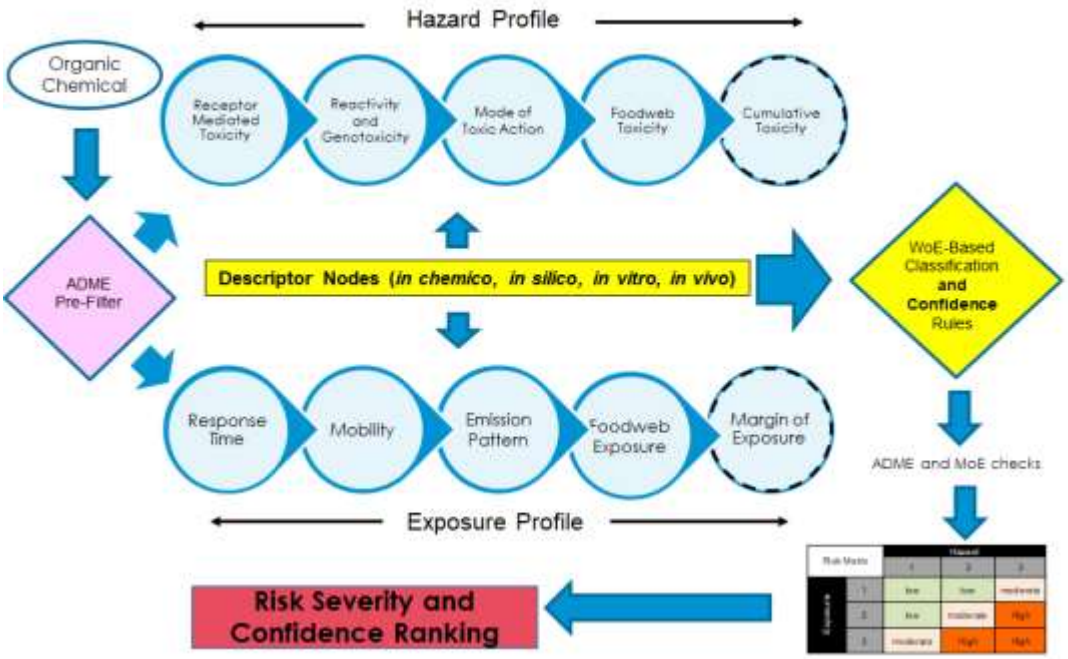
Science Approach Document

Version 2.0 of the Ecological Risk Classification Approach for Prioritizing Organic Substances on the Canadian Domestic Substances List that Did Not Meet Categorization Criteria



Ecological Assessment Division

Science and Technology Branch

Environment and Climate Change Canada

Cat. No.: En84-287/2022E-PDF

ISBN 978-0-660-42124-7

Acknowledgements

Version 2.0 of the Ecological Risk Classification of organic substances (ERC2) represents a considerable evolution from the first version published by Environment and Climate Change Canada (ECCC) in 2016 for the third phase of the Chemicals Management Plan (CMP3). ERC2 generates information for post-2020 risk-based priority setting and work planning of regulatory assessment activities. ERC2 was designed, formalized and computed by Mark Bonnell, Senior Science Advisor, Ecological Assessment Division (EAD), Science and Technology Branch. Several individuals and groups have contributed to computational and logical aspects of ERC2's design and functioning. ECCC wishes to acknowledge their very helpful support during the development of ERC2, which could not be completed without it.

- Alexander Okonski, EAD (computational support)
- Beate Escher, Swiss Federal Institute of Aquatic Science and Technology (mode of action)
- Chris Fraser, EAD (SciAD and IRAP support)
- Bryon Shore, EAD (QA-QC, computational and logic support)
- Carlie Lalone, US EPA Office of Research and Development (computational support)
- Danie Dube, formerly ECCC (Canadian Domestic Substances List tracking)
- Danny Lee, EAD (computational support)
- Detelina Dimitrova Laboratory of Mathematical Chemistry, Bulgaria (computational support)
- Drew MacDonald, EAD (computational support)
- Franklin Bauer, Kreatis, Inc. (computational support)
- Gail Bonnell, formerly ECCC (SciAD support)
- Hristiana Ivanova, Laboratory of Mathematical Chemistry, Bulgaria (computational support)
- Irina Dermen, Laboratory of Mathematical Chemistry, Bulgaria (computational support)
- James Armitage, Armitage Environmental Research, Inc. (computational support)
- James Anupol, EAD (SciAD support)
- Jerod Miksza, EAD (computational support)
- Jessy Kurias, EAD (computational support)
- Jon Arnot, Arnot Research and Consulting, Inc. (computational support)
- Kamel Mansouri, Integrated Laboratory Systems, Inc. (computational support)
- Kristin Isaacs, US EPA Office of Research and Development (computational support)

- Mace Baron, US EPA National Health and Environmental Effects Research (computational support)
- Marc Fernandez, EAD (UVCB support)
- Marisol Eggleton, EAD (SciAD consultation)
- Michael Beking, EAD (computational and logic support)
- Nikolai Nikolov, Danish DTU Food Agency (computational support)
- OECD QSAR Management Group and IATA Case Study Group (concepts)
- Ovanes Mekenyan, Laboratory of Mathematical Chemistry, Bulgaria (computational support)
- Paul Thomas, Kreatis Inc. (computational support)
- Sam Casey, US EPA Office of Research and Development (computational support)
- Todd Martin, US EPA Office of Pollution Prevention and Toxics (computational support)

This document has undergone external written peer review and consultation with government members from the Accelerating the Pace of Chemicals Risk Assessment (APCRA) program[1]. Comments on the technical portions of the document were received from Kellie Fay on behalf of the United States Environmental Protection Agency (US EPA) Office of Pollution Prevention and Toxics and Doris Hirmann on behalf of Evaluation Units from the European Chemicals Agency (ECHA). ECCC gratefully acknowledges the time and effort from both agencies to review the report. Comments from these agencies were taken into consideration for this draft release for public consultation, noting that the final content and outcome of the report remains the responsibility of Environment and Climate Change Canada.

[1] Accelerating the Pace of Chemical Risk Assessment (APCRA) | Safer Chemicals Research | US EPA

SYNOPSIS

In 2016, Environment and Climate Change Canada (ECCC) developed and published a 21st-century science approach to re-prioritizing 640 organic chemicals from the third phase of the Chemicals Management Plan (CMP). Known as the Ecological Risk Classification of organic substances (ERC) approach, this first version of ERC (ERC1) was applied to organic substances that met criteria for further risk assessment following the Government of Canada's categorization of the *Domestic Substances List* (DSL) in 2006, performed under the *Canadian Environmental Protection Act*, *1999* (CEPA).

Since then, ECCC has refined its approach for the ERC, which led to the creation of this second version (ERC2). This Science Approach Document presents the ERC2 approach and the results of its application to approximately 12 200 organic substances from the DSL that did not meet criteria for further risk assessment following the 2006 categorization exercise. The ERC2 approach builds on ERC1 and refines key areas of uncertainty previously identified in ERC1 through better integration and transparency of weight of evidence concepts, expansion of the toxicological and exposure 'space' used for hazard and exposure, increased consideration of model domain of applicability, updates to the decision logic governing classification rules based on new tools and lessons learned from ERC1 and the restructure of these rules, an increased focus on long-term developmental and reproductive toxicity and neurotoxicity, and the introduction of confidence and severity scoring for exposure, hazard and risk classification outcomes.

ERC2 is considered a high-throughput integrated approach to testing assessment (IATA) method, as it uses many sources of 'alternative data' (also known as new approach methodologies or NAM) such as *in silico*, *in chemico*, and *in vitro* data to complement traditional *in vivo* sources and to provide evidence for risk classification. The ERC2 approach gathers multiple lines of evidence in chemical profiles for hazard (toxicity) and exposure. The evidence is compared to logic rules established for hazard, exposure and risk classification to determine if future assessment activities are merited. The details for all classification rules in ERC2 are provided in appendices of the science approach document, while the core science behind ERC2, contained in profile descriptors, is provided in the main body of the document. Each ERC2 classification is also associated with a confidence and severity score. The confidence score is a measure of the consensus of data and data availability. The severity score is a measure of hazard, exposure or risk scale and is used as a means of weighting various classification outcomes in ERC2 when considering possible future assessment activities.

Examples of risk assessment outcomes and targeted assessment activities (e.g., identification of endocrine active substances, potential for cumulative assessment, and regrettable alternatives) are provided to demonstrate how results from the ERC2 approach may be applied by ECCC for future work planning. Risk classification, confidence and severity results for ~12 200 substances on the DSL are contained in a Microsoft Excel® spreadsheet as a supporting document to this Science Approach Document.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

1.1 ERC version 1.0

In 2016, Environment and Climate Change Canada (ECCC) successfully developed and published a 21st-century science approach to re-prioritizing 640 organic chemicals from the third phase of the Chemicals Management Plan (CMP). Known as the Ecological Risk Classification of organic substances (ERC) approach (ECCC 2016), this first version of ERC (ERC1) was applied to organic substances that had been classified as persistent or bioaccumulative and inherently toxic (PiT or BiT) criteria during the Government of Canada's 2006 categorization. Categorization was a prioritization exercise under the *Canadian Environmental Protection Act,* 1999 (CEPA), which identified 4300 of approximately 23 000 substances on the Domestic Substances List (DSL) as priorities for further assessment under the CMP. Canada had previously assessed all of the substances categorized as being persistent, bioaccumulative and inherently toxic (PBiT) during the CMP's 2006-2012 Challenge initiative.

ERC1 is a risk-based approach that used multiple chemical descriptors (multiple lines of evidence) to establish both hazard (potency) and exposure profiles for individual chemicals. The profiles were compared to weighted hazard and exposure classification rules in order to determine an overall risk classification of an organic chemical substance. The approach allowed Canada to further its 2020 commitment to SAICM[2] by refocusing how organic chemical priorities are identified for further risk assessment using the best available and up-to-date science. By doing so, ERC1 better defined chemicals of higher ecological concern for CMP3 and demonstrated that a weight-of-evidence approach was viable and indeed preferable for organic chemical prioritization.

ERC1 was reviewed and published by the Organisation for Economic Co-operation and Development (OECD) Integrated Approaches to Testing and Assessment (IATA) Case studies initiative under the third review cycle (OECD 2018). Figure 1 from ECCC's IATA Case study (OECD 2017) illustrates the logic flow to substance classification using ERC1. More details on ERC1 can be found in ECCC (2016). The impact of re-prioritizing the remaining CMP organic substances using ERC1 was significant. Not only did fewer substances require further ecological risk assessment (~80% less), it resulted in much greater parity (~55%) with organic priorities identified by Health Canada for further human health risk assessment. Ultimately, ERC1 provided the Government of Canada with a more efficient allocation of time and resources for chemicals management by targeting chemicals of concern using a modern 21st-century science approach.

[2] Strategic Approach to International Chemicals Management

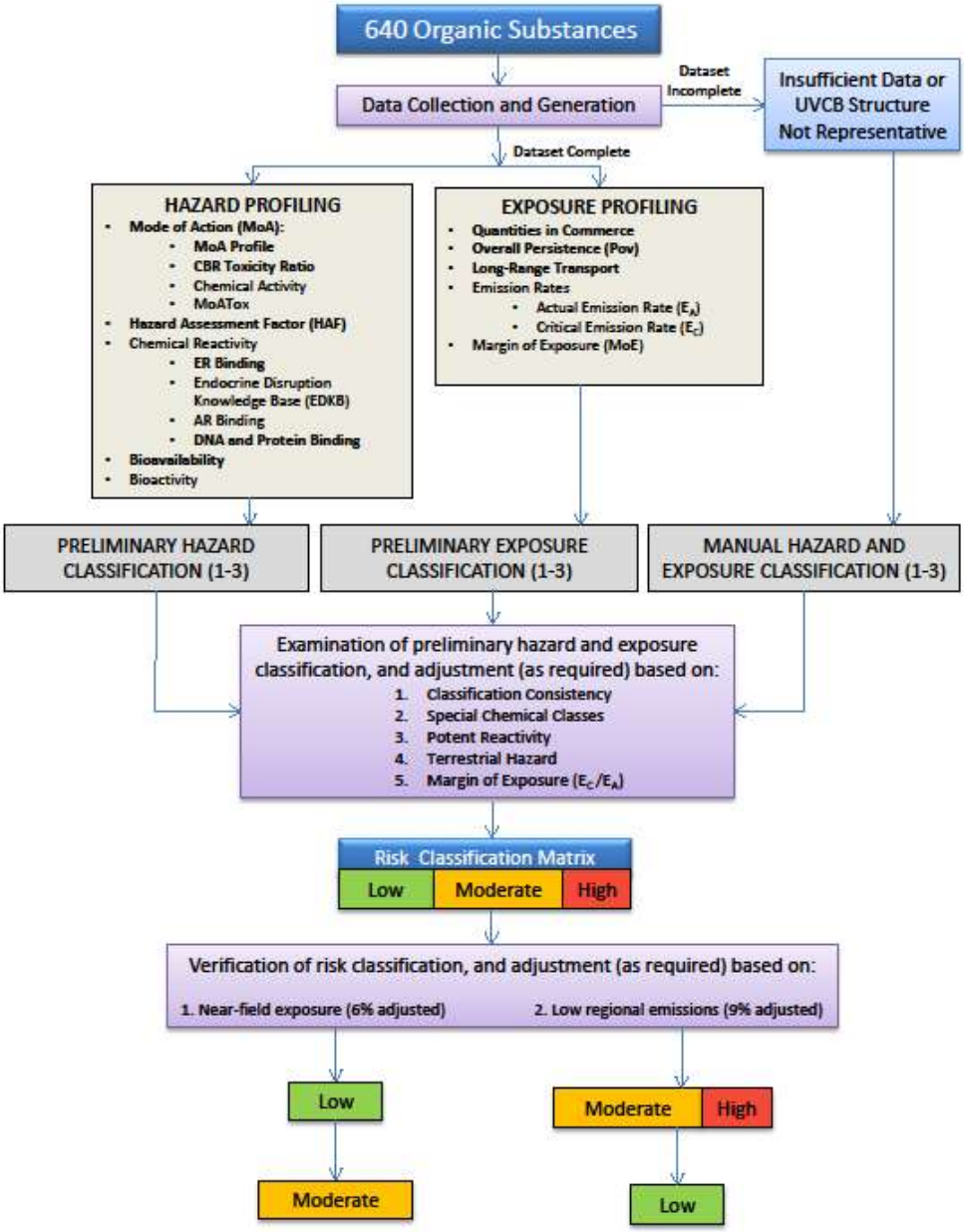


Figure 1: ERC1 Logic Flow Leading to Risk Classification

1.2 Introduction to ERC version 2.0

Based on the successful implementation of ERC1 and lessons learned in the process, the development of version 2.0 of the ERC approach (ERC2) began in 2017 and was completed mid-2019. ERC1 provided ECCC with a proof of concept and could therefore be used as a template to increase the sophistication of the ERC approach and incorporate new sources of information and tools that had become available

