on these lists) for the specified study type.

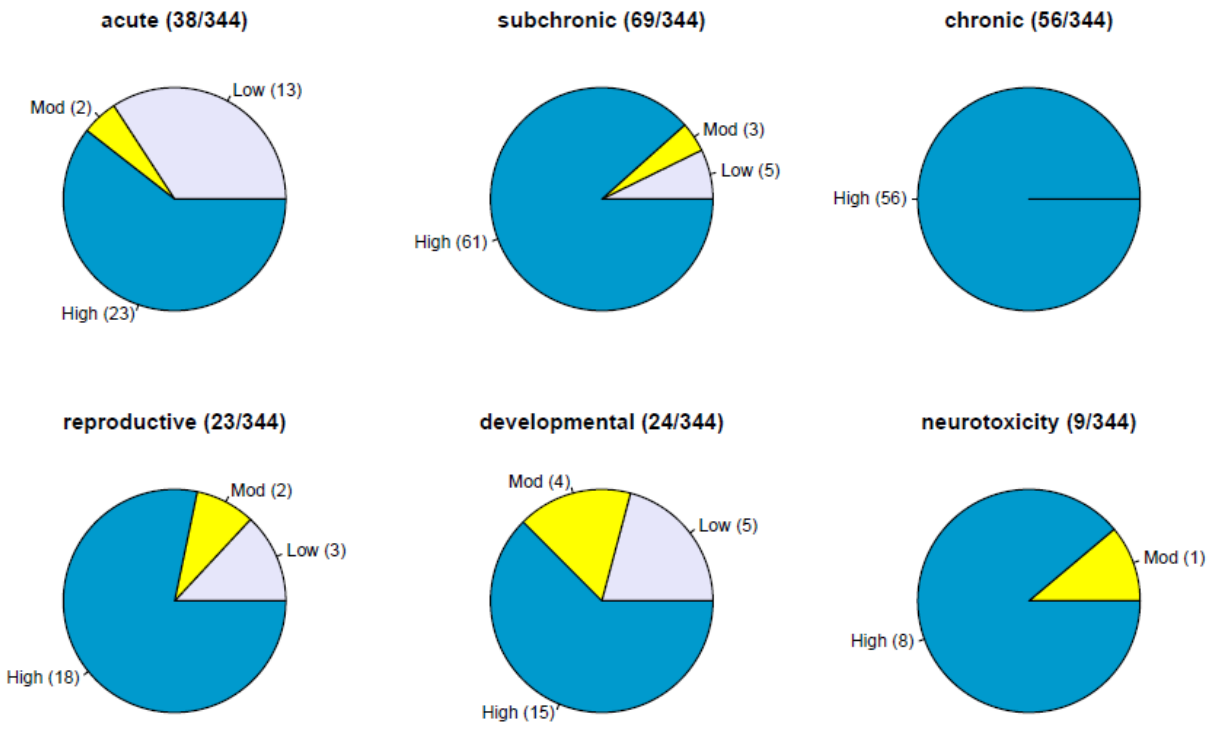


Figure A1.2b: Distribution of component scores based on quantitative effect levels from inhalation *in vivo* toxicity studies. Data is shown for the TSCA Step 2 chemicals. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2) with data in that domain. The second value is the total number of chemicals in the collection.

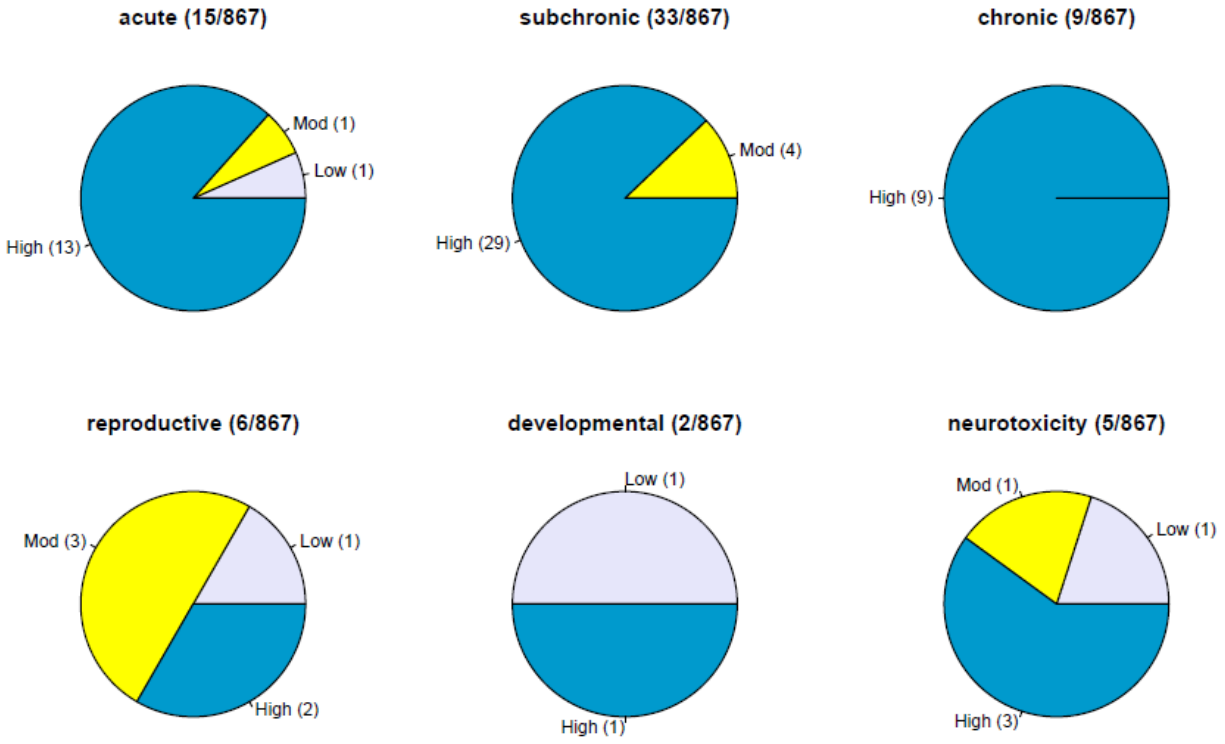


Figure A1.2c: Distribution of component scores based on quantitative effect levels from inhalation *in vivo* toxicity studies. Data is shown for the SCIL chemicals. In the title, the first value in parentheses is the number of chemicals in the chemical collection (SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Genotoxicity

Genotoxicity data are extracted from the EPA ToxValDB database, where it had been compiled from COSMOS (http://www.cosmostox.eu/home/welcome/), ECHA via ChemPortal (https://www.echemportal.org/echemportal/index.action) and NLM ToxNet (https://toxnet.nlm.nih.gov/). Due to the different terminologies used to describe specific tests among the three data sources, a unified taxonomy was developed to which all source test types were mapped. Because a given chemical could have multiple tests available, with potentially disagreeing calls, the following decision logic was used to determine whether a chemical was potentially genotoxic for the purpose of this illustration. Progression through the steps would occur until a decision was made.

Example stepwise decision logic for genotoxicity:

1. If Ames test or *in vivo* micronucleus test is positive: positive
2. If Ames is negative: negative
3. If 2 or more tests other than Ames or *in vivo* micronucleus test are positive: positive
4. If 2 or more tests other than Ames or *in vivo* micronucleus test are negative: negative
5. If 1 test other than Ames or *in vivo* micronucleus test is positive: positive (single report)
6. If 2 test other than Ames or *in vivo* micronucleus test is negative: negative (single report)

For Methods 1 – 4, the positive and negative calls for genotoxicity were scored based on Table A1.3. Missing values are assigned a score of 0, to flag them for future data needs.

Table A1.3: Example component scoring approach for genotoxicity

Score	Genotoxicity Call
1	Negative or negative (single report)
2	Positive (single report)
3	Positive

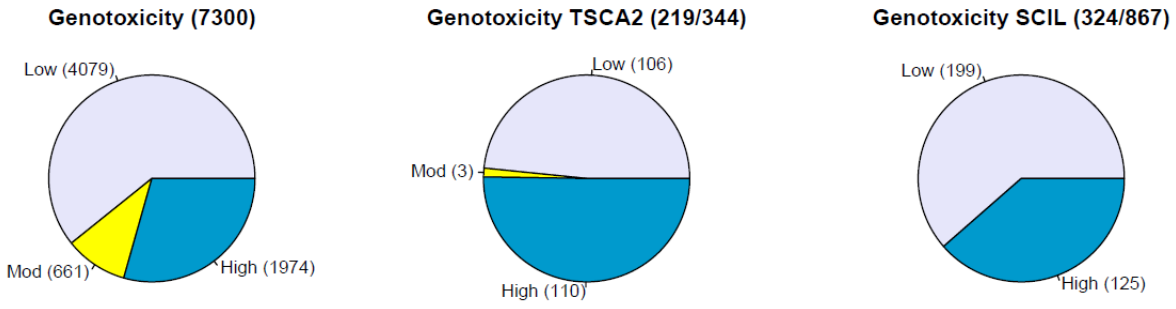


Figure A1.3: Distribution of chemicals across the genotoxicity component score bins. The first pie chart shows data for all chemicals in the ToxValDB database which includes TSCA Step 2

chemicals, SCIL chemicals, and other chemicals not on these lists (a total of 7300). The second and third pie charts show data for the TSCA Step 2 and SCIL chemicals, respectively. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2or SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Carcinogenicity Indications from *In Vivo* Toxicity Studies

Carcinogenicity was scored using cancer determinations reported by IARC, EPA IRIS, NTP, EPA OPP, EPA NCEA/PPRTV, CalEPA, Health Canada, and NIOSH. A total of 965 chemicals had information from one of these sources. For the purposes of this example, chemicals with any positive cancer classification (e.g. "probable human carcinogen", "possible human carcinogen", "known human carcinogen") were given a score of 3; chemicals with annotations of evidence of non-carcinogenicity for humans were given a score of 1; and chemicals with classifications such as "not classifiable as to its carcinogenicity" were excluded (effectively given a score of 0).

Predicted Effect Levels using Quantitative Structure Activity Relationships (QSAR)/Quantitative Biological Activity Relationships (QBAR)

EPA ORD has published a computational model for predicting *in vivo* effect levels for rodent repeat dose toxicity studies based on chemical structure, high-throughput *in vitro* assay data from ToxCast, and high-throughput toxicokinetic data (Truong et al., 2017). For Methods 2 – 4, the predicted effect levels are converted into component scores using the example cutoffs for oral chronic *in vivo* toxicity studies (Table A1.1). In this example, missing data are assigned a value of 0. Depending on the Method, the component score is used as the equivalent of the chronic *in vivo* toxicity studies.

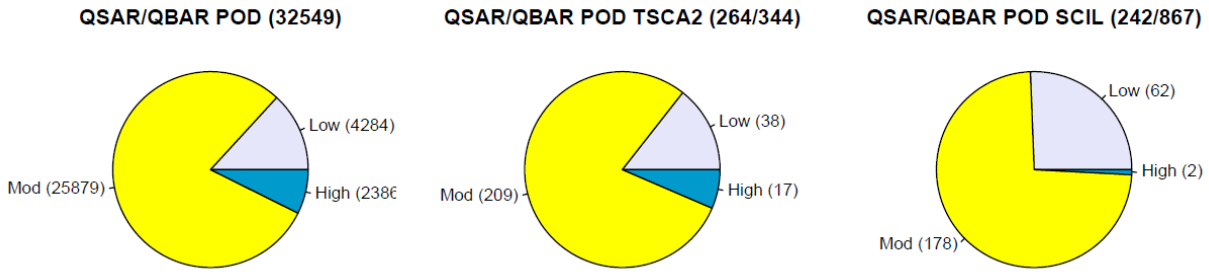


Figure A1.4: Distribution of chemicals across the QSAR/QBAR predicted effect level scoring bins. The first pie chart shows data for all chemicals in the database which includes TSCA Step 2 chemicals, SCIL chemicals, and other chemicals not on these lists (total of 32,549). The second and third pie charts show data for the TSCA Step 2 and SCIL chemicals, respectively. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2 or SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Predicted Effect Levels Using *In Vitro* Bioactivity

Conservative estimates of *in vivo* effect levels are estimated by combining *in vitro* potency values from ToxCast assays (in units of μM) with high-throughput toxicokinetic models. The

high-throughput toxicokinetics approach (HTTK) involves *in vitro* measurements of plasma protein binding and hepatic clearance using primary human hepatocytes. The *in vitro* measurements are scaled using *in vitro*-to-*in vivo* extrapolation (IVIVE) approaches and used to parameterize a one compartment toxicokinetic model. The HTTK models are used to estimate the administered dose equivalent (mg/kg/day) given a steady state blood concentration (Rotroff et al., 2010; Wambaugh et al., 2015; Wetmore et al., 2014; Wetmore et al., 2015). Because not all assay data are of equally high confidence, only ToxCast assays with less than 3 assay quality flags are used (Filter et al., 2016). If there are more than 3 *in vitro* assay hits for a chemical, the second lowest potency value is selected, otherwise the lowest potency value is used. The potency values are then converted into an *in vivo* effect level using the corresponding HTTK model (Judson et al., 2011). The approach generally provides a conservative estimate an *in vivo* effect level for a broad range of study types (Wetmore et al., 2013) and has been proposed for use in the Health Canada CMP (HC, 2016).