since ERC1. ERC2 was then developed by ECCC for the purpose of re-examining ~12 000 organic chemicals on the DSL that had not met the 2006 categorization criteria for persistence (P), bioaccumulation (B) and inherent toxicity (iT). Similar to ERC1, ERC2 applies to discrete organics and organic UVCBs[3] with known structure(s), but also includes the organic moieties of organic-metal salts identified as inorganic priorities under the CMP. ERC2 integrates alternative non-animal data with traditional animal testing data and is considered an "alternative approach to testing and assessment" (IATA) prioritization model. Ecological prioritization results from ERC2 provides the Government of Canada with information for post-2020 chemicals management priority-setting and planning for organic substances. Considering the CMP's current Identification of Risk Assessment Priorities (IRAP) approach, ERC2 is considered an emerging science and monitoring feeder to IRAP, as both a data source and a candidate identification mechanism (Figure 2).

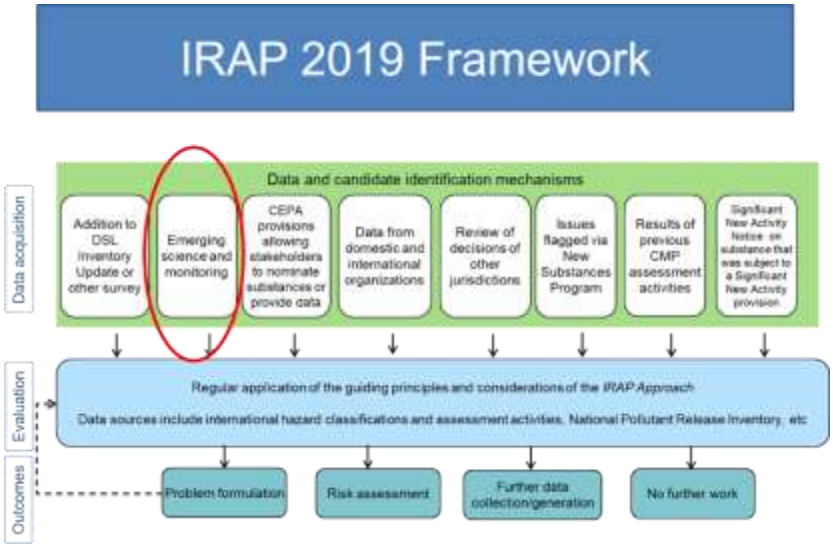


Figure 2: Identification of Risk Assessment Priorities (IRAP) process showing where ERC2 provides information via the emerging science and monitoring feeder

While the primary use of ERC2 outcomes will to be inform the Government of Canada's post-2020 work planning on the ~12 200 organic substances (e.g., identification and further scoping of potential priorities, risk assessment, identifying regrettable alternatives, determining targeted research and monitoring needs, etc.), the approach can also be applied to other organic substances not included in ERC2. For example, the ERC2 approach is also used to coherently generate data and knowledge to inform the problem formulation of alternatives and analogues of bisphenol A (BPA) a chemical group

[3] UVCBs: substances of Unknown or Variable composition, Complex reaction products or Biological materials

beyond the ~12 200 identified as a potential priority via the CMP's IRAP approach (ECCC, HC 2019). ERC2 builds on the concepts of ERC1, but expands these concepts in specific ways, including:

- Refinement of key areas of uncertainty previously identified by ECCC in ERC1 and by OECD members during the OECD IATA 3rd Cycle Case study review
- Better integration and transparency of weight of evidence concepts also identified by OECD members during the OECD IATA 3rd Cycle Case study review
- Expansion of the toxicological and exposure "space" used for hazard and exposure profiles based on the use of additional empirical data and new *in silico* tools
- Increased consideration of model domain of applicability
- Updating and restructuring of the decision logic governing classification rules based on new tools and lessons learned from ERC1
- The introduction of confidence and "severity" scoring
- Expanded toxicokinetics (i.e., absorption, distribution, metabolism, elimination [ADME]) considerations
- An increased focus on long-term developmental and reproductive toxicity and neurotoxicity

The following two sections describes these enhancements and the logic flow of ERC2 in more detail.

2. Core Concepts of the ERC2 Approach

2.1 Use of Computational Methods

ERC2 includes over 10 million data points for various chemical properties and endpoints for ~12 200 organic chemicals listed on the Canadian DSL. Given the high number of organic chemicals evaluated using the approach, computational methods have been an essential aspect of forming the evidence required for substance classification in ERC2. As a result of the lack of experimental data for the vast majority of industrial chemicals, over 90% of the data in ERC2 has been generated using *in silico* tools and therefore ERC2 can be considered to be mainly a computational approach. The computational tools used for generating data in ERC2 are listed in Appendix I.

ERC2 can also be considered a high-throughput IATA method, as it uses many sources of "alternative data" (also known as new approach methodologies or NAM) such as *in silico*, *in chemico*, and *in vitro* data to complement traditional *in vivo* sources. For example, most of the data for the exposure descriptors have been generated from *in silico* tools. The ERC2 logic rules, data and results are contained in Microsoft Excel®, which was chosen for practical reasons (e.g., data sharing and publication, user accessibility and interfacing). Computational *in silico* data contained in the ERC2 Excel database have been pre-generated for many of the chemicals contained in ERC2. This was performed outside of the ERC2 database. As such, ERC2 is not yet a fully automated method, although new organic chemicals can

be manually added to and prioritized in ERC2 at any time. Finally, as ERC2 employs many *in silico* tools to fill data gaps, care was taken to examine model domain of applicability as well as extreme values of chemical properties used as input data to various models (e.g., "out of domain" is reported). Model domain of applicability and extreme chemical property flags are described in the *in silico* classification and confidence rules for individual descriptors in ERC2 in Appendices II-IV. Finally, computational or empirical data that could not be generated or were not available in databases are shown as "NA" in the ERC2 result tables in this document, but in the Microsoft Excel® table, results (Appendix XII) are left blank for computational reasons.

2.2 Weight of Evidence and Acceptable Level of Uncertainty

2.2.1 Weight of Evidence Approach

ERC2 was designed to provide outcomes to guide further regulatory decision-making. ECCC designed ERC2 as a weight of evidence (WoE) prioritization approach, keeping in mind the WoE principles and elements contained in OECD (2019). The following concept of WoE described in OECD (2019) captures quite accurately the ERC2 WoE approach:

> *"Conceptually WoE can be seen as a method for decision-making that involves consideration of known lines of evidence (LoEs) where a "weight" is assigned to each LoE, according to its relevance and reliability. A conclusion can be reached by combining the various LoEs to determine if sufficient strength of evidence is available to address the question posed under the hypothesis (e.g., a molecular initiating event will lead to an adverse outcome)."*

As a consensus-based computational IATA, ERC2 was designed to collect and generate information for key lines of evidence in a coherent manner. For example, the adverse outcome pathway (AOP) construct is used as a means for determining plausible mechanisms to support adverse outcomes that are used for hazard classification, thereby increasing the strength of evidence for classification. OECD (2019) also outlines five guiding principles and several key elements that a WoE approach should contain. The guiding principles suggest that a WoE approach should:

1. Include a hypothesis which involves a clear formulation and statement of the problem for which evidence is needed and possible alternative hypotheses
2. Be systematic and comprehensive in design by documenting a step-wise procedure integrating all evidence and indicating how evidence was collected, evaluated and weighed
3. Include a treatment of uncertainty arising from available data (knowns) and data and/or knowledge gaps (unknowns)
4. Consider the potential for bias during collection, evaluation and weighing of evidence
5. Be transparent by including clear documentation to assist the communication of WoE decisions so that they can be understood, reproduced, supported or questioned by all interested parties.

ERC2 was also designed with the OECD's WoE key elements in mind. Figure 3 shows how ERC2 uses these elements for decision-making.

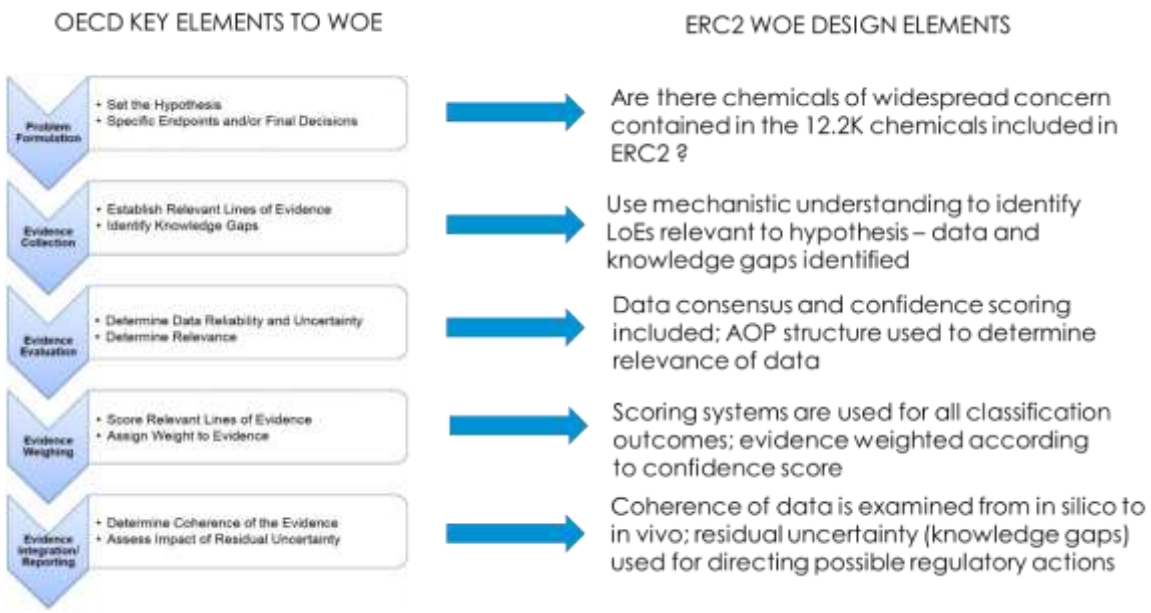


Figure 3: Key elements of WoE from OECD (2019) showing how ERC2 uses these elements

Ultimately, the goal of designing a prioritization approach with the above OECD WoE principles and elements in mind is to provide a coherent evidence-based story line that is transparent and traceable when communicating ERC2 outcomes to all stakeholders.

2.2.2 Acceptable Level of Uncertainty

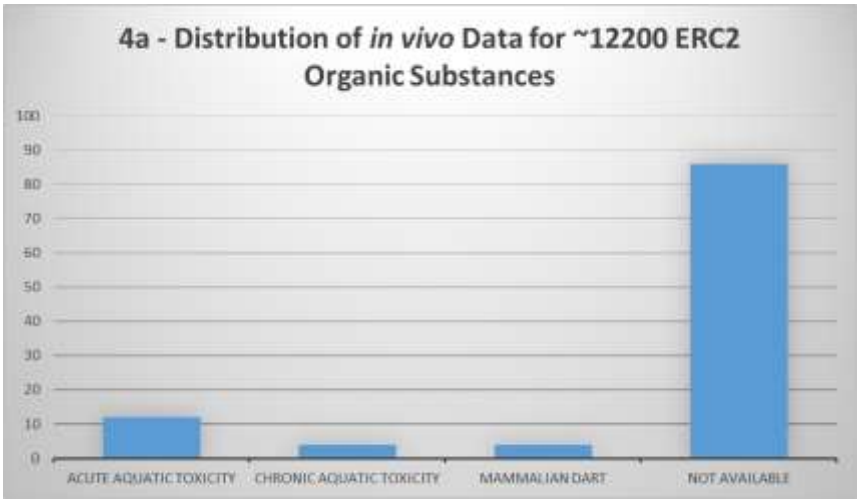
Although designed as a tool for prioritizing organic substances, ERC2 can itself be considered a problem formulation for a single chemical or chemical group. During the problem formulation for risk assessment or when laying out a rationale for further ecological evaluation in, for example, a substance profile summary, a statement of the acceptable or tolerable level of uncertainty can be made (OECD 2019). This is important because the level of confidence required to accept or reject a hypothesis is dependent on regulatory acceptance of the level of uncertainty. The level of uncertainty is in turn related to the regulatory context, which for ERC2, is prioritization.