For Methods 2 – 4, the estimated effect levels are converted into component scores using the cutoffs for oral chronic studies (Table A1.1). In this example, missing data are assigned a value of 0. Depending on the Method, the component score is used as the equivalent and has been *in vivo* toxicity studies. Due to the conservative nature of the *in vitro* bioactivity derived effect levels (Wetmore et al., 2013), most chemicals scored using this approach are generally binned as High (Fig. A1.5).

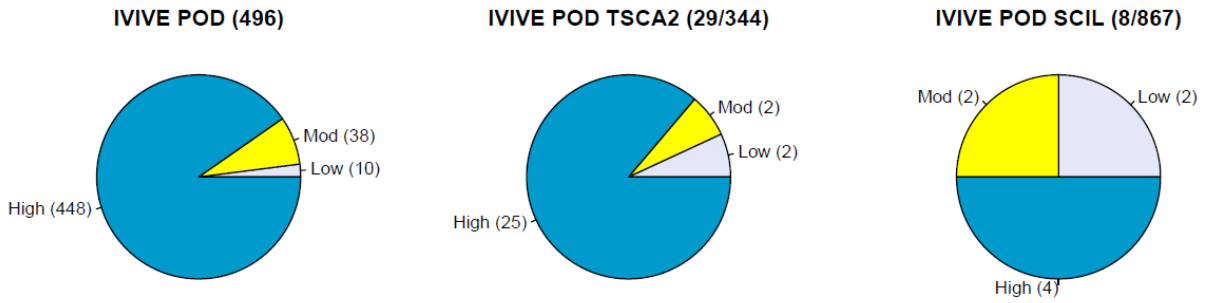


Figure A1.5: Distribution of chemicals across the *in vitro* bioactivity derived effect level component scoring bins. Numbers of chemicals in the database in each bin are given in parentheses. The first pie chart shows data for all chemicals in the database which includes TSCA Step 2 chemicals, SCIL chemicals, and other chemicals not on these lists. The second and third pie charts show data for the TSCA Step 2 and SCIL chemicals, respectively. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2 or SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Generalized Read-Across Predictions of Qualitative *In Vivo* Toxicity

EPA ORD has developed a generalized read-across (GenRA) approach to predict whether a chemical will show adverse effects in a specific study type (Shah et al., 2016). The approach uses both chemical structure and *in vitro* bioactivity (where available) to identify chemical analogs to determine/predict *in vivo* responses. The output from the GenRA model specifies probabilities from 0 to 1 for specific organ responses (e.g., liver, kidney) in a particular study

type (e.g., subchronic, developmental). The study types predicted by GenRA are subchronic, chronic, reproductive, developmental and developmental neurotoxicity. For Methods 2 – 4 in this illustration, the maximum probability across all organ responses in a specific study type was selected and used to calculate a component score based on Table A1.4. Depending on the Method, the component score is used as the equivalent of the *in vivo* subchronic/chronic, reproductive, developmental, and neurotoxicity studies. In this example, missing data are assigned a value of 0. The GenRA predictions resulted in a high proportion of chemicals in the High bin (Fig. A1.6).

Table A1.4: Example component scoring approach for GenRA predictions of qualitative *in vivo* toxicity

Score	Range of maximum GenRA probabilities
1	>0 to 0.5
2	0.5-0.9
3	>0.9

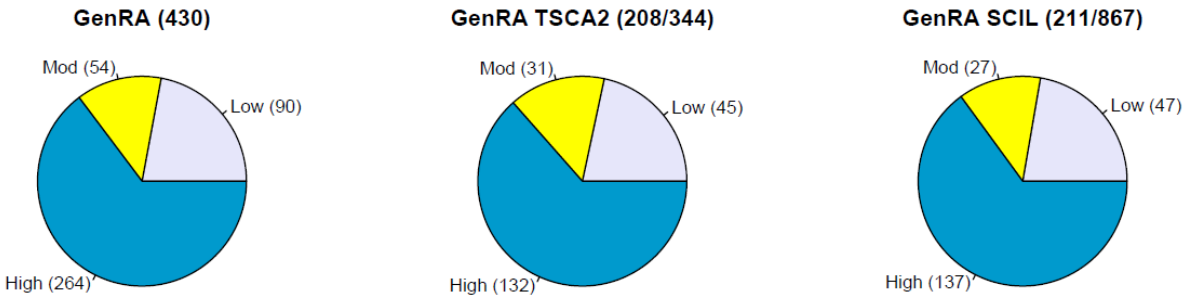


Figure A1.6: Distribution of chemicals across the GenRA component scoring bins. The first pie chart shows data for all chemicals in the database which includes TSCA Step 2 chemicals, SCIL chemicals, and other chemicals not on these lists (total of 430). The second and third pie charts show data for the TSCA Step 2 and SCIL chemicals, respectively. Note that the current GenRA data was only generated for the chemicals in these two collections. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2 or SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Endocrine Disruption Potential Based on In Vitro Bioactivity

EPA ORD has screened ~1,800 chemicals across a suite of high-throughput *in vitro* assays for estrogen receptor (ER) and androgen receptor (AR) bioactivity. Based on the results from the high-throughput *in vitro* assays, consensus computational models for ER and AR activity were developed in both agonist or antagonist modes (Judson et al., 2015; Kleinstreuer et al., 2017). The computational models yield scores in the range of 0 (inactive) to 1 (as active as potent natural hormones). The standard cutoff for ER or AR activity is a model score of 0.1 (i.e., chemicals with scores ≥0.1 have significant evidence for interacting with the appropriate receptor in the agonist or antagonist mode). For Methods 2 – 4, the ER and AR model scores are

converted into a component score based on Table A1.5. If the chemical has not been tested in the ER or AR assays, it is assigned a component score of 0. Most of the chemicals evaluated were in the Low bin for ER and AR bioactivity potential (Fig. A1.7).

Table A1.5: Example component scoring for ER and AR *in vitro* bioactivity

Score	ER agonist	ER antagonist	AR agonist	AR antagonist
1	Score<0.1	Score<0.1	Score<0.1	Score<0.1
3	Score≥0.1	Score≥0.1	Score≥0.1	Score≥0.1

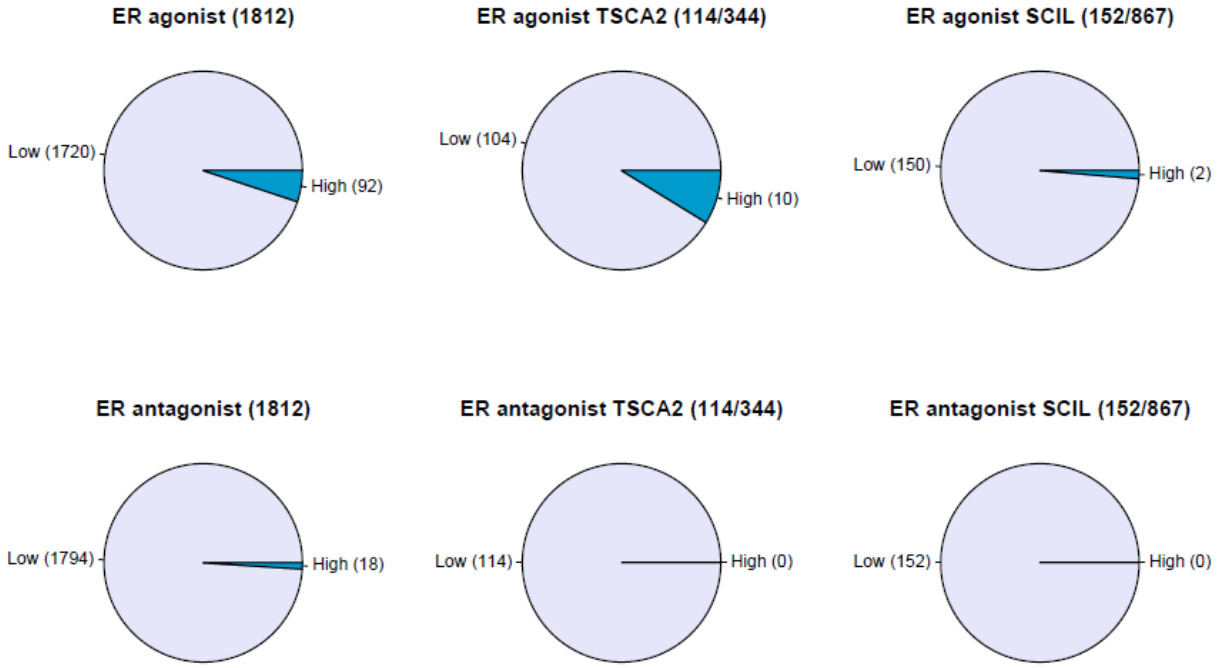


Figure A1.7: Distribution of chemicals for the ER and AR bioactivity component scoring bins. The left-hand column shows binning for all chemicals in the database which includes TSCA Step 2 chemicals, SCIL chemicals, and other chemicals not on these lists (1812 for ER, 1855 for AR). The second and third pie charts show data for the TSCA Step 2 and SCIL chemicals, respectively. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2or SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Endocrine Disruption Potential Based on QSAR models.

EPA ORD worked with a large international consortium of QSAR modelers to build many QSAR models, and a consensus model of ER activity, based on the 1812 chemical data set described above (Mansouri et al., 2016). QSAR models were built for both ER agonist and antagonist activity. Component scores were assigned based on Table A1.6. Most of the chemicals evaluated were in the Low bin for both ER agonist and antagonist potential (Fig. A1.8). Missing chemicals were assigned a score of 0.

Table A1.6: Example component scoring for ER QSAR models

Score	ER agonist	ER antagonist
1	Predicted inactive	Predicted inactive
3	Predicted active	Predicted active

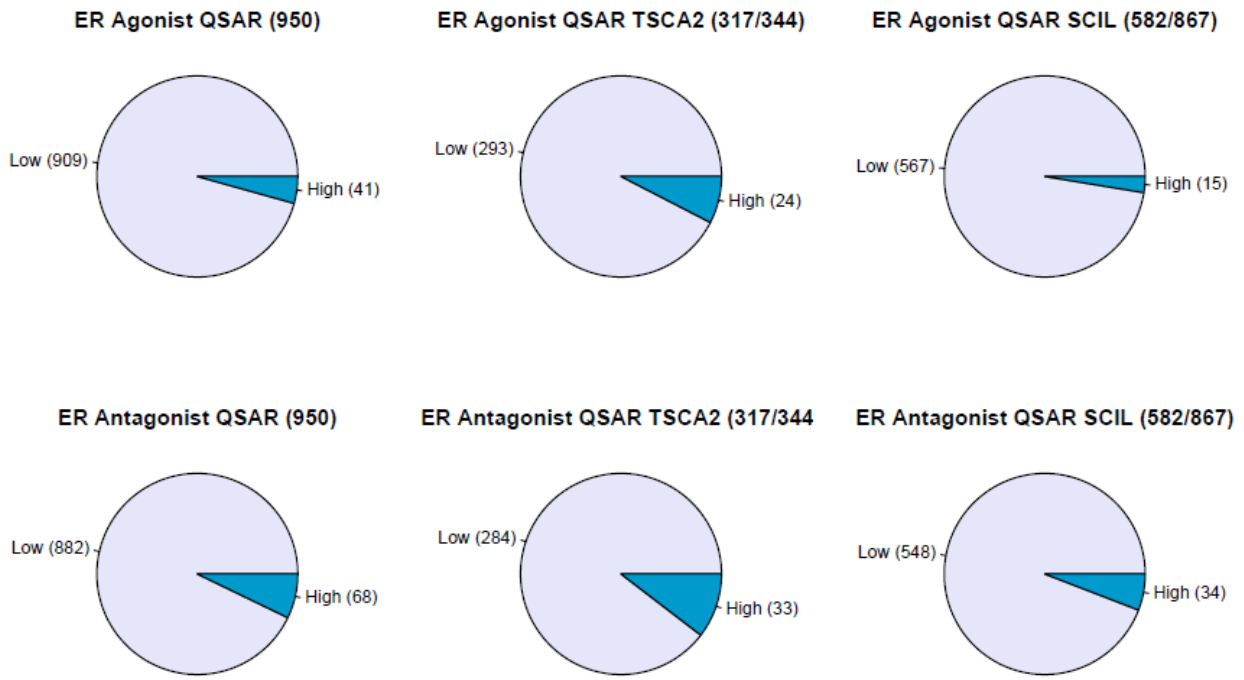


Figure A1.8: Distribution of chemicals for the ER and AR QSAR model component scoring bins. The left-hand column shows binning for the overlap of the chemicals evaluated by the ER and AR QSAR model (~32,000) with the TSCA Step 2 and SCIL chemicals. The second and third pie charts show data for the TSCA Step 2 and SCIL chemicals, respectively. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2or SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Ecological Hazard

Ecological hazard data are extracted from the EPA ToxValDB database where it had been compiled from the EPA ECOTOX database. Although data are available for a variety of species, only data for aquatic species are used in the current illustration. The data can come from any of the following study types: mortality:acute, mortality:chronic, reproductive:acute, reproductive:chronic, growth:acute, growth:chronic (all from ECOTOX). The types of effect levels are LDxx/LCxx/ECxx/EDxx where xx can range from 1% to 100%, and LOEL/NOEL/LOEC/NOEC. Values must be in units of mg/L. For each chemical, the lowest toxicity value was separately determined for acute and chronic studies, regardless of species. The ecological hazard domain score is then calculated based on cutoffs from Table A1.7. In this example, missing data are assigned a component score of 0. Similar to human hazard, these cutoffs put the majority of tested chemicals into the High hazard bin (Figure A1.9)

Table A1.7: Example component scoring cutoffs for ecological hazard based on the TSCA 2012 Work Plan Methods Document.

Score	Range of effect levels (mg/L) (acute)	Range of effect levels (mg/L) (chronic)
1	POD>100	POD>10
2	10<POD<100	1<POD<10
3	POD<10	POD<1

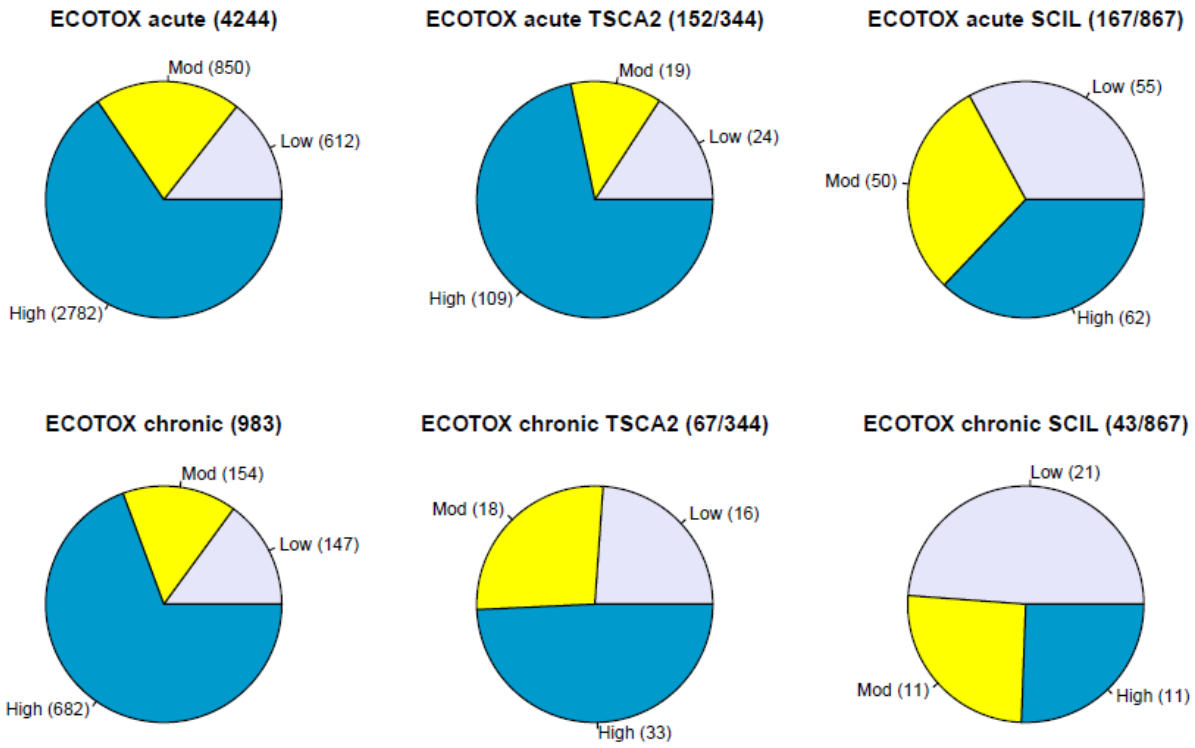


Figure A1.9: Distribution of ecological hazard domain scores for acute (top) and chronic (bottom) aquatic toxicity studies. The first pie chart shows data for all chemicals in the database which includes TSCA Step 2 chemicals, SCIL chemicals, and other chemicals not on these lists (4244 for acute, 983 for chronic). The second and third pie charts show data for the TSCA Step 2 and SCIL chemicals, respectively. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2 or SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Exposure Domain

Data components in the exposure domain consist of chemical release data, chemical use data, media occurrence reports, human biomonitoring and ecological monitoring data, and quantitative exposure predictions.

Toxics Release Inventory

Data from the Toxics Release Inventory (TRI) was obtained from the EPA website (https://www.epa.gov/toxics-release-inventory-tri-program). The TRI component scoring is only based on the volume of release (Table A1.8).

Table A1.8: Example TRI component scoring based on the TSCA 2012 Work Plan Methods Document.

Score	TRI Releases
1	Total < 5000
2	5000 ≤ total <100000
3	Total ≥ 100000

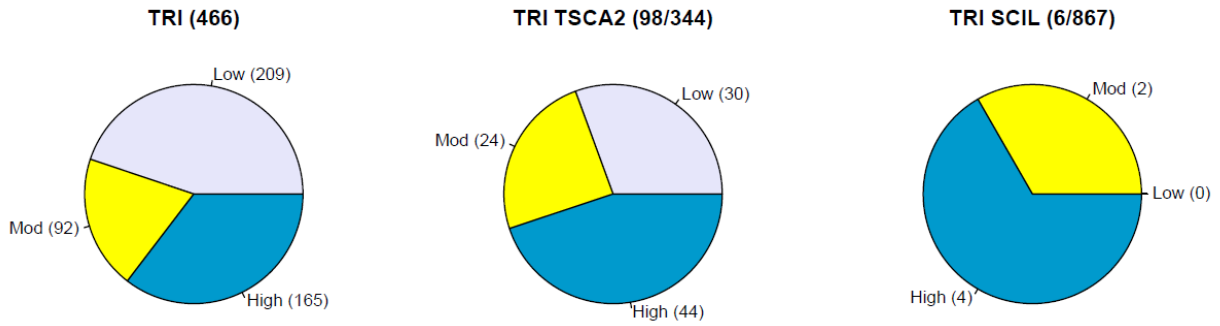


Figure A1.10 Distribution of component scores in the TRI categories. The first pie chart shows data for all chemicals in the database which includes TSCA Step 2 chemicals, SCIL chemicals, and other chemicals not on these lists (466). The second and third pie charts show data for the TSCA Step 2 and SCIL chemicals, respectively. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2 or SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Chemical Data Reporting

The 2016 Chemical Data Reporting (CDR) information was downloaded from the EPA website (https://www.epa.gov/chemical-data-reporting/2016-chemical-data-reporting-results). The CDR component scoring is based on three classes of information: production volume, number of manufacturing sites, and specific use classes, which are “commercial”, “industrial”, “consumer” and “child”. The component scoring categories for production volume and the number of manufacturing sites are provided in Table A1.9. Note that chemicals for which the PV values were withheld were assigned a score of 3. For the categorical variables, component scores were assigned as 1 for no use, 3 for use, and 0 for missing data.

Table A1.9: Example CDR component scoring for quantitative variables based on the TSCA 2012 Work Plan Methods Document.

Score	Production Volume (lbs)	Number of manufacturing sites
1	PV<500,000	1 to 99

2	500,000≤PV<1,000,000	100 to 999
3	PV≥1,000,000 or "Withheld"	≥1000

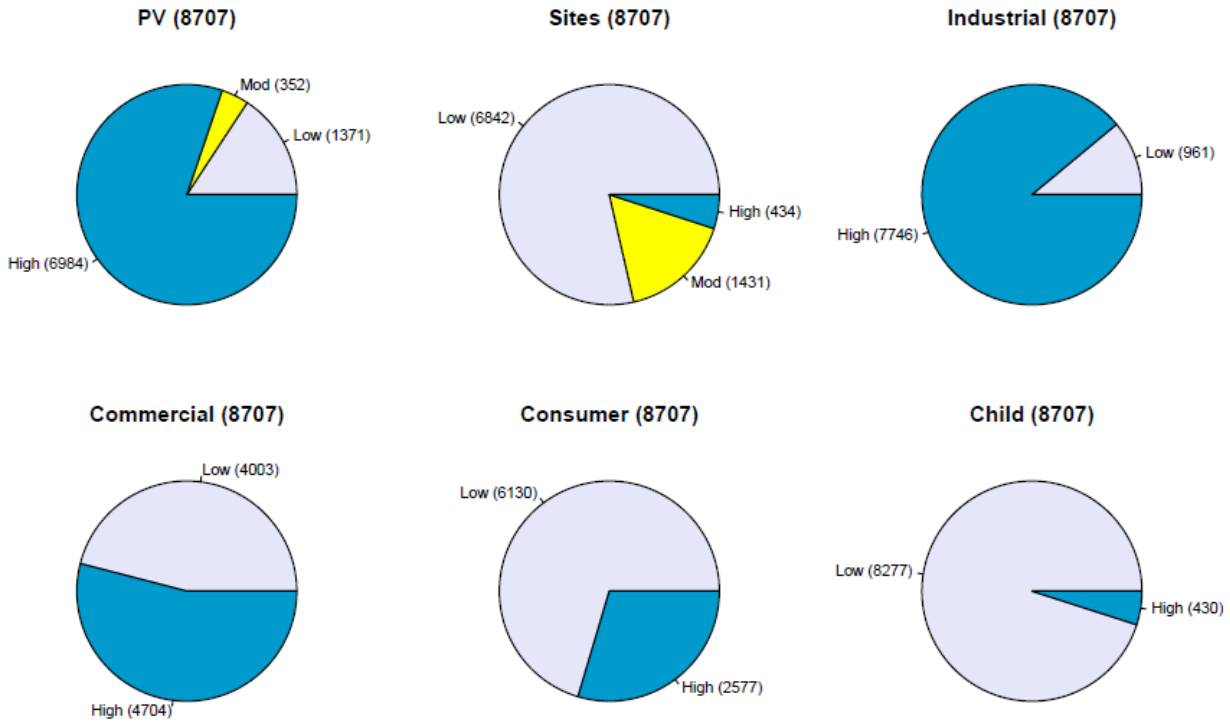


Figure A1.11a: Distribution of component scores across the CDR categories. Numbers of chemicals in the database in each bin are given in parentheses. Data is shown for the complete database of 8707 chemicals which includes TSCA Step 2 chemicals, SCIL chemicals, and other chemicals not on these lists.