In the context of prioritization for the CMP, a wider margin of uncertainty is accepted in comparison to making a risk-based conclusion under CEPA toxicity as part of a regulatory assessment. An IRAP review and further scoping of potential priorities for publishing a substance profile do not typically involve a

deep level of data investigation compared with assessment. Nonetheless, confidence scoring is integrated into ERC2, allowing the level of acceptable uncertainty to be selected according to the question asked (hypothesis) and the decision-context (e.g., prioritization or risk assessment). For example, it may be desirable to have a higher ERC2 confidence score for those chemicals classified as low concern by ERC2 rules to avoid both false positive and negative outcomes. Additionally, ERC2 was designed to protect the multiple ecological receptors expected to come into contact with an anthropogenic organic substance released to the environment. Conservative assumptions are used in ecological assessment to extrapolate across the myriad of species in the environment to bring a broad-based protection goal for the environment. Consequently, in an ecological assessment there is low species specificity and inherently a higher level of uncertainty in comparison to a human health assessment, where there is a high degree of species specificity and generally a lower level of acceptable uncertainty even at the prioritization stage.

2.3 Biological extrapolation and the adverse outcome pathway concept

There is a paucity of hazard data for ecological receptors, particularly at *in vitro* and *in vivo* levels. For example, Figure 4 shows the relative distribution of available *in vivo* data (Figure 4a) and *in vitro* data (Figure 4b) for the ~12 200 substances in ERC2, noting that a single substance can have more than one type of *in vivo* or *in vitro* data. A result of the high-throughput *in vitro* TOXCAST program from the US EPA is that there is considerably more *in vitro* data than *in vivo* data (18%), but overall 81% of substances in ERC2 have no laboratory tested biological data.



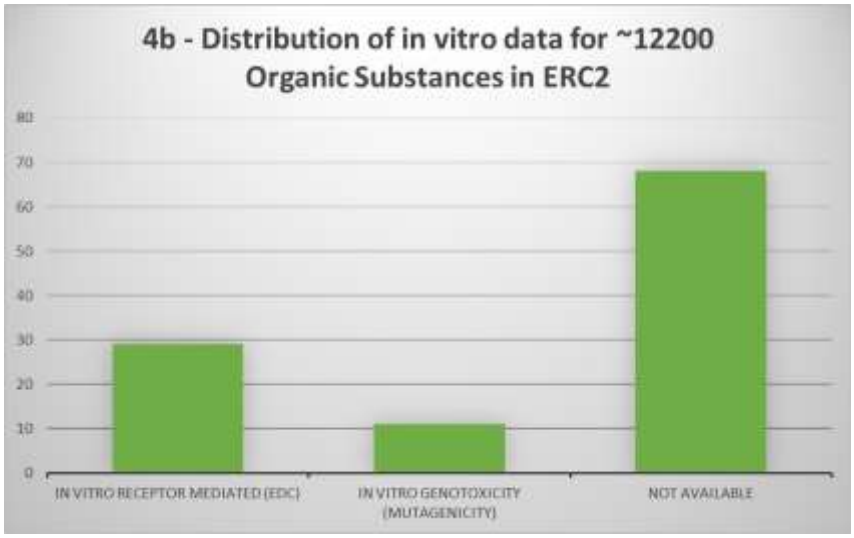


Figure 4a,b: Percentage distribution of laboratory tested biological data for the ~12 200 substances in ERC2.

The above data figures include data that until recently would traditionally be considered in the human health domain (e.g., estrogen receptor binding *in vitro* data, mammalian developmental and reproductive toxicity). As a consequence of the lack of experimental aquatic toxicity data for the majority of substances in ERC2, biological read-across or cross-species extrapolation is used to fill data gaps for mammals and other receptors. This is performed with the understanding that many xenobiotic interactions and pathways are conserved across species (e.g., the use of mammalian estrogen receptor binding data for aquatic receptors) (Ankley and Gray 2013). Common susceptibility between mammalian and aquatic receptors to chemical toxicity for various modes of action has also been explained using chemical activity (Mackay et al. 2012). ERC2 thus promotes the notion of "one toxicology", where cross-species susceptibility can be further examined for specific interactions using, for example, US EPA's SeqAPASS (Sequence Alignment to Predict Across Species Susceptibility) tool (Lalone et al. 2016). In ERC2, SeqAPASS may be used to verify the degree of cross-species susceptibility to specific interactions (estrogen receptor, thyroid inhibition) by checking the commonality of amino acid sequences between species and providing a score to indicate strong or poor species susceptibility relationships.

More recently, Sapounidou et al. (2021) collected an extensive array of mechanistic toxicity evidence across species to demonstrate the taxonomic applicability of plausible mechanisms leading to adverse effects. A main conclusion from this work is that many mechanistic interactions are conserved across environmental species, particularly vertebrates, with some specific mechanisms and molecular initiating events (MIEs) limited to aquatic plants (algae). Given a need for extrapolation across species, it is therefore necessary to establish plausibility (mechanistic)-causality (adverse outcome) relationships to achieve higher confidence for hazard classification. ERC2 has therefore employed the adverse outcome pathways (AOP) concept (Figure 5) (used with permission from the OECD; after Ankley et al. 2010) as a useful method for organizing toxicological data for prioritization in order to satisfy the plausibility-

causality relationship aspect of the WoE requirements (Figure 5). Using the AOP structure also allows substances in ERC2 to be grouped according to common chemical interactions with biological tissues or so-called MIEs that could potentially lead to adverse outcomes. However, it is strongly emphasized that using the AOP structure in ERC2 to organize toxicological data does not provide complete nor detailed AOP information for each substance. Rather, the AOP structure is primarily used to organize toxicological information such that it provides possible linkages between mechanistic and adverse outcome data. An adverse outcome, or lack thereof, may or may not be explained by the mechanistic interactions linked to it as suggested in ERC2. Many molecular interactions leading to adverse outcomes are possible; ERC2 simply suggests that linked interactions are plausible and therefore the adverse outcome are explainable. One of the more important aspects of employing the AOP structure in ERC2 is that confidence scoring for hazard classification can be directly linked to available toxicological data at various biological levels (i.e., MIEs to whole organism responses) and used to examine the coherence of the data. This is explained further under confidence scoring for hazard in section 4.1.

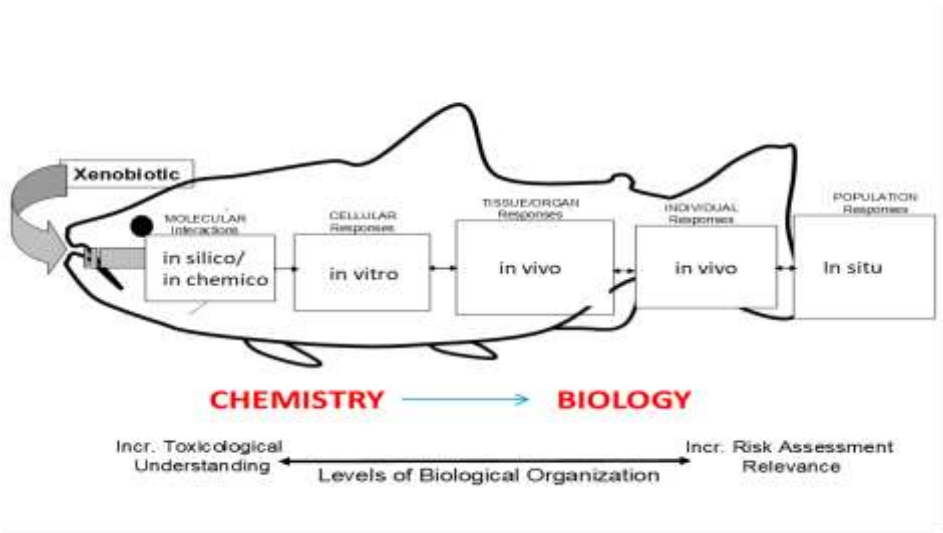


Figure 5: ERC2 hazard data cascade for relating key mechanistic interactions adverse outcomes using the adverse outcome pathway (AOP) structure

2.4 Chemicals of Concern in 1995 vs. 2021

When prioritization of organic chemicals under the Toxic Substances Management Policy (TSMP) first began in Canada in 1995, chemical persistence (P) and bioaccumulation (B) criteria were developed to identify chemicals considered to be of higher concern based on the TSMP[4]. Considering the domain of chemicals used to define the policy at the time, the TSMP was primarily aimed at identifying chemicals with "POP-like" behaviour, those capable of eliciting unpredictable exposures and effects to species in

[4] Toxic substances management policy - Canada.ca

remote areas over a long time period. These chemicals (listed in Appendix I of the TSMP) are mainly comprised of neutral organics, many of which are highly halogenated and found to be mobile in air or water or both. Application of the P and B criteria is precautionary by design and meant to address the uncertainty of exposure in the far field over time. When the *Canadian Environmental Protection Act*, 1999 (CEPA) (Canada 1999) mandated the categorization of all substances on the DSL by 2006, the TSMP P and B criteria were legally formalized as the *Persistence and Bioaccumulation Regulations* (Canada 2000) and used in combination with non-legal criteria for "inherently toxicity to non-human organisms" (iT) to identify ecological priorities for risk assessment.

Many of the POP-like chemicals have been identified globally and are being managed, although many regulatory programs continue to prioritize these chemicals such as the "Substances of Very High Concern" (SVHC) under the European REACH[5] as well as via United Nations Persistent Organics Pollutants Review Committee (POPRC)[6]. In Canada, use of the criteria is primarily performed to determine the applicability of risk management measures required for a substance under Track 1 of the TSMP or Schedule 1 of CEPA. While efforts continue to identify new PBT and POP-like substances in chemical inventories and the environment (Muir et al. 2006), more recent studies use or suggest other metrics or descriptors to identify chemicals of concern (CoCs), such as persistence, mobility and toxicity (PMT) (Matthies et al. 2016; McLachlan et al. 2014; Reppas-Chrysovitsinos et al. 2019; Rüdel et al. 2020) or simply just high persistence as being sufficient for regulating and managing perfluorinated acids (Cousins et al. 2020).

Persistence, bioaccumulation and toxicity are important properties of chemical behaviour that can be predictive of adverse effects in the environment. However, substances that meet PBT or P&B criteria represent a small fraction of a chemical inventory. For example, in Canada, approximately 400 PBiTs were identified during categorization in 2006. Given there were approximately 23 000 substances in categorization, this represents <2% of the Canadian DSL organics and inorganics. When subsequent ecological risk assessments were conducted on ~100 organic PBiTs under the CMP's Challenge initiative (the other ~300 substances had already been assessed and managed in Canada and did not require further risk assessment), only ~5% of the substances resulted in a risk-based CEPA Toxic conclusion for the environment. This raises the question whether after more than 20 years since categorization efforts began, a hazard-based approach is truly an effective means for identifying CoCs under CEPA, a risk-based Act for chemicals management. Additionally, it must be questioned whether there are other CoCs on the DSL that are potentially of high concern for the environment, but do not have a PBiT profile yet.

These are important questions that ERC2 was designed to address because many industrial chemical inventories contain substances that are "drug-like or "pesticide-like" in their fate and behaviour. Chemicals of this nature were designed to be highly potent and consequently have a very narrow margin of exposure in the environment, meaning that exposures to very low concentrations (e.g., low ng/L) in the environment can lead to adverse effects. These substances typically do not meet current regulatory

[5] https://echa.europa.eu/candidate-list-table

[6] http://chm.pops.int/TheConvention/POPsReviewCommittee/OverviewandMandate/tabid/2806/Default.aspx

hazard criteria and many are “ionogenic” meaning ionized at typical environmental pH. Importantly, some of these substances can be endocrine active leading to endocrine disruption and ultimately developmental and reproductive effects. Substances of this type are on industrial inventories because they can have multiple use patterns that cross over into the industrial chemical realm. One recent example from the third phase of the CMP is the risk assessment of Dinoseb used for industrial chemical processing (ECCC, HC 2019), a substance banned by many agencies when used as a pesticide.

One of the benefits of the ERC approach (v1 and v2) is that it is risk-based, thus allowing both hazard and exposure potential to be evaluated at the prioritization step. The ERC approaches incorporate descriptors to identify highly potent CoCs at various spatial and temporal scales of exposure, with one of its goals to target substances that have the potential to affect the *planetary boundary*[7]. The planetary boundary concept was first described by Rockstrom et al. (2009) and later extended to identifying chemicals that pose a threat to the planetary boundary by MacLeod et al. (2014). This work then lead to recommendations for technological and societal changes to avoid global chemical pollution by Diamond et al. (2015) and an approach for prioritizing chemicals based on this concept by Reppas-Chrysovitsinos et al. (2019). ERC2 examines a specific aspect of the planetary boundary chemical threats described by situations C2-2 and C3-2 in MacLeod et al. (2014) and illustrated here as Figure 6 (reproduced with permission from the author). This scenario describes a threat from persistent and mobile substances (and their transformation products) that have a hazard potential capable of causing permanent genetic damage in exposed organisms, the effects of which may also be expressed in subsequent generations (i.e., epigenetic inheritance). This type of effect is known to be “irreversible or poorly reversible” and may result in a “regime shift” in populations (MacLeod et al. 2014). The impact of this type of chemical profile is that local and global effects continue after emissions have ceased due to regulatory or voluntary actions (dashed lines in Figure 6). In other words, *effects are distributed widely across populations and are <u>independent</u> of global or local concentrations* (MacLeod et al. 2014). Concentrations in the far field, normally quite diluted compared to near field or local concentrations, become highly relevant because exposure at very low concentrations can result in adverse effects (i.e., very narrow margin of exposure). ERC2 prioritizes substances of this type as potential “threats to the planetary boundary” where the risk can scale from the near field to the far field.

[7] The planetary boundary refers to a boundary which delimits “the safe operating space for humanity” and by extension the environment. Anthropogenic activities, many of which are related to chemical pollution, can lead to impacts at a planetary scale and threaten the vital earth system processes that allow humanity to continue to exist (Rockstrom et al. 2009).

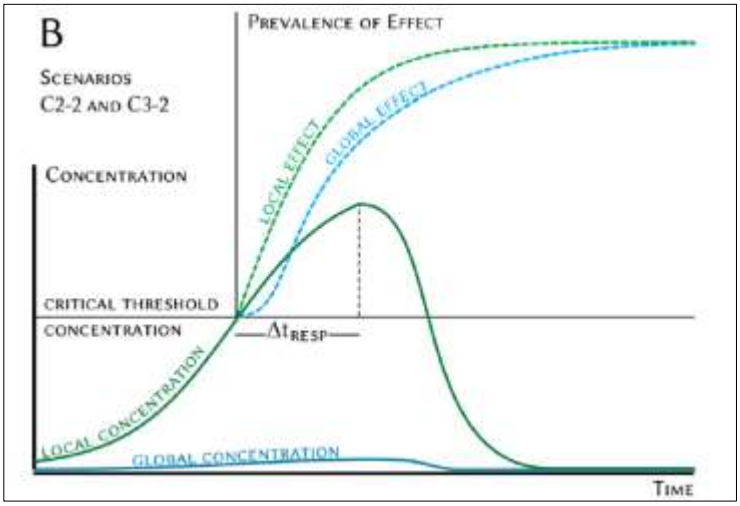


Figure 6: Scenarios C2-2 and C3-2 from MacLeod et al. (2014) describing the rapid global distribution of effects (C2-2) that are poorly reversible (C3-2)

3. ERC2 Workflow and Logic

The ERC approaches were designed using a chemical profiling concept. In this approach, evidence is gathered in chemical profiles for hazard (toxicity) and exposure and used to compare with logic rules established for risk classification. In addition to risk classification, ERC2 adds confidence and severity scoring for final outcomes of hazard, exposure and risk. Figure 7 describes the basic workflow of ERC2 that is conducted for all substances with a valid Chemical Abstract Services registry number (CAS RN) and a chemical structure. Following Figure 7, the chemical identity of a substance is first determined (section 3.1), followed by manual classification of risk for those UVCBs without an acceptable representative structure (section 3.2). UVCBs with acceptable representative structures are treated similarly to discrete organic substances, noting that this is a practical solution for dealing with the uncertainty and general lack of available data for UVCB component distributions. Physical-chemical property data are then generated for all discrete organic substances and representative structures of UVCBs. This information is later used as input for specific hazard or exposure descriptors. Following this, substances are filtered for specific aspects of ADME (absorption, distribution, metabolism and elimination) (section 3.3). Information contained in the hazard and exposure profiles (called descriptor data) is compared to pre-established logic rules for classifying hazard and exposure (Appendices II-X) and subsequently risk using a risk matrix approach (section 4.0). Exposure classification outcomes of low concern are checked for adequate margins of exposure before final exposure classification. Chemical profile information is also compared to logic rules for scoring confidence and severity for hazard, exposure and risk (sections 4.0 and 5.0). The following sections provide details for the generalized workflow in Figure 7.

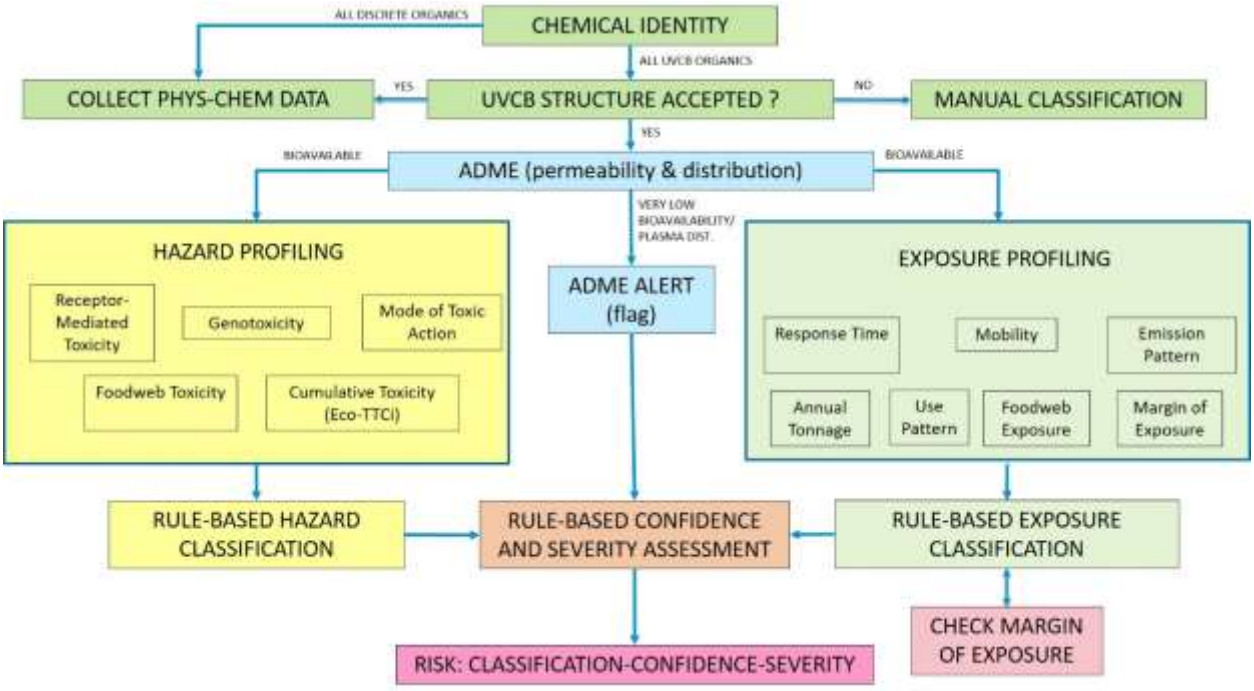


Figure 7: Generalized workflow for ERC2 leading to substance risk classification and scoring of confidence and severity

3.1. Chemical Identity

ERC2 contains 13 162 organic chemicals identified by CAS RN, chemical name and, if available, SMILES (Simple Molecular-Input Line-Entry System), but results could only be generated for ~12 200 of these chemicals. This total also includes the organic counterions from all organic metal salts as well as all organic counterions from metal salts that met the categorization criteria (but were later assessed based on the inorganic moiety) from phases one to three of the CMP. The organic counterions, mostly anionic, were added to ERC2 under the assumption that complete dissociation of the metal salt occurs in the environment, thus exposing organisms to the organic moiety. Substances with multiple counterions required selection of a representative ionic moiety. This was performed manually using expert judgement and consideration of basic physical-chemical properties to determine a bioavailable realistic worst-case structure. Figure 8 shows a pie distribution of substance types contained in ERC2 and Figure 9 shows the percent distribution of ionogenic substances and neutral substances which is similar to other inventories such as REACH (Franco and Trapp 2010).

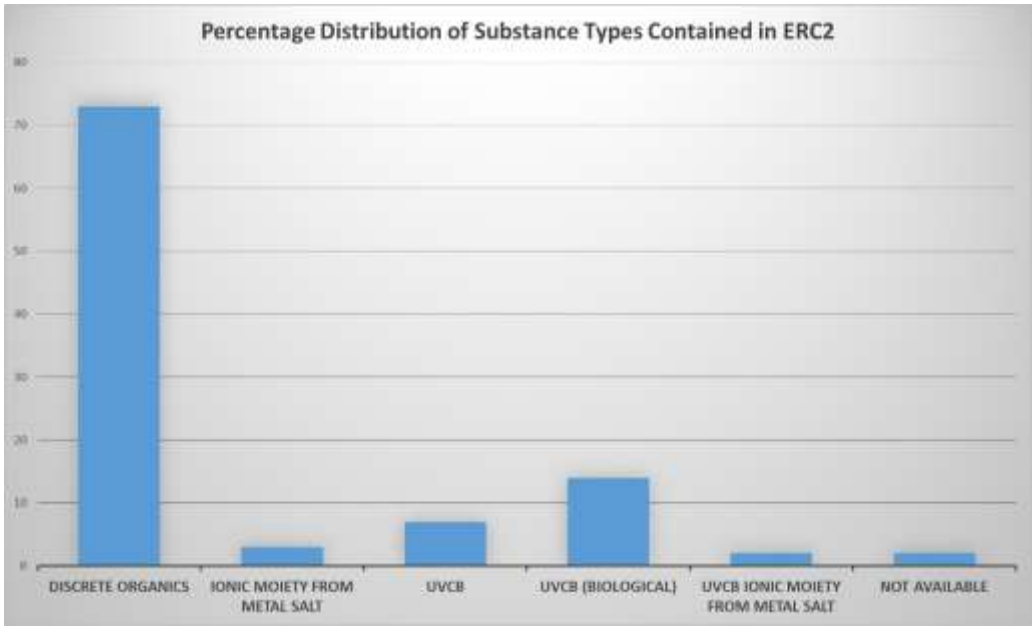


Figure 8: Percentage distribution of substance types considered for ERC2

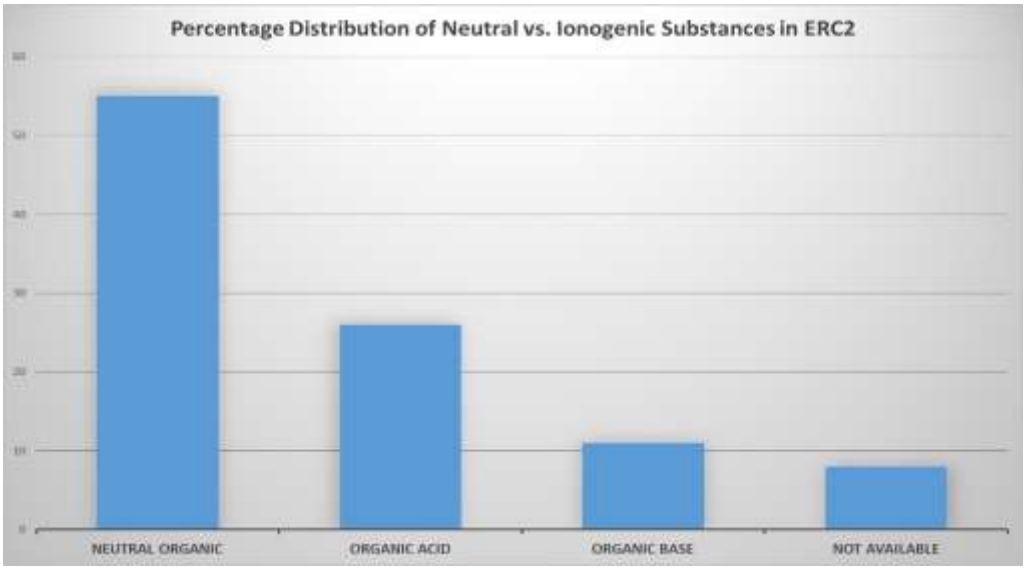


Figure 9: Distribution of neutral vs. ionogenic (ionized) substances considered in ERC2

SMILES were available for ~12 200 of the ERC2 substances, the majority of which were generated and curated during the 2006 categorization of the DSL (Government of Canada 2017). This list of SMILES included a representative structure for less complex UVCBs, selected as a "worst-case" representative from a PBT perspective at the time of categorization. There were 959 of ~12 200 substances with no available SMILES, predominantly for very complex organic UVCBs. These were excluded from ERC2 and

will require a separate prioritization activity (section 3.2). The remaining ~12 200 structures therefore contained discrete neutral and ionizing organic substances and organic UVCBs with a representative structure. These structures were further curated to allow application of various *in silico* tools. Positive and negatively charged counterions of organic salts (e.g., Na^+, Ca^{2+}, K^+, Li^+, $NH4^+$, simple organic acids, etc.) were removed so that the remaining organic structure represents the moiety that may be of concern in the environment from dissociation of the salt.

The ~12 200 substances were entered as a new inventory in the OECD QSAR Toolbox v4.4.1 using the CAS RN, name and SMILES generated during DSL categorization in 2006. The OECD QSAR Toolbox allows for cross-validation of the CAS RN-SMILES relationship and provides canonical (unique) SMILES as output. The Toolbox canonical SMILES were used for further ERC2 *in silico* data generation. Figure 10 shows the outcomes of the CAS RN-SMILES relationship verification for substances used for further modelling. The results of this verification show that most (~ two-thirds) CAS RN-SMILES relationships are of high quality when a relationship can be formed (i.e., did not result in a "not available" outcome). However, almost one quarter of the substances in ERC2 have low quality CAS RN-SMILES relationships. Just under half of these are UVCBs for which a representative structure was selected and will not by default have a high quality CAS RN-SMILES relationship due to the multicomponent nature of UVCBs. The remaining low quality outcomes are largely a result of the de-salting of ERC2 organic salts which will not represent the CAS RN structures in the OECD QSAR Toolbox.