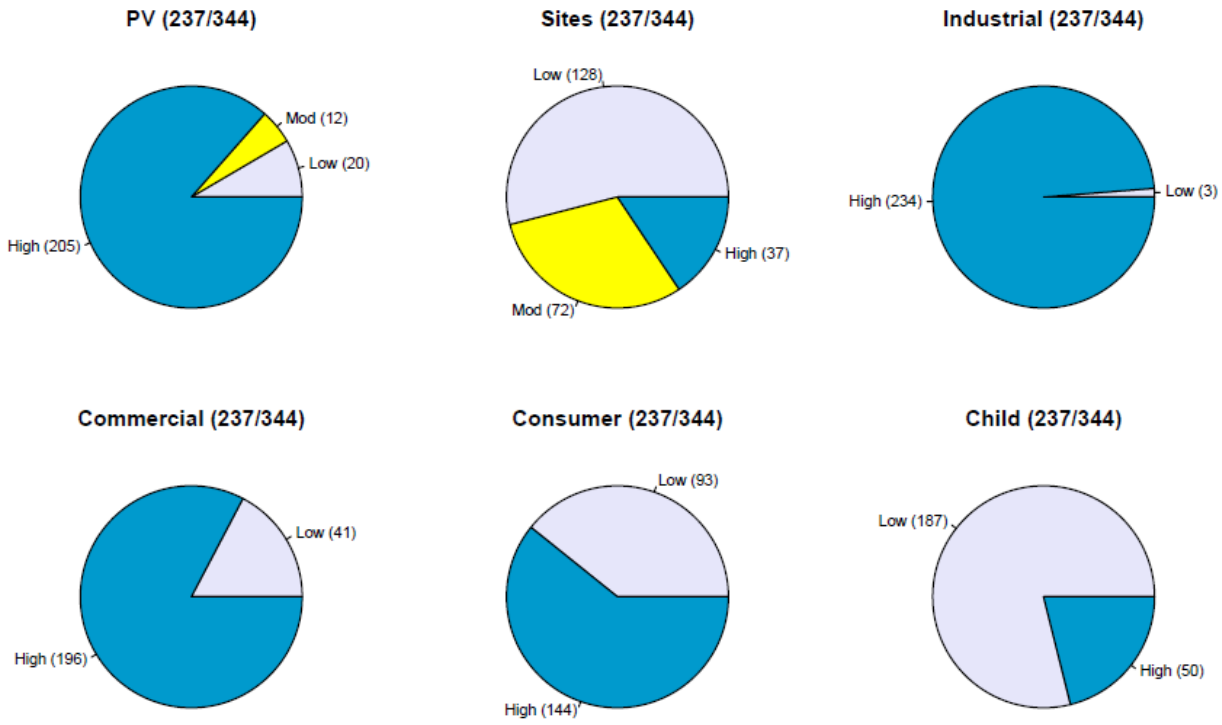


Figure A1.11b: Distribution of component scores across the CDR categories. Data is shown for the TSCA Step 2 chemicals. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2) with data in that domain. The second value is the total number of chemicals in the collection.

Figure A1.11c: Distribution of component scores across the CDR categories. Data is shown for the SCIL chemicals. In the title, the first value in parentheses is the number of chemicals in the chemical collection (SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Qualitative Metrics Related to Human Biomonitoring, Residential Exposure, Consumer Use, Dietary Exposure, Child Exposure, and Environmental Exposure

A range of data sources were compiled that represent aspects of human and ecological exposure including human biomonitoring, residential exposure, consumer use, dietary exposure, child exposure, and environmental exposure (Table A1.10). For each of the variables, the values across all chemicals in the database are transformed into 10 quantiles. Values in the upper (90%-100%) quantile are assigned a value of 3, those in the 50%-90% quantiles are assigned a value of 2, and those below the 50% quantile are assigned a value of 1. For each class (e.g. biomonitoring), the maximum score for the corresponding variables for a chemical is assigned as that chemical's class score. By definition, 50% of chemicals are in the Low bin, 40% in the Moderate bin and 10% in the High bin for each of these metrics. Note that all of the class scores, with the exception of "Environmental Exposure" are used in the human exposure domain scoring. Functional use data as specified in the preceding sections could also be used to inform relevant exposure categories.

Table A1.10: Definitions of variables used in defining qualitative human and ecological exposure component scores.

Variable Name	Class	Description
OPPT measured biosolids N	Biomonitoring	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Biosolids
OPPT measured breastmilk N	Biomonitoring	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Breast milk
OPPT measured human blood N	Biomonitoring	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Human blood (whole/serum/plasma)
OPPT measured human other N	Biomonitoring	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Unclassified Human Samples
OPPT measured human urine N	Biomonitoring	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Human Urine
OPPT measured skin wipes N	Biomonitoring	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Skin Wipes
CPDat reported categories child N	Child exposure	Number of unique consumer product formulation categories for children's products (e.g., children's sunscreens or powders, diaper cream, play dough) (source: EPA ORD CPDat database)
NERL potential child exposure	Child exposure	The chemical has potential for children's exposure based on use profile.
CPDat reported categories consumer a N	Consumer use	Number of unique consumer product formulation categories (e.g., personal care products, cleaning products, home improvement products, etc.) associated with chemical (out of 313 total) (source: EPA ORD CPDat database)
CPDat reported categories consumer b N	Consumer use	Number of unique individual consumer product formulations containing chemical (as identified by MSDS or ingredient lists) (source: EPA ORD CPDat database)
CPDat reported categories indoor N	Consumer use	Number of associated unique consumer product formulation categories with indoor releases (source: EPA ORD CPDat database)
NCCT predicted exposure diet prob	Dietary exposure	Probability of exposure via dietary pathway, from random forest algorithm. (source: EPA ORD predictions)
OPPT measured food N	Dietary exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Food product

OPPT measured water drinking N	Dietary exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Drinking water
OPPT measured ecological other N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Unclassified Ecological Species
OPPT measured landfill leachate N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Landfill leachate
OPPT measured other environmental N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Other-environmental
OPPT measured sediment N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Sediment
OPPT measured soil N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Soil, or outdoor settled dust
OPPT measured vegetation N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Vegetation and food other than fish
OPPT measured water ground N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Groundwater
OPPT measured water surface N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Surface water
OPPT measured water waste N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Wastewater (influent, effluent)
OPPT measured wildlife aquatic invertebrates N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Wildlife (Aquatic Invertebrates)
OPPT measured wildlife aquatic vertebrates N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Aquatic vertebrates
OPPT measured wildlife birds N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Terrestrial Vertebrates
OPPT measured wildlife fish N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Wildlife (Fish)

OPPT measured wildlife terrestrial vertebrates N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Wildlife (Birds)
OPPT measured wildlife worms N	Environmental exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Wildlife (Worms)
USGS measured water samples ALD N	Environmental exposure	Total number of samples with measured value greater than limit of quantification in samples (Nsamps) collected between 1995 and 2014 in the 48 contiguous United States from the NWQMCWQP, collected by USGS and USEPA sampling programs.
USGS measured water samples N	Environmental exposure	Total number of samples evaluated for this chemical between 1995 and 2014 in the 48 contiguous United States in the National Water Quality Monitoring Council's Water Quality Portal (NWQMCWQP), aggregating data from USGS and USEPA sampling programs.
USGS measured water sites N	Environmental exposure	Total number of geographic sites with at least one sample evaluated for this chemical in the NWQMCWQP between 1995 and 2014 in the 48 contiguous United States.
USGS measured water sites UC N	Environmental exposure	Total number of geographic sites with at least one sample evaluated for this chemical in the NWQMCWQP between 1995 and 2014 in the 48 contiguous United States (upper confidence interval)
NCCT predicted exposure farfield prob	Industrial use	Probability of exposure via far-field industrial pathway, from random forest algorithm. (source: EPA ORD)
NERL mean concentration ug per g	Product composition	Mean weight fraction of the chemical in a product (e.g. a surfactant in a detergent).
NERL number of products	Product composition	Number of products with the chemical (e.g. how many detergents in a database of consumer products contain this chemical)
NCCT predicted exposure residential prob	Residential exposure	Probability of exposure via residential ("near field") pathway, from random forest algorithm. (source: EPA ORD)
OPPT measured air ambient N	Residential exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Ambient Air
OPPT measured air indoor N	Residential exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Indoor air

OPPT measured indoor dust N	Residential exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Indoor Dust
OPPT measured products N	Residential exposure	Number of data sources (studies) in which chemical was observed in media in EPA OPPT monitoring database v1.3: Consumer Products

Quantitative Metrics Related to Human Exposure and Biomonitoring

A range of data sources were compiled that provide quantitative estimates of human exposure and biomonitoring values (Table A1.11). Similar to what was done for the qualitative exposure metrics, for each variable, the values across all chemicals in the database are transformed into 10 quantiles. Values in the upper (90%-100%) quantile are assigned a value of 3, those in the 50%-90% quantiles are assigned a value of 2, and those below the 50% quantile are assigned a value of 1. For each class (e.g. biomonitoring), the maximum score for the corresponding variables for a chemical is assigned as that chemical's class score.

Table A1.11: Definitions of variables used in defining quantitative human exposure component scores.

Variable Name	Class	Description
FDA measured exposure	Exposure quantile	Quantile for the Food and Drug Administration Cumulative Estimate of Daily Intake (CEDI) relative to all chemicals with CEDI intake rates. 99th percentile is highest exposure.
FDA measured exposure quantile	Exposure quantile	Quantity for the Food and Drug Administration Cumulative Estimate of Daily Intake (CEDI) relative to all chemicals with CEDI intake rates. 99th percentile is highest exposure.
SHEDS predicted exposure residential quantile	Exposure quantile	Quantile for SHEDS-HT residential pathway prediction, relative to all SHEDS-HT residential chemical predications. 99th percentile is highest exposure. (Isaacs et al., 2014)
NHANES predicted exposure median	Exposure quantitative	Median exposure rate (mg/kg bodyweight/day) inferred for the total U.S. population that can be inferred from the Centers for Disease Control and Prevention National Health and Nutrition Examination Survey. (Wambaugh et al., 2014)
NHANES predicted exposure upper95	Exposure quantitative	Upper 95% confidence limit on exposure rate (mg/kg bodyweight/day) inferred for the total U.S. population that can be inferred from the Centers for Disease Control and Prevention National Health and Nutrition Examination Survey. (Wambaugh et al., 2014)

SEEM3 predicted exposure median	Exposure quantitative	Median consensus model exposure rate (mg/kg bodyweight/day) based on third generation Systematic Empirical Evaluation of Models (SEEM). (Ring et al., in prep.)
SEEM3 predicted exposure upper95	Exposure quantitative	Upper 95% confidence limit on consensus model exposure rate (mg/kg bodyweight/day) based on third generation Systematic Empirical Evaluation of Models (SEEM). (Ring et al., in prep.)
SHEDS predicted exposure residential exposure	Exposure quantitative	Quantile for SHEDS-HT residential pathway prediction, relative to all SHEDS-HT residential chemical predications. 99th percentile is highest exposure. (Isaacs et al., 2014)

Summary Component Scores for Human and Ecological Qualitative and Quantitative Metrics

Using the qualitative and quantitative exposure metrics defined in Tables A1.10 and A1.11, a set of summary component scores are generated for each class (Table A1.12). Each of these scores for a chemical will have a value of 1 to 3. Missing scores are assigned a value of 0. For Methods 1 – 3, a chemical is assigned the maximum score across all of the classes. For Method 4, the sum across all of the component scores is used. For Method 5, the predicted upper limit of exposure for the chemical in mg/kg/day is used from the third generation of the Systematic Empirical Evaluation of Models (SEEM3) approach (Ring et al., in prep.). The SEEM3 approach is similar to that used in SEEM2 (Wambaugh et al., 2014). If a value is not calculated, no default is used. As stated above, all of these scores, with the exception of the ecological component score are used in the human exposure domain.

Table A1.12: Summary of sources / classes of qualitative and quantitative exposure metrics

Class Name	Use	Sources
Biomonitoring	Score increases with the number of cases where the chemical is seen in biomonitoring samples	OPPT Biomonitoring database
Residential	Score increases with the number of cases where the chemical is seen in residential samples	OPPT Biomonitoring database, SEEM3 predictions
Consumer	Score increases with the number of cases where the chemical is seen in consumer product samples	CPDat
Diet	Score increases with the number of cases where the chemical is seen in dietary samples	OPPT Biomonitoring database, SEEM3 predictions
Children	Score increases with the number of cases where the chemical is seen in samples in children	CPDat, NERL exposure models
Ecological	Score increases with the number of cases where the chemical is seen in ecological samples	OPPT Biomonitoring database, USGS environmental sampling
Quantile	Quantile exposure levels	NHANES, FDA, SHEDS, SEEM3
Quantitative	Quantitative (mg/kg/day) exposure estimates	NHANES, FDA, SHEDS, SEEM3

The distribution of component scores in exposure classes are shown in Figure A1.12. Note that the rolled-up classes do not break out into the 50/40/10 distribution of the individual quantile binned source data, because multiple sources go into each category.

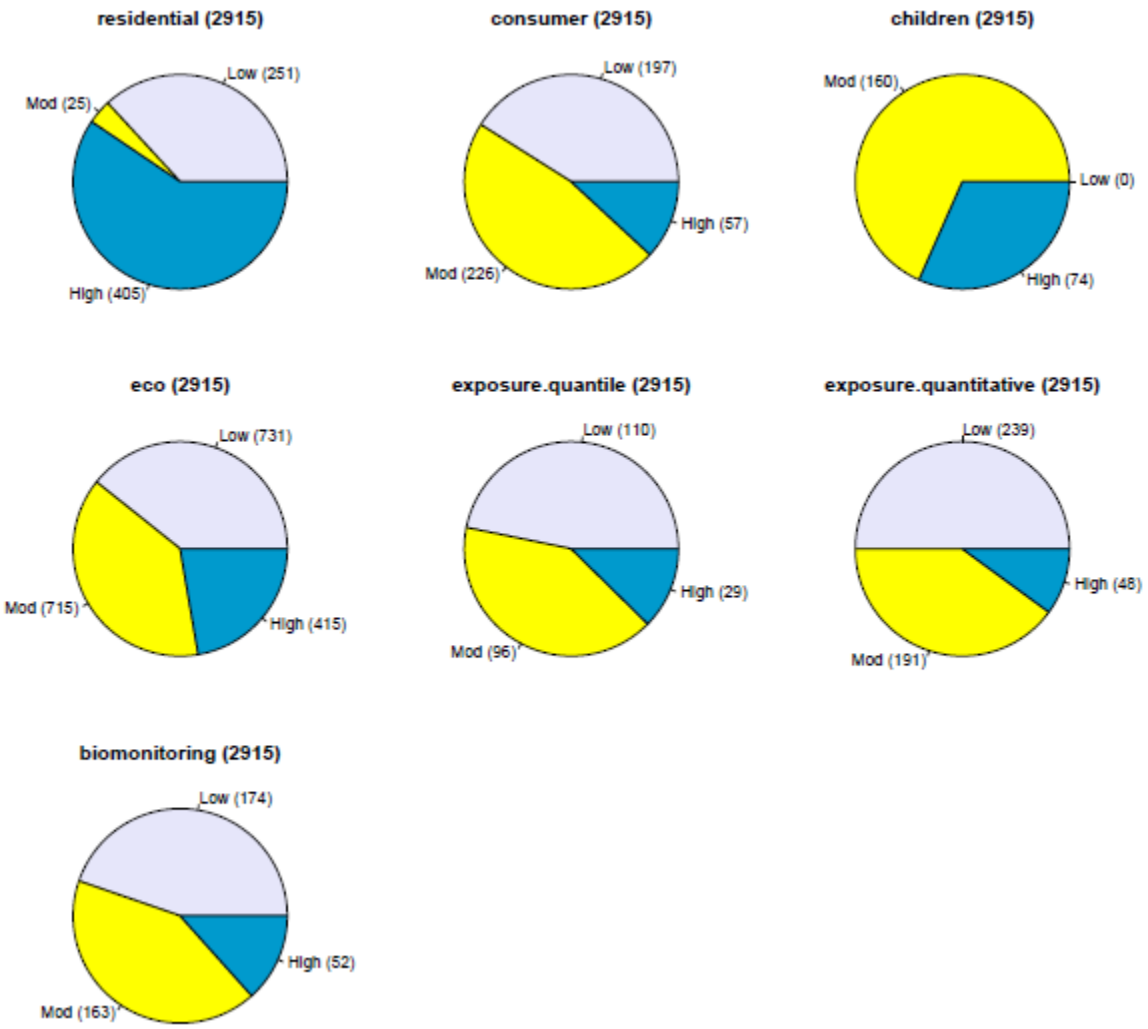


Figure A1.12a: Distribution of component scores across the exposure classes. Data is shown for all chemicals in the database (2915) which includes TSCA Step 2 chemicals, SCIL chemicals, and other chemicals not on these lists.

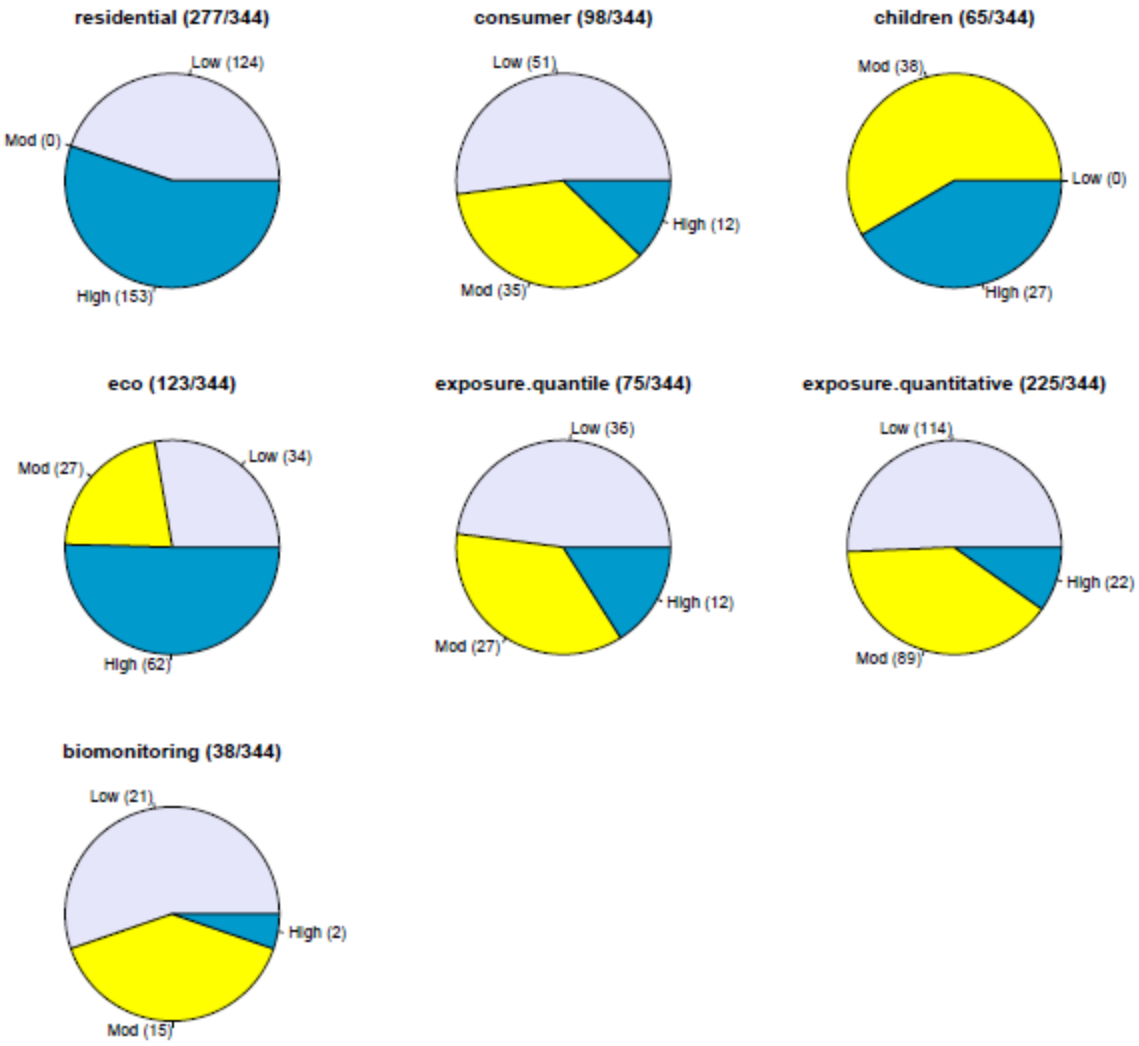


Figure A1.12b: Distribution of component scores across the exposure classes. Data is shown the TSCA Step 2 chemicals. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2) with data in that domain. The second value is the total number of chemicals in the collection.

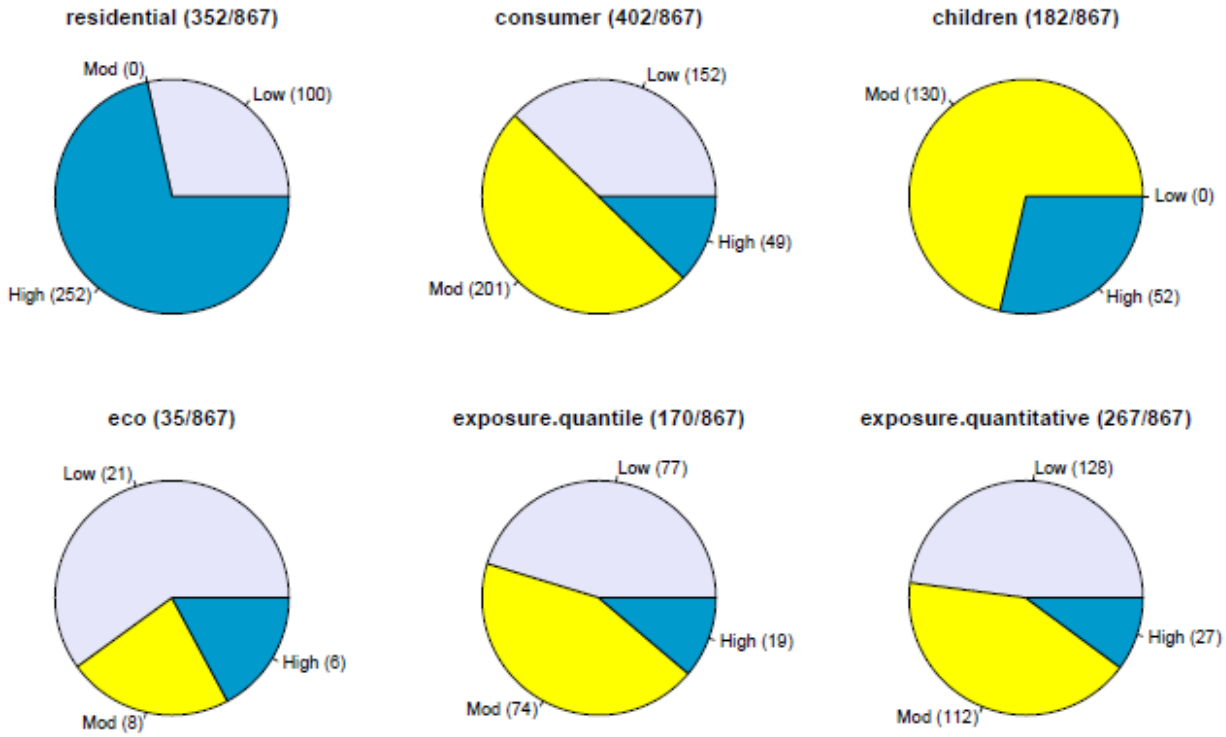


Figure A1.12c: Distribution of component scores across the exposure classes. Data is shown the SCIL chemicals. In the title, the first value in parentheses is the number of chemicals in the chemical collection (SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Persistence and Bioaccumulation Domain

Currently, variables representing environmental persistence and bioaccumulation are calculated from chemical structures using the OPERA suite of models (https://comptox.epa.gov/dashboard). Future efforts will fold in predictions and experimental measurements from EPI Suite due to their use in the TSCA 2012 Work Plan Methods. Because these are calculated properties, values are only available for distinct chemical structures, and for organic molecules (i.e., not for organometallics or inorganic chemicals). Organic salts are first converted to their corresponding neutral form before calculating properties. The rules for converting from physical values to component scores are given in Table A1.13. In this example, missing values are assigned a component score of 0.