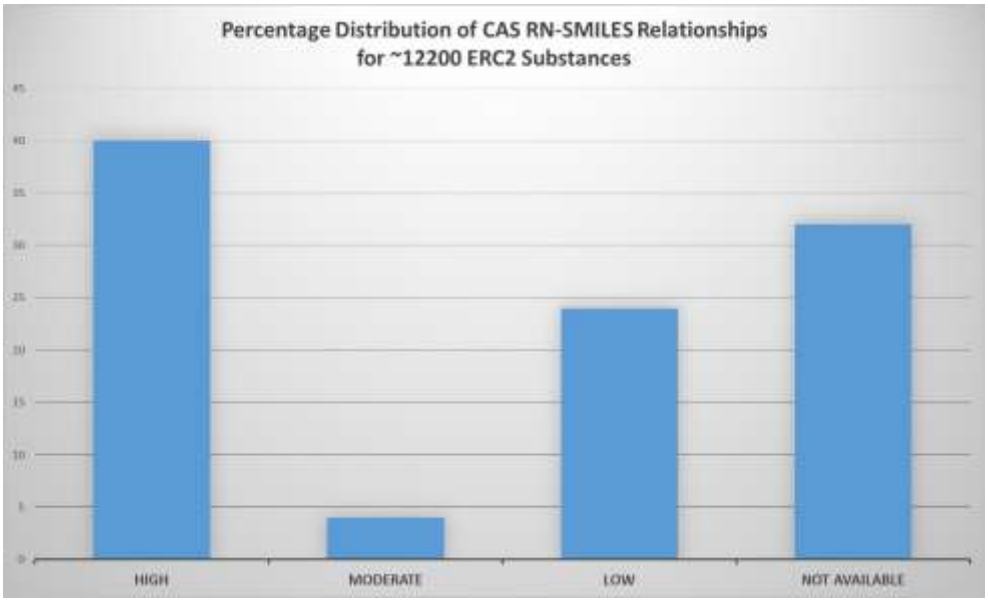


Figure 10: Results of the OECD Toolbox v4.4.1 CAS RN-SMILES relationship quality for ~12 200 ERC substances

3.2 Classification of UVCBs with Representative Structures

Approximately 23% of the chemical substances covered by ERC2 are substances of unknown or variable composition, complex reaction products, or biological materials (UVCBs) which pose unique risk challenges to regulators (Figure 8). UVCBs are regarded as multi-component or multi-constituent substances. Two-thirds of these UVCB substance were represented in ERC2 by a single representative chemical determined during the categorization of the DSL from 1999-2006. According to Salvito et al. (2020), "*Given the uncertainty and variability in the composition of UVCBs, constituent identification and substance naming are difficult, which presents a major challenge for risk assessment because the chemical structure of the individual constituents determines their fate, exposure, and toxicity*". As previously mentioned, the single UVCB structure in ERC2 was selected from the available information to represent a reasonable "worst-case" structure from an environmental and organism bioavailability perspective as well as hazard potential (i.e., using the attributes of the structure and estimated chemical properties). This approach is based on the "representative chemical constituent-based approach" (Salvito et al. 2020).

In ERC1 (ECCC 2016), each UVCB structure was manually examined to re-verify its suitability as a representative of the UVCB. However, the selection was often performed with only partial knowledge of composition of the entire UVCB substance. Unfortunately, information on UVCB composition is rarely available, and is difficult and costly to produce. Consequently, for prioritization of thousands of chemicals, the representative structure approach is the most suitable and practical method to screen the UVCBs based on the assumption that the structure relied on represents a worse case from a risk classification perspective. Nonetheless, several UVCBs will require manual risk classification even if a representative structure was available. The testing and assessment approaches to addressing UVCB substances is advancing and new approaches may soon be used to verify the adequacy of the representative UVCB structure and help accurately classify the risk of UVCBs. Not all structures have been re-verified at the time of ERC2 publication, as this is a complex resource-intensive task given the number of UVCB structures in ERC2. Therefore, ERC2 results for UVCBs based on a representative chemical structure are regarded as uncertain and many are likely to require manual reclassification based on an updated approach to primarily ensure current ERC2 outcomes are not resulting in false negative conclusions. A flag related to this uncertainty has been added to the specific substances where this applied in Appendix XII.

3.3 Profiling ADME (Absorption, Distribution, Metabolism and Elimination)

Once the substance identity has been performed, an ADME pre-classification filter is engaged to provide alerts for substances that present a highly confident outcome of very low internal and environmental bioavailability or that upon uptake, become distributed in the blood plasma of organisms (protein

plasma binding). Substances with a high molecular fraction occurring in the ionized state at environmentally relevant pH are also identified at this stage. In such cases, these chemicals may be outside of the capability of ERC2 to provide confident classification outcomes using data generated from *in silico* tools or laboratory testing. Often, ecological models assume 100% substance bioavailability and water-lipid partitioning (e.g., bioaccumulation and ecotoxicity models), and bioavailability of "poorly soluble" substances is often artificially enhanced in order to conduct lab tests.

Figure 11 shows the generalized logic flow for providing ADME alerts in ERC2. The degree of bioavailability, both internally in organisms and externally in the environment, is determined using a series of rules described in Appendix II. These rules encompass both tissue permeability and 2D and 3D (3D conformational arrangements) physical-chemical properties that govern environmental partitioning and organism uptake. ADME *in silico* tools used for these determinations are listed in Appendix I. A consensus of greater than five (up to six) measures of bioavailability must result for a substance to be classed as having low bioavailability with high confidence (see Appendix II). Much less than 1% of ERC2 substances meet this criteria of high confidence, with most being UVCB oils and biologicals (e.g., fatty acid tri-ester CAS RN RN 101-34-8).

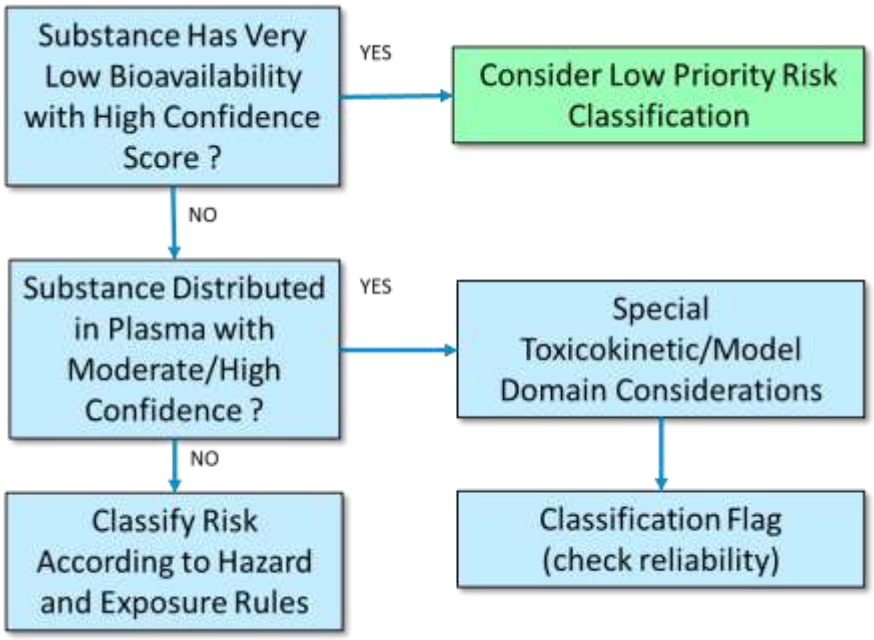


Figure 11: ERC2 Rule structure for selected ADME endpoints

The degree of and confidence with plasma distribution of substances was determined using some key *in silico* indicators and consensus-based logic rules (see Appendix II). The rules identify substances that are not expected partition to lipids with moderate or better confidence, but instead partition to blood proteins such as albumin in blood plasma. Most perfluoro acids are examples of substances that predominantly circulate in the blood of organisms (a small fraction will partition to phospholipids) rather than target lipids such as adipose. Figure 11 shows that for these substances, special toxicokinetic

considerations need to be taken outside of ERC2. Many of the ERC2 *in silico* results for these substances may not apply or will have a high degree of uncertainty. Flags are given to these substances during final classification to determine the reliability of the risk classification outcome.

3.4 Profiling Hazard

3.4.1 Defining Toxicological Space for Hazard Profiling

Once ADME profiling has been completed, a substance is profiled for toxicological hazard using specific descriptors designed to take into account the principle interactions (high level MIEs) a chemical can have with various biological tissues. The interaction(s) and the potency of these interactions can be linked to observed outcomes in existing toxicity data. If toxicity data are absent, these interactions or events can be predictive of adverse outcomes, as discussed in section 2.3. Figure 12 illustrates the principle toxicological linkages from specific chemical interactions with target biological tissues (from Nendza et al. 2014). For example, covalent binding to protein or nucleic acids can lead to "reactive toxicity", often associated with genotoxicity. Steric fit interactions (i.e., 3D conformational arrangements or "docking" interactions) can lead to receptor binding in nuclear receptors (e.g., estrogen receptor) leading to "specific toxicities". Hydrogen bonding or electrostatic interactions with triglyceride-based adipose or cell membrane phospholipids can lead to "non-specific toxicity" (narcosis).

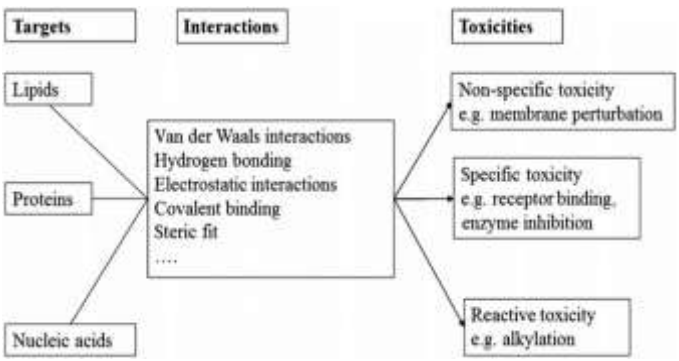


Figure 12: Principle chemical-biological tissue interactions defining toxicological space in ERC2 (from Nendza et al. 2014)

By covering the principle interactions in Figure 12, which are conceptually similar to the AOP concept, the hazard profile in ERC2 is able to suggest *plausible* mechanistic reasoning to explain adverse outcomes in observed toxicity data and thereby increase the confidence of a hazard classification (Section 4.0). Because profiling can be performed from the chemical structure alone, it is comparable to

the idea of chemical product design before commercialization and is therefore useful to predict outcomes when *in vivo* data are missing. Mechanistic reasoning can also be used in read-across frameworks when forming groups of chemicals that have the potential to interact along the same pathway(s) (Sapounidou et al. 2021).

3.4.2 Toxicological Data Used for Hazard Profiling

ERC2 integrates *in silico, in chemico, in vitro* and *in vivo data* (mammalian and aquatic) into the hazard profile to arrive at consensus AOP organized conclusions for hazard classification. Appendix I lists the available *in silico* tools, *in chemico* assays, and database sources of *in vitro* and *in vivo* data used for hazard profiling. Evaluation of the quality of individual data points or databases was not undertaken given the difficulty of performing this task for ~12 200 substances and over 10 million data points. However, lack of consensus profiling outcomes in ERC2 could point to lower quality data as well as data gaps.

A data preference hierarchy is used for hazard classification and hazard confidence scoring according to AOP theory:

in silico < in chemico < in vitro < in vivo

The above data hierarchy is only implemented when there is lack of consensus or there are significant data gaps among the various data types. The above sequence is consequently influential for determining ERC2 hazard outcomes in these situations.

3.4.3 Hazard Descriptors

The following sections outline the five hazard descriptors that account for the toxicological space and types of data used to create a substance hazard profile. Hazard profiling is potency-based. This means the classification rules are driven by potency of chemical interactions or observed effects. This is described in greater detail in the appendices associated with each of the hazard descriptors. The following descriptors are used to profile hazard:

- Receptor-mediated interactions
- Chemical reactivity/genotoxicity
- Mode of toxic action
- Food web toxicity
- Cumulative toxicity

The above descriptors are discussed in the following section noting that profiling of cumulative toxicity is only conducted when information on co-occurrence in the environment is available and thus acts only as an early indicator of the potential for future cumulative assessment activities.

3.4.3.1 Receptor-Mediated Interactions

Receptor-mediated interactions refers to the ability of a chemical substance to bind or dock (i.e., molecular docking) with specific nuclear receptors in proteins and nucleic acids. This often requires a specific 3D steric orientation of the molecule in order to interact with the receptor. For example, in the case of estrogen receptor binding, xenobiotic chemicals can bind to or block the alpha and beta estrogen receptor ligands in various tissues in organisms (Książek and Bryl 2015). The potency and nature of binding can often be reliably predicted using *in silico* tools (e.g., Mekenyan and Sirafimova 2009; Wang and Wang 2010; Cotterill et al. 2019). Receptor-mediated interactions of xenobiotics with key receptor ligands, such as the estrogen or androgen receptors associated with sex hormones, is well described in the literature and is not repeated here.

Endocrine interactions may or may not lead to adverse outcomes associated with reproduction and development, but the interaction can suggest a plausible mechanism for effects in species not yet tested or observed. Recommendations for the scientific evaluation of endocrine active substances for regulatory purposes, summarized from a SETAC Pellston workshop in 2016, is presented by Mathessien et al. (2017) and was considered during the development of this descriptor. A Government of Canada and CMP Science Committee perspective on endocrine disrupting chemicals (EDCs) found in the July 2018 meeting report from the Committee[8] was also considered.

ERC2 examines the potential for endocrine activity of substances by considering the EAT in EATS - that is, interactions with estrogen (E) receptors, androgen (A) receptors and thyroid (T) receptors, but also includes the aryl hydrocarbon receptor (AhR). Interference with steroidogenesis (S) was not considered in this version of the ERC, as too few data and few viable *in silico* tools currently exist to account for steroidogenesis. This interaction can be considered in a future update to ERC2. Binding to AhR, also well documented in the literature, is associated with “dioxin-like” or “aryl hydrocarbon” toxicity resulting from interactions of planar or co-planar molecules (e.g., dioxins, some PCBs congeners and PAHs) with protein receptors that regulate gene expression (e.g., that regulate cell differentiation and immunity responses) resulting in developmental effects (e.g., Giani-Tagliabue et al. 2017).

In silico, *in vitro* and *in vivo* data for both mammalian and aquatic species were organized along the AOP concept to plausibly explain adverse developmental and reproductive effects associated with endocrine and AhR interactions. Appendix I outlines the tools and databases used for receptor-mediated effects and Appendix III describes the hazard and confidence classification rules used for each of the data types mentioned. Figure 13 below shows the relative distribution of data available for profiling receptor-

[8] https://www.canada.ca/en/health-canada/services/chemical-substances/chemicals-management-plan/science-committee/meeting-records-reports/committee-report-july-18-19-2018.html

mediated effects for ~12 200 chemicals in ERC2 according to data type. A high percentage of *in vitro* data were available for ERC2 chemicals largely because of the CERAPP and CoMPARA databases used to develop the *in silico* tools contained in the OPERA model for estrogen receptor and androgen receptor interactions (Appendix I).

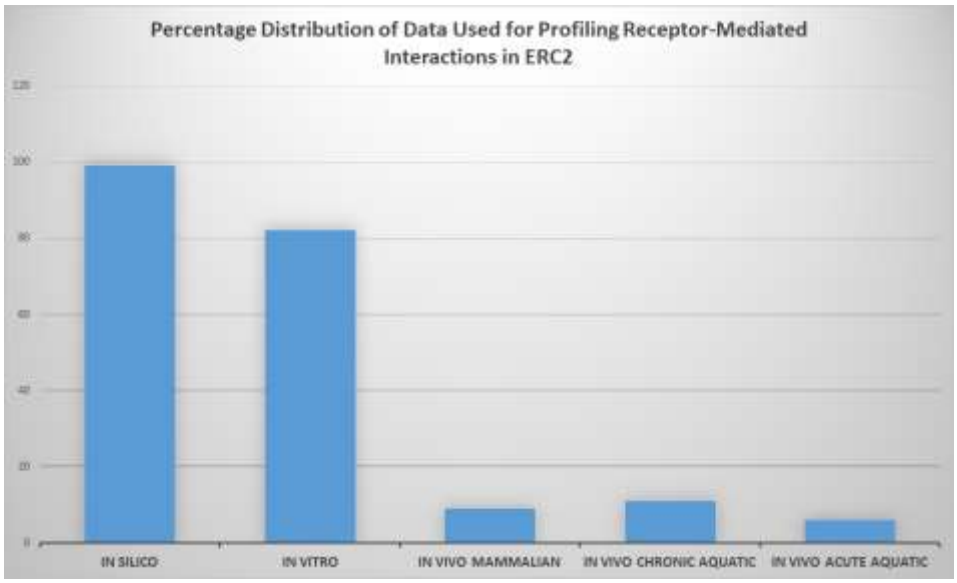


Figure 13: Relative percent data distribution for profiling receptor-mediated interactions for ~12 200 substances in ERC2 by data type

3.4.3.2 Chemical Reactivity and Genotoxicity

ERC2 examines the potential for developmental and reproductive effects associated with general chemical reactivity (Figure 12) and genotoxicity. Both genotoxic endpoint-specific information (e.g., DNA damage, chromosomal aberrations) and endpoint agnostic information (e.g., DNA or protein binding) is used to cover the toxicological space associated with interactions of chemicals with nucleic acids and proteins in or at the surface of organism tissues. Chemical reactivity, such as covalent bonding of a substance or its metabolites with nucleic acids and proteins, can lead to various developmental and reproductive effects. For example, protein binding can lead to non-lethal effects in aquatic organisms such as *Tetrahymena pyriformis* (Schultz et al. 2010; Roberts et al. 2010; Richarz et al. 2014) and terrestrial organisms such as earthworms (Princz et al. 2014). ECCC conducted an endpoint vs. endpoint correlation[9] of *in chemico* protein binding potency data (i.e., median reactivity concentration or RC50) and median lethal and sub-lethal effects in data for fish, daphnids and algae using this feature in v4.3 of the OECD QSAR Toolbox. Figure 14 shows these correlations and while data are sparse for algae and

[9] This refers to correlating two types endpoint data in a linear regression to determine their correlation.

daphnia, the available evidence suggests binding potency is positively correlated with increased median effects concentrations (LC50, EC50) in these organisms.

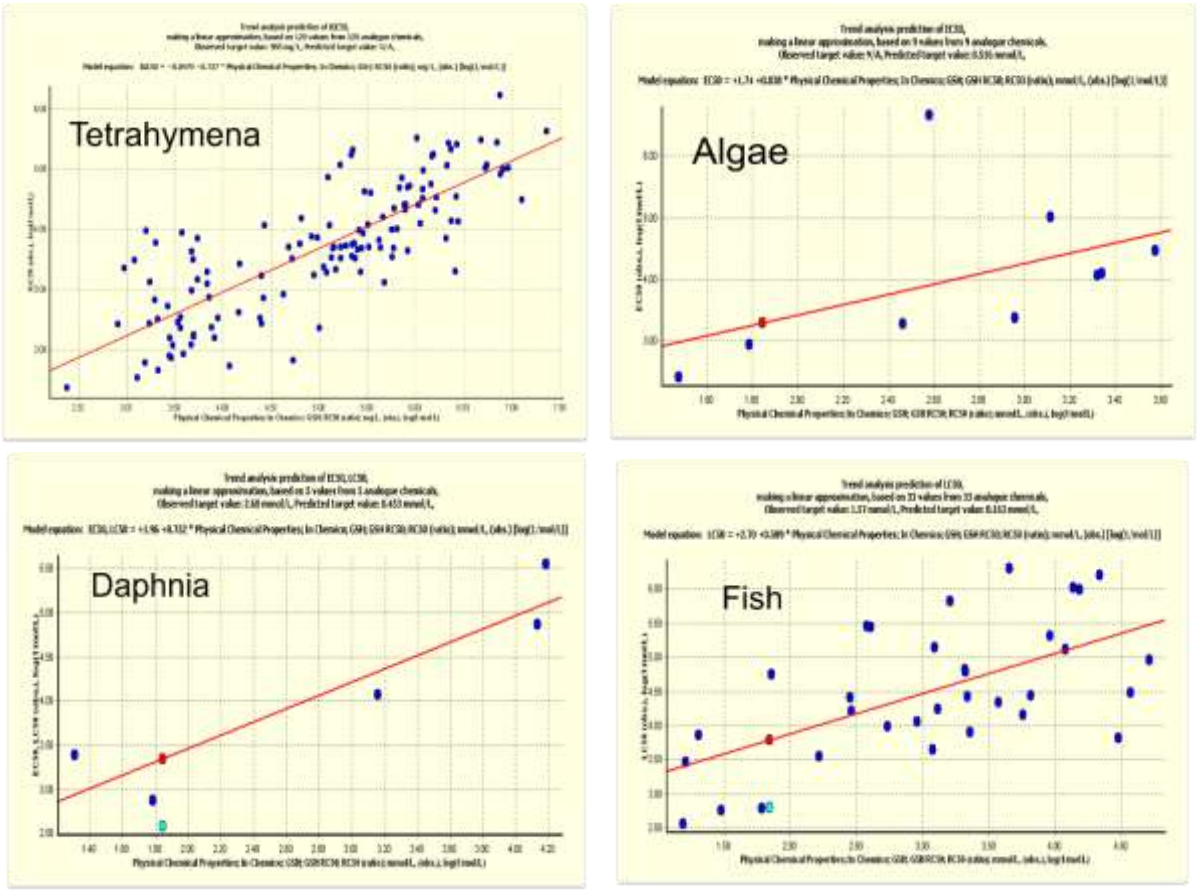


Figure 14: Endpoint vs. endpoint correlation analysis of RC50 protein binding (log 1/mol) data and median effects concentrations for four aquatic species EC50 median effect concentration (log 1/mol)

Genotoxicity refers to chemical substances that can damage genetic information (e.g., DNA, mRNA) within a cell, causing mutations which may lead to cancer. Genotoxicity is often confused with mutagenicity. It is commonly cited that all mutagens are considered genotoxic, while not all genotoxic substances are said to be mutagenic. ERC does not pursue the carcinogenic aspects of genotoxicity. For the purposes of profiling genotoxicity for non-human organisms, ERC2 examines “eco-genotoxic” responses related to developmental and reproductive effects and thus mainly mutagenic responses (Kirkland 1998). Most non-human organisms are less likely to develop cancerous outcomes due to much shorter lifespans. However, there are exceptions to this general observation, such as carcinomas in long-lived sea turtles (Arthur et al. 2008). As noted earlier in this document, exposure to chemicals that can interact with genetic materials, even at very low concentrations, may result in developmental and

reproductive toxicity. When coupled with chemical properties that lead to wide environmental distribution and long residence time in the environment, adverse effects may be transferred beyond the time frames of actual exposure and mutagenic responses are genetically transferred to subsequent generations not exposed to the chemical agent.

Similar to receptor-mediated interactions, *in silico*, *in chemico*, *in vitro*, and *in vivo* data for both mammalian and aquatic species were organized along the AOP concept to plausibly explain mutagenic, developmental and reproductive effects associated with chemical reactivity and genotoxicity. Appendix I lists the *in silico, in chemico, in vitro* and *in vivo* data types, sources and application in ERC2 for genotoxicity profiling for mutagenic responses. As described earlier, the concordance between and within the data sources is analysed to determine plausible mechanistic reasoning for observed adverse outcomes and not yet observed, but plausible adverse outcomes.

In silico profiling is comprised of 14 mechanistic profilers and QSARs that examine DNA and protein binding, the potency of protein binding and functional groups on a molecule that can be linked to developmental and reproductive effects. These models largely examine interactions of the parent substance (e.g., the electrophile in nucleophilic addition-substitution reactions with biomolecules) with some approaches accounting for metabolite interactions as well via S9 activation of the parent molecule (e.g., Mekenyan et al. 2004; Serafimova 2007; Mekenyan et al. 2007; Gerberick et al. 2004; Natsch et al. 2008; Natsch et al. 2015; Jaworska et al. 2015; Urbisch et al. 2015; Dimitrov et al. 2016; Benigini et al. 2012; Benigini and Bossa 2012). The mechanistic alerts resulting from the *in silico* profiling can be linked to *in vitro* and *in vivo* genotoxic outcomes such as DNA adduct formation, DNA damage (e.g., strand breaks, chromosomal aberrations, presence of micronuclei, gene mutations).

In vitro data collected for this interaction are endpoint-based and focus on mutagenicity, but overlap into carcinogenicity as well (e.g., chromosomal aberrations). Appendix I lists the data sources used to collect *in vitro* information which include similar endpoints as *in silico* profiling. These data may or may not use metabolically active cell lines and thus may or may not include exposures to metabolites of the target chemicals tested. Quantitative developmental and reproductive *in vivo* toxicity data collected for receptor-mediated interactions are again used for genotoxicity with the understanding that any or all of these interactions could plausibly lead to an observed adverse *in vivo* outcome. Gentoxicity data however also include categorical genetic damage data, which are not related to receptor-mediated interactions. Figure 15 shows the relative distribution of data collected for profiling chemical reactivity and genotoxicity by data type.