Table A1.13: Example component score ranges for persistence and bioaccumulation

Score	Persistence	Bioaccumulation
1	Half-life<60 days	BAF<1000
2	Half-life from 60-180 days	1000≤BAF<5000
3	Half-life≥180 days	BAF≥5000

The distributions of component scores and physical parameter values, show that very few chemicals have moderate or high persistence or bioaccumulation component scores (Fig. 13).

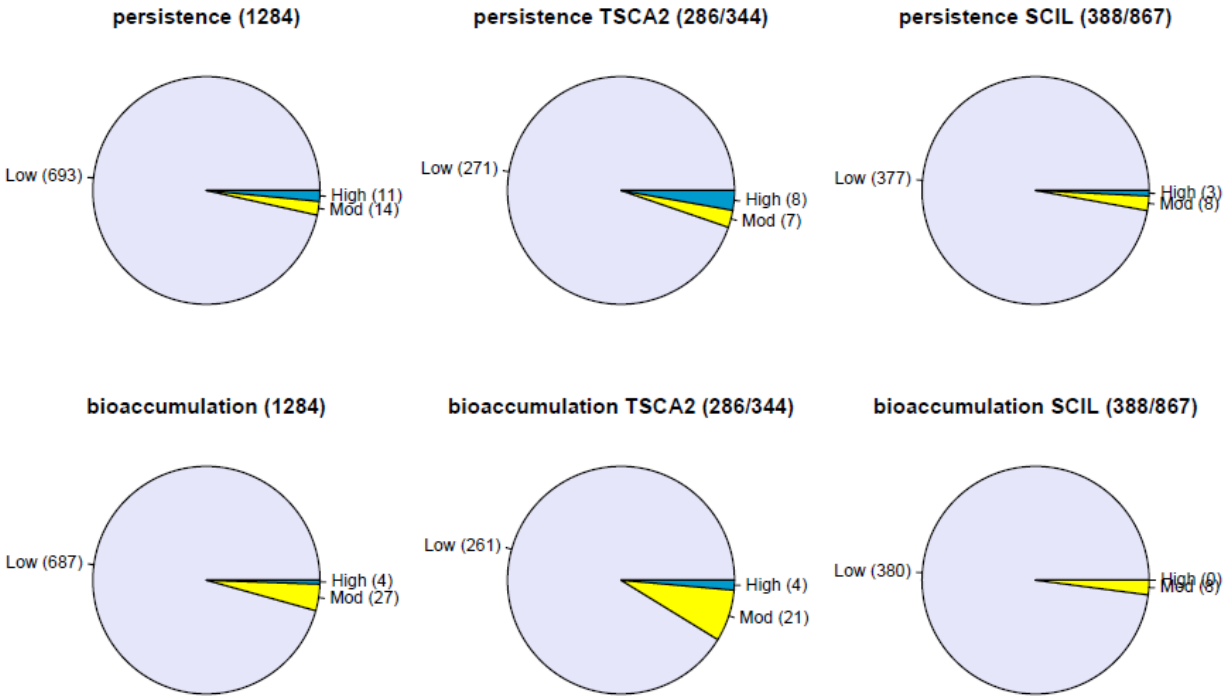


Figure A1.13: Distribution of persistence and bioaccumulation component scores. The first pie chart shows data for the merged list of TSCA Step 2 and SCIL chemicals. The second and third pie charts show data for the TSCA Step 2 and SCIL chemicals, respectively. In the title, the first value in parentheses is the number of chemicals in the chemical collection (TSCA2 or SCIL) with data in that domain. The second value is the total number of chemicals in the collection.

Appendix II

Prototype User Interface Design

The prototype application, RapidTox, is a workflow management tool and provides access to multiple workflows and pre-defined methods. The methods outlined above will be available.

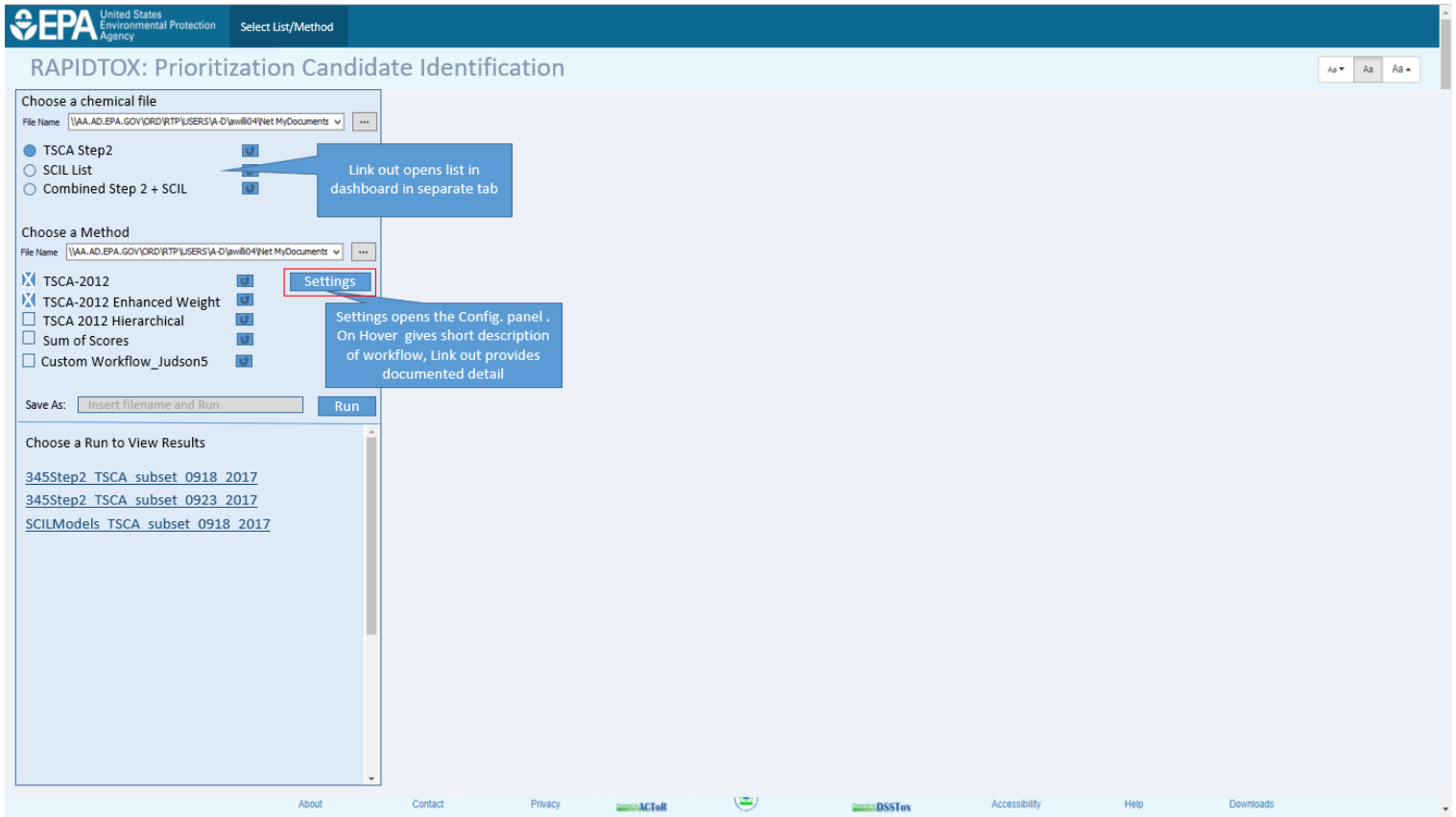


Figure A2.1: The user chooses a chemical list for processing through one or more methods. Selecting the out link adjacent to the chemical list will navigate the user to the chemical list on the CompTox Chemistry Dashboard. This will provide a listing of the

chemicals that can be downloaded in various formats and allows the user to navigate through available data for each of the chemicals one at a time. The out links adjacent to each of the methods provides documentation regarding the logic behind each of the methods.

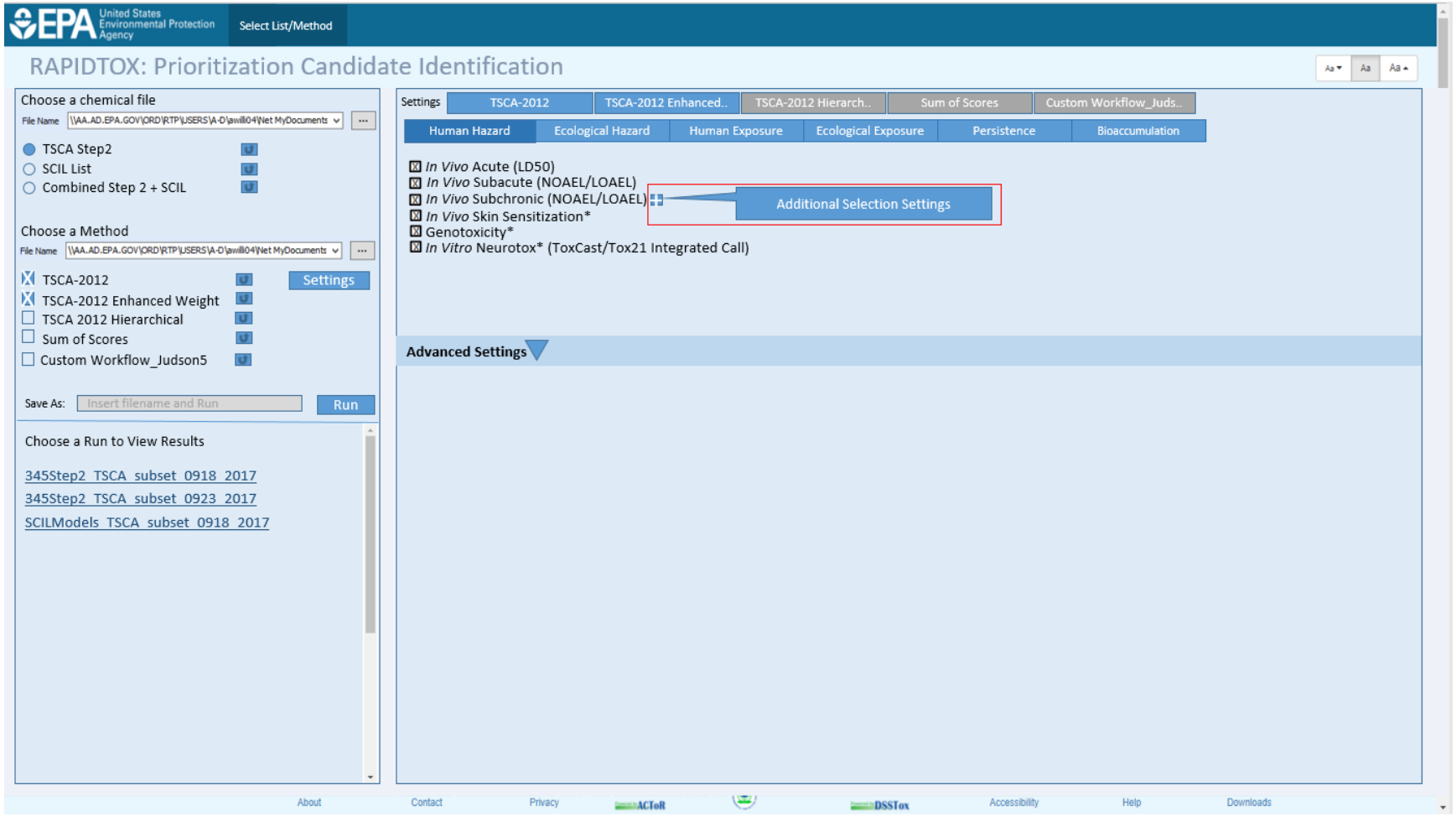


Figure A2.2: For each of the selected methods (only two are chosen in the figure above) the user navigates through the individual data stream tabs and adjusts the settings to their needs for a particular run. All settings utilized during a particular run are saved into the

results file. The user enters a file name into the Save As entry field before initiating the run. Previous runs are listed in the results panel (bottom left hand side) and can be easily loaded for review.

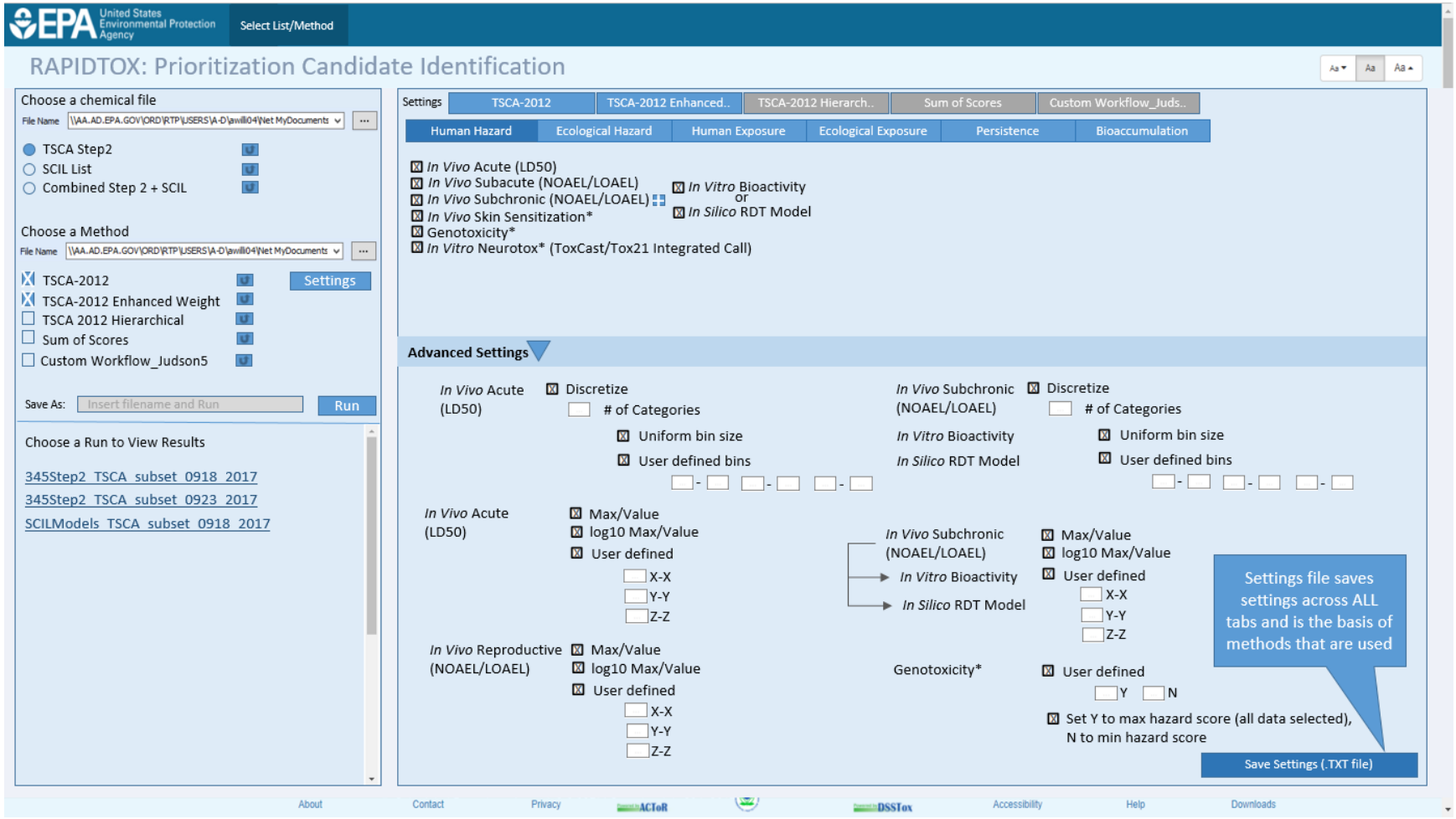


Figure A2.3: The prototype design includes Advanced Settings to increase flexibility and develop a customized prioritization workflow that may be implemented into a future version of the application.

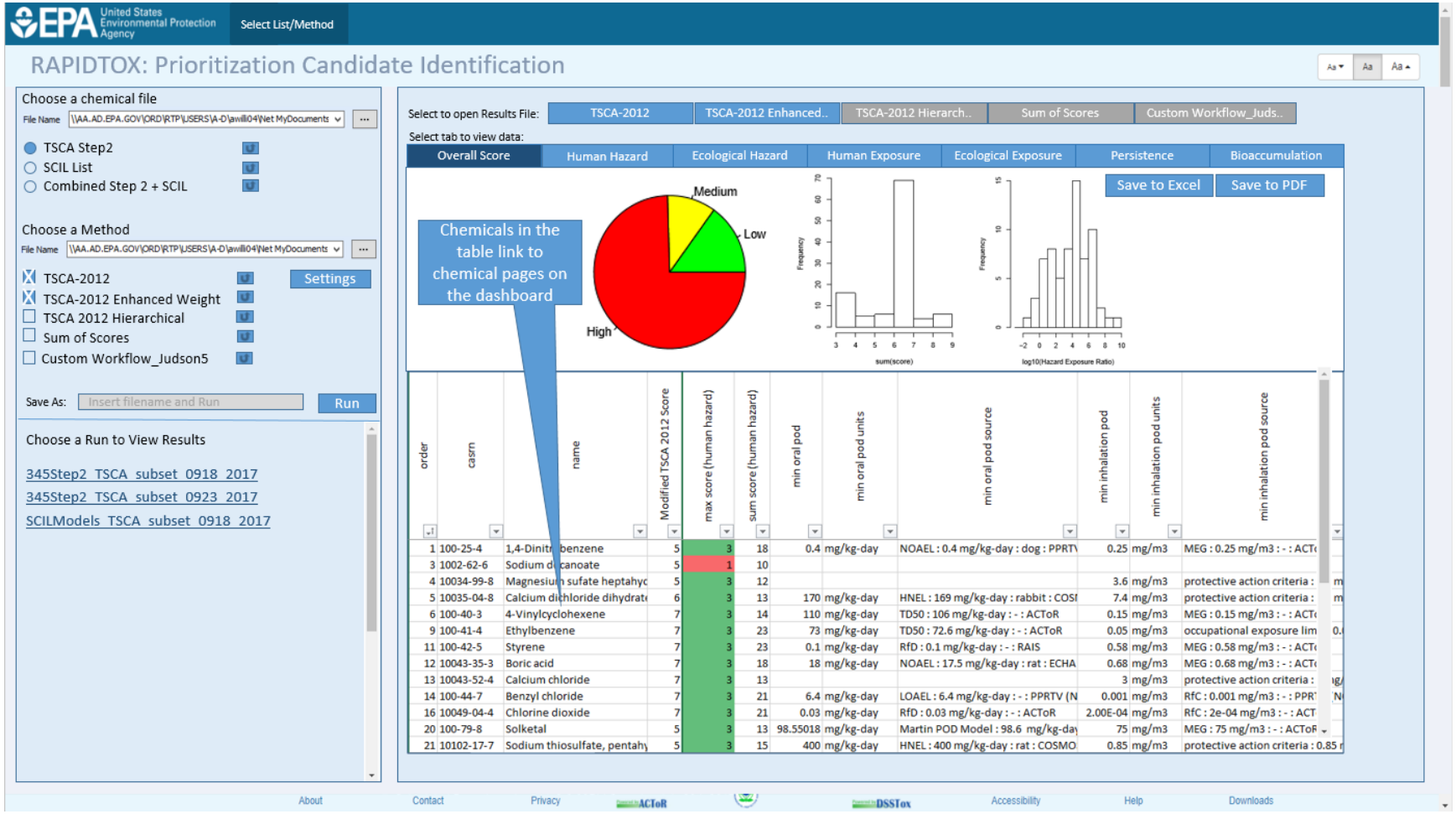


Figure A2.4: The results page will provide access to an *overall score* and relevant graphical representations of the results. Individual tabs will allow review of the results for each data type. The individual chemicals in the results table will be hyperlinked to the relevant

chemical page on the CompTox Chemistry Dashboard. The results from a particular run can be saved in various formats (e.g. as Excel and PDF)

Appendix III

ToxValDB Database

ToxValDB collects summary data from a large number of other *in vivo* databases/data sets. It includes multiple types of quantitative toxicity values and effect levels including LOAEL, NOAEL, NOEL, LOEL, RfD, RfC, etc. Values may be repeated (copied) from one database to another, so there is some amount of repeat data, and this repeating is tracked in one of the database tables. For the purpose of the current prioritization process, only a subset of the data in ToxValDB is used, restricted to "toxval types" of LO(A)EL, NO(A)EL, NO(A)EC, LO(A)EC, LD50, LC50, TD50, BMD (benchmark dose), cancer slope factors and cancer unit risk values. In the source databases, a variety of units are used, but these have been converted to mg/kg-day and mg/m3 for oral and inhalation values.

"Risk assessment class" values included are

- Acute (database value=acute)
- Subchronic (database value=subchronic, subacute, short-term, repeat dose)
- Chronic (database value=chronic)
- Reproductive (database values= reproductive, reproductive developmental)
- Developmental (database value=developmental)
- Neurotoxicity (database value=neurotoxicity)

For each chemical and for each of these risk assessment classes, we find the lowest oral and lowest inhalation toxicity values (if they exist).

All data in ToxValDB has a "use_me" flag which can be used to filter data based on priority. By default, data are given a value of use_me=1. Data to be excluded from any use is set to 0, and values to be used for prioritization is given a value from 2 to 5 (IRIS, PPRTV=5, and the other sources in Table A1.1 values of 2,3 or 4).

Data sources

Table A3.1: Data source, the toxval types and the number of unique chemicals across all of the toxval types for that data source.

Source	ToxvalDB Types	Chemicals
ECHA	NOEL,NOAEL,LOAEL,NOAEC,LOAEC,LEL,LEC,NOEC, LOEC, LOEL,BMDL,MTD,NEL,NOTEL,BMDL-10	2183
COSMOS	LOAEL,LD50,HNEL,LEL,NOEL,LOEL,NOAEL	1104
ToxRefDB	LOAEL,NOAEL,NEL,LEL	1086
Wignall	RfD,NOEL,RfC,NOAEL,cancer slope factor,BMCL,LOAEL,LEL,BMDL,BMD	959
HPVIS	LD50,LC50,NOAEL,LOAEL,LOAEC,NOAEC	799
DSSTox CPDBAS	TD50	788
RAIS	RfD,RfC,cancer slope factor	635

HESS	NOEL,LOEL	522
IRIS	RfC,RfD,cancer slope factor	439
PENN Properties	cancer slope factor,RfD	389
EPA OW HH Benchmarks Pesticides	cancer slope factor,RfD	375
EPA Pesticides HHBP	cancer slope factor,RfD	375
CalOEHHA ToxCriteria 2015	cancer slope factor,RfD,unit risk	287
CalOEHHA CancerPotValues	cancer slope factor	274
PPRTV (ORNL)	RfD,LOAEL,NOAEL,BMDL,BMDL-05,POD, BMDL-10,BMDL-01,NOEL,RfC,BMCL,NOAEL/LOAEL	266
EPA HAP ChronicRisk	cancer slope factor	248
HEAST	NOAEL,RfD,LOAEL,NOEL,RfC	208
PPRTV (NCEA)	RfC,BMCL,RfD,BMDL,NOEL,LOEL,NOAEL,LOAEL, cancer slope factor,BMDL-10	194
USGS HBSL 2012	cancer slope factor,RfD	179
EFSA	NOAEL,NOEL,NOEC,LC50,LOAEL,LD50,RfD, BMDL-10,BMDL-05	169
Alaska DEC CLs Guidance	cancer slope factor,RfC,RfD	164
MINN HealthRisk	cancer slope factor,RfD	94
UK PestUsage	LC50	34
HAWC	LOEL,NOEL	17
Health Canada	cancer slope factor	14
DOD Gulf Pest	cancer slope factor,LOEL,NOAEL,NOEL,RfD	12
CalOEHHA chRD	RfD	10

Data Curation

Data are imported into ToxValDB from many individual sources, with a variety of annotations (toxval types, units, duration or study class, exposure route, exposure method, etc.). Toxval types and units are transformed to standard types (and numerical values appropriately transformed). If ToxVal types and units do not match, the data are excluded (use_me=0). All of these transformations are carried out in software so that they can be rerun as more data are added. For prioritization, a key parameter is the type of study or duration, so information is extracted into a value called risk assessment class, which takes on the values of acute, subacute, subchronic, chronic, developmental, reproductive, reproductive developmental, short-term, repeat dose, and neurotoxicology.