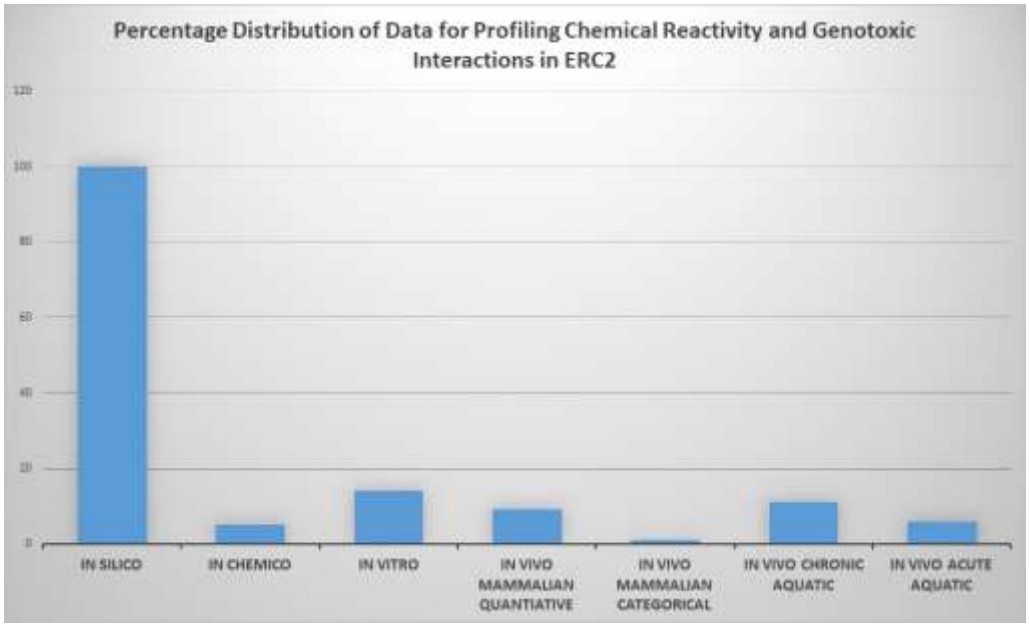


Figure 15: Relative percent data distribution for profiling chemical reactivity and genotoxicity for ~12 200 substances in ERC2 by data type

3.4.3.3 Mode of Toxic Action

Unfortunately, there is no consistent definition of mode of toxic action (MoA); however, there are more recently accepted definitions that can be used for the purposes of this document. MoA refers to a functional change (e.g., adverse outcome) resulting from exposure to organic chemicals interacting with biological tissues at the cellular level (Kienzler et al. 2019). This differs from mechanisms of action (MechoA) described in the previous two sections, which involve interactions at the molecular level and can therefore be referred to more accurately as MIEs (e.g., Bauer et al. 2018). Most *in silico* tools used to profile MoA for non-human species are based on interactions observed at the whole organism level, mostly acute studies in fish (Baron et al. 2015; Kienzler 2017; Martin et al. 2015). MoA considerations have also been used to propose ecological Thresholds of Toxicological Concern (eco-TTC) for various modes of action. Eco-TTCs have been advocated as a useful approach to screening and prioritization of chemicals (de Wolf et al. 2005; Williams et al. 2011; Belanger et al. 2015).

MoA is an important descriptor in ERC2 because it acts as a cellular-level screen for adverse outcomes in addition to those MIEs used for receptor-mediated toxicity and toxicity due to chemical reactivity and genotoxicity. By combining the descriptors for receptor-mediated toxicity, chemical reactivity/genotoxicity and MoA, ERC2 essentially creates a “Mech-MoA” outcome, quite similar to mechanistically-based classification schemes from Bauer et al. (2018) and the more extensive mechanistic MoA classification scheme from Sapounidou et al. (2021). In ERC1, MoA profiling for specific

modes of toxic action was responsible for 40% of the 195 substances resulting classified as high hazard (ECCC 2016; OECD 2017).

In ERC2, MoA is determined using a consensus of *in silico* tools (QSARs) and MoAs determined using tissues residue-based toxicity ratios. Tissue residue (TR) refers to a concentration of a chemical measured in an organism (e.g., mmol/kg, ug/kg), typically on a whole body wet weight basis or lipid basis (mmol/kg lipid). The TR approach was used for ERC1 (ECCC 2016), but was considerably expanded to include additional QSAR and TR approaches (Armitage et al. 2018). Approaches for calculating TRs was performed using three approaches:

- Critical body residue (CBR) approach (e.g., McCarty et al. 1992)
- Critical membrane concentrations (CMCs) (e.g., Endo et al. 2011)
- Lethal chemical activity (e.g., Mayer et al. 2011)

Different approaches to estimate partition coefficients were used to calculate the TRs in the approaches listed above. This resulted in seven methods for TR estimation which are described in more detail in Appendix V.

Once TR values have been calculated, toxicity ratios are then estimated to determine narcotic vs. specifically acting substances. Toxicity ratios refer to the difference in concentration between a baseline narcotic and a chemical exerting a more specific MoA by comparing the TR concentration estimated using one or all of the above approaches and the known internal effects concentration associated with median lethal effects (e.g., Maeder et al. 2004).

$$\text{Toxicity Ratio} = TR/IEC_{50}$$

where TR is the estimated TR (mmol/kg) and IEC_{50} is the internal effects concentration for median lethality (mmol/kg) for the chemical of interest determined by the three bulleted methods below. With this formulation, values greater than unity indicate specific modes of action whereas values equal to or less than one indicate baseline toxicity. An acute to chronic ratio of 10 was applied to the lower bound of the baseline toxicity to account for the extrapolation to chronic lethal toxicity. Therefore, the thresholds for determining specially acting substances from non-specifically acting were set as follows:

- Median critical body residue (CBR50) = (2.5 mmol/kg + $50/K_{ow}$)/10 (McCarty et al. 1992 ; Maeder et al. 2004; McCarty et al. 2013)
- Median critical membrane concertation (CMC50) = 10 mmol/kg lipid (Endo et al. 2011)
- Median lethal activity (LA50) = 0.001 (Schmidt et al. 2015; Schmidt et al. 2016)

The rules for MoA determination, hazard classification and confidence scoring are given in Appendix V. These rules first seek to determine the consensus within QSAR and TR approaches, and then between them. Greater confidence is assigned when both approaches are in agreement. A similar approach to ERC2's MoA QSAR consensus and confidence assignment was adopted for MoA assignments in the ENVIROTOX database (Connors et al. 2019) as described in Kienzler et al (2019). The MoA "quaternary ammonium", based on the ASTER and TEST QSAR MoA assignments, refers to alkyl ammonium

surfactants (Russom et al. 1997; Barron et al. 2015). IEC for these chemicals are not available in Escher et al. (2011) and required a toxicity ratio approach to determine potency using the CMC50 approach. Figure 16 shows the relative distribution of substances with a “non-specific” MoA (narcosis) and selected specific MoAs in ERC2 determined according to the classification rules outlined in Appendix V.

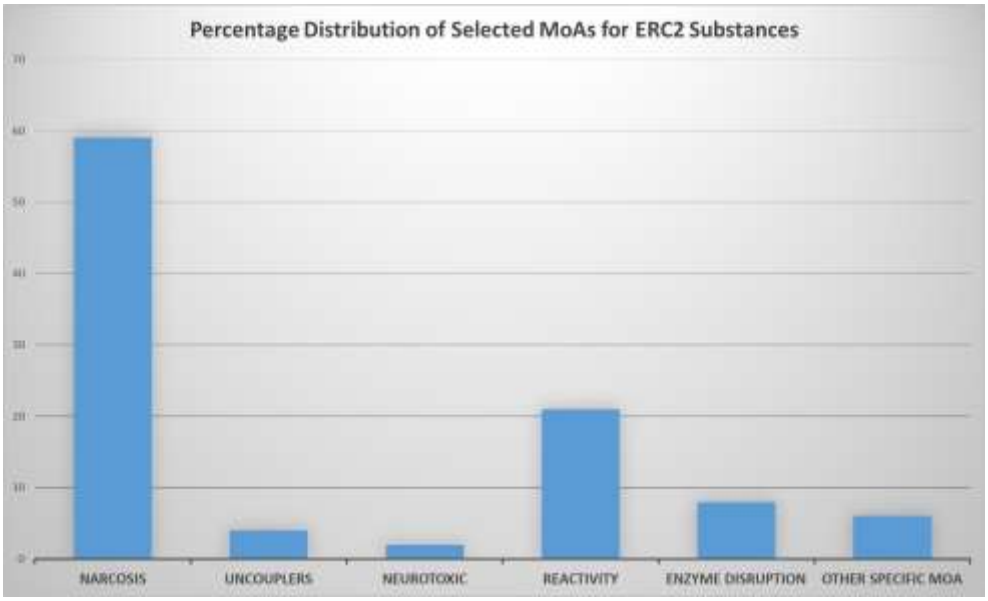


Figure 16: Distribution of MoA assignments in ERC2 including consensus outcomes

3.4.3.4 Food Web Toxicity

This descriptor integrates the properties of persistence, bioaccumulation and internal toxicity to determine the potential for food web effects by tracking the fate of an organic chemical using a coupled environmental fate and food web mass balance model known as the Risk Assessment IDentification And Ranking (RAIDAR) model (Arnot et al. 2006; Arnot et al. 2009). RAIDAR quantifies chemical transport from diffuse sources in an evaluative, regional scale environment to representative ecological receptors and humans (ARC 2018). The default evaluative environment uses a regional scale landscape with an area of 100 000 km^2 (90% land, 10% water with underlying sediment), which is roughly the size of southern Ontario. The geographical conditions of the default environment are “generic” (not specific to any particular region) and are similar to those of the EQuilibrium Criterion (EQC) model (MacKay et al. 1996). The RAIDAR model, however, differs from the EQC model because it includes vegetation, and aquatic and terrestrial organisms comprising ecological and agricultural food webs and humans. A variety of representative plant, invertebrate and vertebrate species including fish, wildlife, agricultural crops and livestock, and humans are included in RAIDAR.

RAIDAR simulates food web exposures using bioaccumulation sub-models to estimate concentrations in representative species given a default and actual rate of emission to the environment. The food web bioaccumulation module links the bioaccumulation sub-models for each representative species by

trophic interactions (i.e., feeding relationships). All organisms are comprised of three proximate phases including lipids, non-lipid organic matter (or organic carbon) and water. The default volumes of these phases (e.g., lipid contents) in the biota are indicative of measurements for the representative species (Arnot et al. 2006; Arnot et al. 2009; ARC 2018). Version 3.0 of RAIDAR was used for ERC2 food web toxicity profiling. This more recent version includes expanded food webs by including the mechanistic dietary bioaccumulation model AQUAWEB (Arnot and Gobas 2004) and other models. Version 3.0 also includes expanded capabilities for addressing ionizing substances (Arnot et al. 2011; ARC 2014; ARC 2018) and expanded biotic partitioning properties to account for non-lipid distribution in organisms.

The RAIDAR model requires that an internal threshold of toxicity be selected as an input parameter for calculating a chemical's potential to result in food web effects. IECs (mmol/kg) were therefore calculated for fish and invertebrates using empirical and predicted acute fish median lethality toxicity data performed for the ERC2 MoA descriptor (see Appendix V). However, for avian and mammalian wildlife species, a default intake rate (mg/kg/d) of 0.15 mg/kg/d was used. This value is approximately equivalent to the 5th percentile of no observed effect levels (NOELs) in the Munro Threshold of Toxicological Concern (TTC) database (Leeman et al. 2014).

The mode of entry to the environment used for model simulations was "mostly water", where air-water partitioning of the substance was used to adjust the fraction emitted to water (see section 3.5.3.3). Water is often required in industrial processes or carries substances from down-the-drain releases to aquatic environments via wastewater treatment systems and represent the majority of industrial chemical emissions to the environment. Releases to soil were nonetheless included in ERC2 to account for potential terrestrial exposures and effects, for example, when significant adsorption to biosludge (e.g., >70%) occurs which is in turn applied to agricultural lands. Other than biosolids application, which is not amenable to a regional scale model simulation, little information is available for ERC2 substances to fully account for terrestrial impacts. It is expected that the classification of hazard using an aquatic emission scenario in RAIDAR as well as other hazard descriptors described previously will be protective of terrestrial species given cross-species susceptibility.

To determine the potential for acute toxicity in both aquatic and terrestrial food webs, ERC2 uses the hazard assessment factor (HAF) from version 3.0 of RAIDAR. The HAF is a ratio of the tissue residue (TR) concentration estimated in the most sensitive species in the RAIDAR aquatic or terrestrial food web (mmol/kg), based on a default emission (C_U) (e.g., 1 kg/hr), and the MOA-based IEC (mmol/kg) (C_T). HAF values of 1 or higher suggest acute effects in sensitive species are expected using the default rate of emission. These species can be regarded as the "most vulnerable" based on this integrated PBT descriptor. HAFs are independent of the actual chemical emission rate, but span several orders of magnitude for the ERC2 organic substances characterized. The HAF concept and rules for hazard classification remain largely unchanged from ERC1. Further details on how HAFs are calculated can be found in Arnot and Mackay (2008) and, specifically, as it pertains to the substances being addressed in this report, in ARC (2018). HAF calculations and other food web estimates from RAIDAR account for biotransformation by model organisms in the food webs. Figure 17 below illustrates the range of HAF values (unitless) calculated for an aquatic emission for ERC2 substances in domain of the model. Fewer than ~12 200 results could be generated by the RAIDAR model because some model input values could

not be generated or found in the literature. Appendix VI describes the hazard classification and confidence rules used for food web toxicity in detail.