Chemical information from the sources is CASRN and / or name. When one of these identifiers is missing, we use the batch download tool in the CompTox dashboard to generate the missing values. The ToxValDB database then holds a CASRN/name pair for each record.

For many chemicals, there are multiple studies, sometimes even for the same risk assessment class. One may expect these values to show some relationship with one another, so we carry out a QC process looking at all of the toxicity values for each chemical and look for suspicious outliers. This could arise from mistakes in the original database, transcription errors in creating ToxValDB, unit conversion errors, etc. During this QC process, we attempt to trace the source of the error and fix where possible. However, there are cases where the actual data are discrepant, and so no changes are made to the data unless an obvious error is found. An additional process for manual curation is currently being defined.

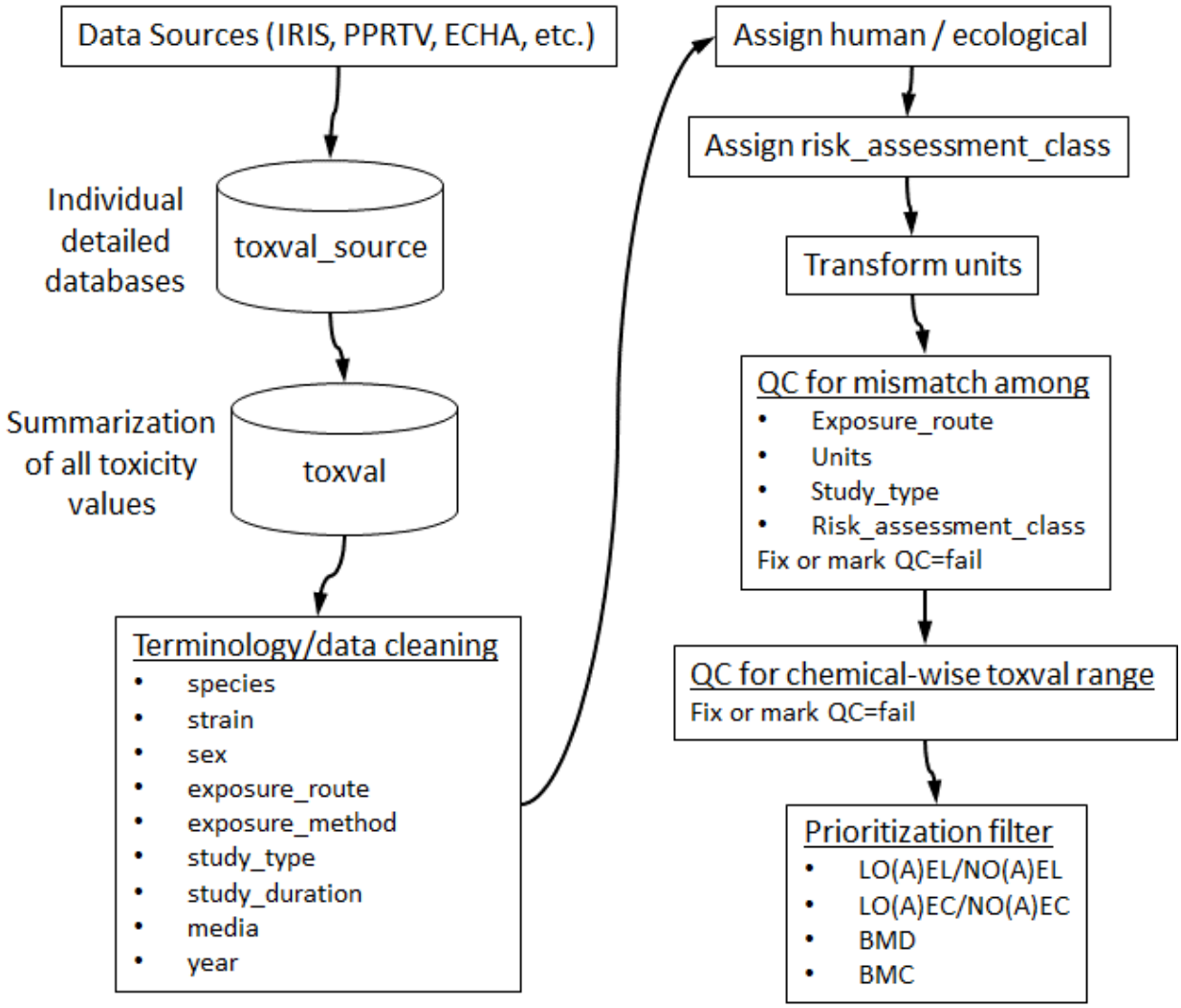


Figure A3.1: Schematic of ToxValDB process

References

ECHA. (2016). Topical Scientific Workshop on New Approach Methodologies in Regulatory Science: Background Document. Retrieved from https://echa.europa.eu/documents/10162/22816069/tsws_background_document_en.pdf/53102457-43f9-4709-8e4f-cf2b58458fbf

EPA. (2012). TSCA Work Plan Chemicals: Methods Document. Retrieved from https://www.epa.gov/sites/production/files/2014-03/documents/work_plan_methods_document_web_final.pdf

EPA. (2016). TSCA Work Plan: 2012 Scoring of Potential Candidate Chemicals Entering Step 2. Retrieved from https://www.epa.gov/sites/production/files/2016-08/documents/2012_workplan_step_2_chemicals-for_web-final.pdf

EPA. (2017). Safe Chemical Ingredients List. Retrieved from https://www.epa.gov/saferchoice/safer-ingredients

Filer, D. L., Kothiya, P., Setzer, R. W., Judson, R. S., & Martin, M. T. (2016). tcpl: the ToxCast pipeline for high-throughput screening data. *Bioinformatics*. doi:10.1093/bioinformatics/btw680

HC. (2016). Integrating New Approach Methodologies within the CMP: Identifying Priorities for Risk Assessment, Existing Substances Risk Assessment Program. Retrieved from http://www.ec.gc.ca/ese-ees/172614CE-D9F6-43DA-A528-B8175438210F/Objectives%20Paper%20November%202016_EN.pdf

Isaacs, K. K., Glen, W. G., Egeghy, P., Goldsmith, M. R., Smith, L., Vallero, D., . . . Ozkaynak, H. (2014). SHEDS-HT: an integrated probabilistic exposure model for prioritizing exposures to chemicals with near-field and dietary sources. *Environ Sci Technol, 48*(21), 12750-12759. doi:10.1021/es502513w

Judson, R. S., Kavlock, R. J., Setzer, R. W., Cohen Hubal, E. A., Martin, M. T., Knudsen, T. B., . . . Dix, D. J. (2011). Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. *Chem Res Toxicol, 24*(4), 451-462. doi:10.1021/tx100428e

Judson, R. S., Magpantay, F. M., Chickarmane, V., Haskell, C., Tania, N., Taylor, J., . . . Thomas, R. S. (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*(1), 137-154. doi:10.1093/toxsci/kfv168

Kleinstreuer, N. C., Ceger, P., Watt, E. D., Martin, M., Houck, K., Browne, P., . . . Judson, R. (2016). Development and Validation of a Computational Model for Androgen Receptor Activity. *Chem Res Toxicol*. doi:10.1021/acs.chemrestox.6b00347

Kleinstreuer, N. C., Ceger, P., Watt, E. D., Martin, M., Houck, K., Browne, P., . . . Judson, R. (2017). Development and Validation of a Computational Model for Androgen Receptor Activity. *Chem Res Toxicol, 30*(4), 946-964. doi:10.1021/acs.chemrestox.6b00347

Mansouri, K., Abdelaziz, A., Rybacka, A., Roncaglioni, A., Tropsha, A., Varnek, A., . . . Judson, R. S. (2016). CERAPP: Collaborative Estrogen Receptor Activity Prediction Project. *Environ Health Perspect, 124*(7), 1023-1033. doi:10.1289/ehp.1510267

Rotroff, D. M., Wetmore, B. A., Dix, D. J., Ferguson, S. S., Clewell, H. J., Houck, K. A., . . . Thomas, R. S. (2010). Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol Sci, 117*(2), 348-358. doi:10.1093/toxsci/kfq220

Shah, I., Liu, J., Judson, R. S., Thomas, R. S., & Patlewicz, G. (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. doi:10.1016/j.yrtph.2016.05.008

Truong, L. O., G., Pham, L. L., Clouzeau, J., Loisel-Joubert, S., Blanchet, D., Nocairi, H., . . . Martin, M. (2017). Building a Point-of-Departure Model: Predicting Study-Level Point-of-Departures of Systemic Effects using Chemical, Biological, Kinetic and Study Covariates. *Arch Toxicol, In Press*.

Wambaugh, J. F., Wang, A., Dionisio, K. L., Frame, A., Egeghy, P., Judson, R., & Setzer, R. W. (2014). High throughput heuristics for prioritizing human exposure to environmental chemicals. *Environ Sci Technol, 48*(21), 12760-12767. doi:10.1021/es503583j

Wambaugh, J. F., Wetmore, B. A., Pearce, R., Strope, C., Goldsmith, R., Sluka, J. P., . . . Setzer, R. W. (2015). Toxicokinetic Triage for Environmental Chemicals. *Toxicol Sci*. doi:10.1093/toxsci/kfv118

Wetmore, B. A., Allen, B., Clewell, H. J., 3rd, Parker, T., Wambaugh, J. F., Almond, L. M., . . . Thomas, R. S. (2014). Incorporating population variability and susceptible subpopulations into dosimetry for high-throughput toxicity testing. *Toxicol Sci, 142*(1), 210-224. doi:10.1093/toxsci/kfu169

Wetmore, B. A., Wambaugh, J. F., Allen, B., Ferguson, S. S., Sochaski, M. A., Setzer, R. W., . . . Andersen, M. E. (2015). Incorporating High-Throughput Exposure Predictions With Dosimetry-Adjusted In Vitro Bioactivity to Inform Chemical Toxicity Testing. *Toxicol Sci, 148*(1), 121-136. doi:10.1093/toxsci/kfv171

Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S., Li, L., Clewell, H. J., 3rd, Judson, R. S., . . . Thomas, R. S. (2013). Relative impact of incorporating pharmacokinetics on predicting in vivo hazard and mode of action from high-throughput in vitro toxicity assays. *Toxicol Sci, 132*(2), 327-346. doi:10.1093/toxsci/kft012

Zang, Q., Mansouri, K., Williams, A. J., Judson, R. S., Allen, D. G., Casey, W. M., & Kleinstreuer, N. C. (2017). In Silico Prediction of Physicochemical Properties of Environmental Chemicals Using Molecular Fingerprints and Machine Learning. *J Chem Inf Model, 57*(1), 36-49. doi:10.1021/acs.jcim.6b00625