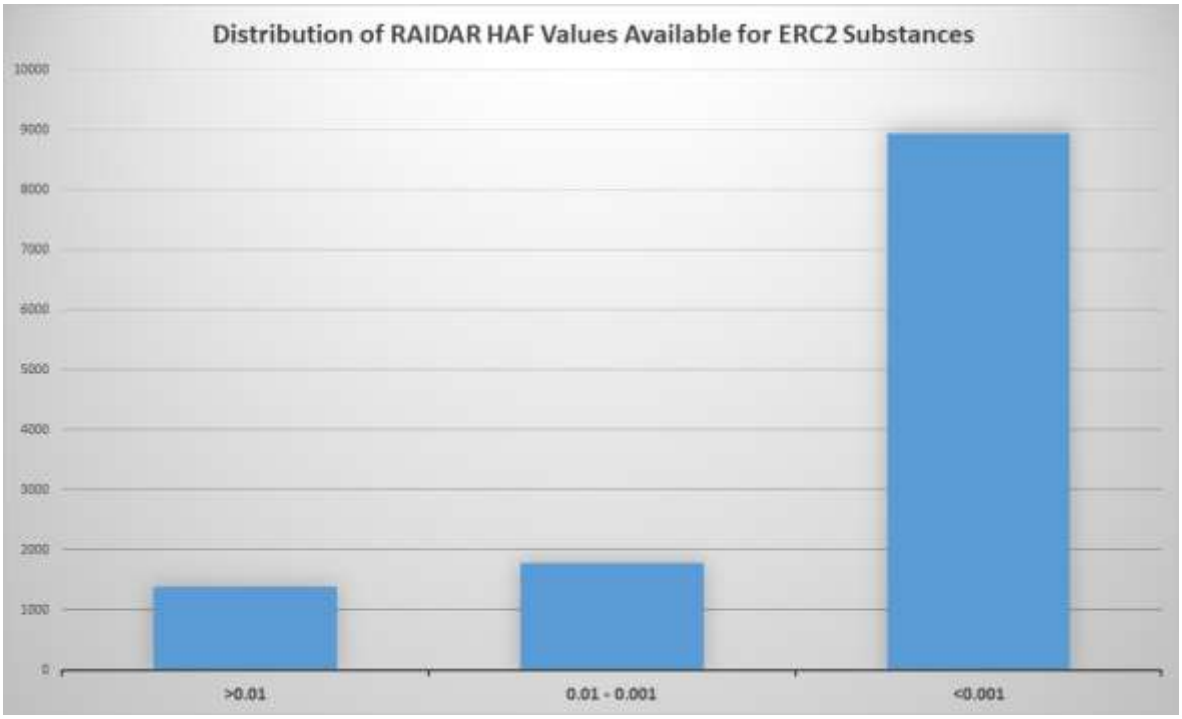


Figure 17: Distribution of RAIDAR Hazard Assessment Factors (HAF) available for ERC2 substances with RAIDAR results

3.4.3.5 Accounting for Cumulative Toxicity from Combined Chemical Exposures

The toxicity from the combined exposure to multiple chemicals in the environment is closer to reality than single chemical exposures, but the predictive assessment of chemical mixtures presents a challenge to regulators (Backhaus and Faust 2012). A review of lessons learned from cumulative risk assessment (CRA) experience by the US EPA is given in Gallagher et al. (2015). The authors also conclude that CRA challenges current regulatory paradigms, but that an iterative and tiered framework can help address the complexity and uncertainty in CRA. In Canada, the subject of addressing environmental impacts from exposure to chemical mixtures (not referring to UVCBs) that occur in the environmental as a result of human activities and natural processes is a current priority area of work for ECCC and will continue to be so in the future. In 2015, the Government of Canada's CMP Science Committee[10] addressed CRA using the Government of Canada's CRA for phthalates as an example (EC, HC 2015). One of the main

[10] Chemicals Management Plan Science Committee - Canada.ca

conclusions from the Committee on CRA was that knowledge of the co-occurrence of chemicals is and will remain a main driver for conducting a CRA.

"*In the context of the CMP, the specific motivation would be evidence of co-exposure to a group of chemicals."* – Government of Canada's CMP Science Committee Report 2015[11]

In 2017, Canada's Standing Committee on Environment and Sustainable Development (ENVI)[12] produced a report suggesting critical amendments to CEPA[13]. Recommendations 45, 46 and 56 made by the ENVI Committee relate to the need to better address CRA under CEPA:

"*The committee recommended that CEPA be amended to '[…] require investigation of the effects of any proposed or final regulation or instrument on vulnerable populations and marginalized communities […and…] aggregate exposures, and cumulative and synergistic effects, in determining how to regulate a toxic substance' (recommendation 56)."*

With respect to CRA, the Government of Canada's 2018 response supports the intent of the committee's recommendation to amend CEPA "[…] *by adding a new requirement that the Ministers or their delegates, when determining if a substance is toxic, assess aggregate exposure to and cumulative and synergistic effects of the substance, and that the Ministers use a process that looks at multiple exposure points of a chemical substance" (recommendation 46)*.[14]

As a system for prioritizing the risk of substances in the environment, ERC2 can address aspects of CRA for "targeted mixtures", that is, effects from chemical mixtures where prior knowledge of their co-occurrence in the environment is known or expected (e.g., phthalates). ERC2 was not designed to address effects from "non-target mixtures" where there is no prior knowledge of a chemical mixture and potential effects must be determined from chemical analysis in, for example, "hot-spot" areas of the Canadian environment by sampling environmental media.

Target mixture toxicity is accomplished in ERC2 using an internal (tissue residue) toxic unit approach (Dyer et al. 2010) also used in the phthalates assessment by ECCC (EC, HC 2015). ERC2 uses the IEC_{50} for various modes and mechanisms of action based on the data compilation in Escher et al. (2011) and those calculated using acute fish toxicity data. IEC_{50} values were available for non-specific and specific MoAs for most substances in ERC2 distributed according to Figure 16. Appendix VII gives detail on their derivation. The IEC can be regarded as an internal eco-toxicological threshold of concern (eco-TTC_i), which is a suggested alternative approach in Kienzler et al. (2019) to the traditional media concentration eco-TTC using, for example, no observed effect concentrations (NOECs). Given that IEC_{50} values reflect acute exposures, assessment factors (AFs) should be applied to lower the point of departure to chronic

[11] https://www.canada.ca/en/health-canada/services/chemical-substances/chemicals-management-plan/science-committee/meeting-records-reports/committee-report-november-18-19-2015.html
[12] ENVI - Home - House of Commons of Canada (ourcommons.ca)
[13] Committee Report No. 8 - ENVI (42-1) - House of Commons of Canada (ourcommons.ca)
[14] Follow-up report to the Standing Committee on the Canadian Environmental Protection Act - Canada.ca

no effect (IEC_{NOEC}) concentrations in a similar manner to traditional concentration data (Okonski et al. 2021).

The IEC_{50} represents the threshold of effects to compare to the total tissue burden from aggregated exposures in targeted chemical mixtures. The term "cumulative toxicity" is used in this document to avoid confusion with "cumulative effects". Cumulative effects, also referred to as cumulative environmental effects and cumulative impacts, can be defined as changes to the environment caused by the combined impact of past, present and future human activities and natural processes.[15] In Canada, cumulative effects assessment is governed under the *Canadian Environmental Assessment Act* (CEAA) and generally refers to the assessment of impact from human projects.[16]

The IEC_{50} is dependent on mode of action and therefore are able to address chemical class CRA (e.g., hindered phenols), endpoint or mechanism based CRA (e.g., endocrine effects), and combinations of both, as recommended by the CMP Science Committee in their 2015 report. As such, it is also a useful approach for grouping chemicals for further cumulative assessment based on these metrics.

Cumulative risk profiling can be performed in ERC2 by summing the tissue residues (mmol/kg) for targeted mixtures estimated using the RAIDAR model or data from other sources such as monitoring data. The summed internal mixture concentration or aggregate TR (TR_A) can be estimated for a single or multiple sensitive species for representative aquatic and terrestrial food web species from the RAIDAR model output. A tissue residue-based CRA profile is then derived as follows:

$$CRA_{TARGETED} = TR_{NOEC} / TR_A$$

Appendix VII provides greater detail on the approach for deriving IEC_{50} for determining the TR_{NOEC} according to the mode of action described in Appendix V.

Finally, profiles of CRA in ERC2 are currently not computed. This descriptor was added to ERC2 to allow "on the fly" estimation of CRA values for targeted mixtures when information becomes available to suggest a CRA for a targeted mixture is appropriate. Consequently, this descriptor does not impact the classification of hazard or risk in ERC2 as it was not used in the consensus model for hazard classification.

3.5 Profiling Exposure

3.5.1 Defining Exposure Space for ERC2

[15] Cumulative Effects - Canada.ca

[16] https://www.canada.ca/en/impact-assessment-agency/services/policy-guidance/assessing-cumulative-environmental-effects-ceaa2012.html

Similar to hazard profiling, a series of descriptors is used to define the spatial and temporal scale of exposure in ERC2. This is achieved by using a combination of multimedia environmental fate simulations, mode of entry in the environment, and emission rate (estimated) data (kt/yr). Multimedia fate and food web exposure simulations were carried out using v3.0 of the RAIDAR model (ARC 2018) described briefly in section 3.4.3.4. Figure 18 illustrates the scale of exposure considered in ERC2 for classification purposes.

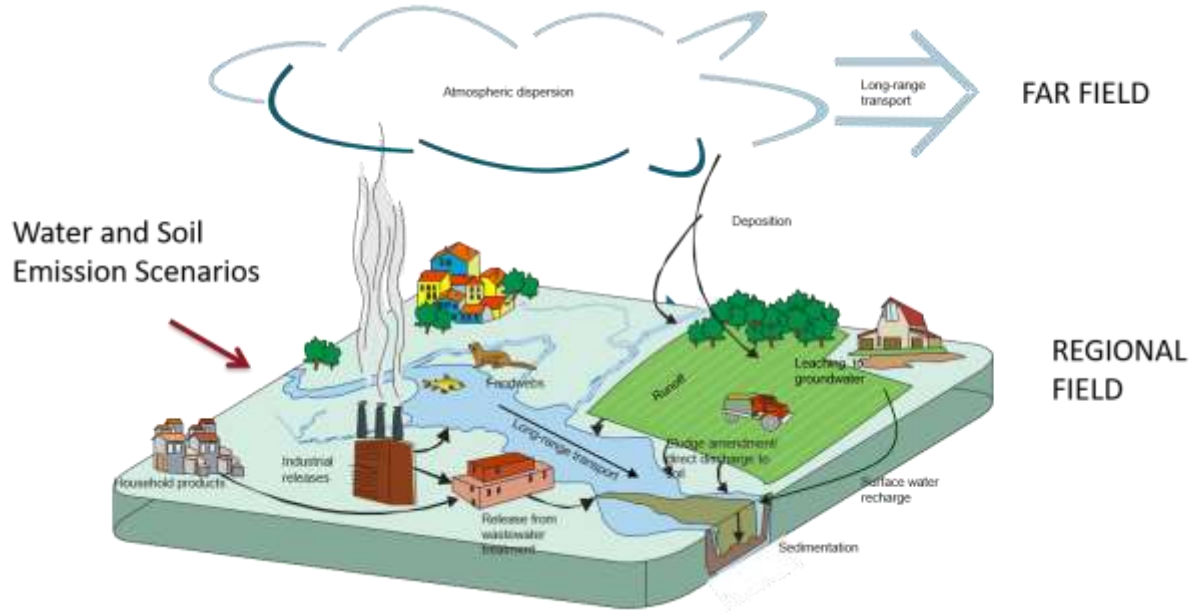


Figure 18: Scale of exposure considered in ERC2

Classification of exposure concern is based on the likelihood of organism contact, primarily driven by the persistence and distribution of a substance in the environment. Exposure scale ranges from the near field (~local environments near emission sources) to the far field (regional to global environments) by considering integrated persistence and mobility descriptors. External concentrations (e.g., mg/L) in various media traditionally used as predicted environmental concentrations (PECs) are not used in ERC2 exposure classification; rather, internal measures of exposure (whole body tissue residues in mmol/kg or ug/kg), which better integrate chemical fate both external and internal to the organism, are applied. Similar to hazard profiling, tissue residue concentrations are estimated in aquatic and terrestrial food webs at a regional scale (100 000 km^2) according to the food web organisms used in the RAIDAR model v3.0. ERC2 has the ability to aggregate tissue residue concentrations for known or targeted chemical mixtures and compare to internal toxicity thresholds of concern, as discussed in the previous section.

3.5.2 Parameterizing the RAIDAR v3.0 Model

The RAIDAR model requires input of selected physical-chemical properties (e.g., acid dissociation constant [pKa], molecular weight), selected partition coefficients (e.g., octanol-water partition coefficient [K_{ow}], air-water partition coefficient [K_{aw}], organic-carbon-water partition coefficient [K_{oc}]) and environmental media and biota half-lives. Given that the K_{ow} is an important partitioning property used by the RAIDAR model, median empirical values representing multiple data sources (collected via the OECD QSAR Toolbox v4.3) were gathered and used as input, when available. Empirical K_{ow} values were available for approximately 9% of ERC2 substances. When empirical values were not available, predicted values from the EPIWIN KOWWIN and ACD (LogD consensus) models (see Appendix I) were used. If a substance's mass fraction was >10% ionized at pH 7.4, then LogD values were used (logarithm of the octanol-water dissociation constant), else the mean of K_{ow} values from the two models was used. QSAR models were used to generate input data for all other physical-chemical properties, largely using the EPIWIN suite of models (ARC 2018).

As previously mentioned, RAIDAR v3.0 includes the ability to address the behaviour of ionizing organic chemicals (IOCs). Input of partition coefficients and physical-chemical properties for IOCs were used unless partition coefficients for the charged species (e.g., LogD at pH 7.4) could be obtained. For example, input of K_{oc} for the charged form of either an acid or a base (K_{oci}). Data for K_{oc} and more specifically K_{oci} are quite limited. Specific QSARs were used in this case (Karicoff 1981; Seth et al. 1999; US EPA 2011; Trapp et al. 2010; Franco and Trapp 2011). RAIDAR can only accept a single dissociation constant for an IOC and it is assumed that the IOC is either an acid or a base. Molecules with complex and multiple ionization centers and zwitterions are simulated using the dissociation constant of the most acidic or basic ionogenic functional group. This was considered a necessary simplifying assumption as several IOCs have multiple ionization centers containing both acidic and basic groups. Greater detail on physical-chemical input parameterization of RAIDAR v3.0 for ERC2 is available in ARC (2018).

All versions of the RAIDAR model require the input of half-lives for several media and biota. Degradation half-life data for air, water, soil and sediment compartments and biotransformation half-lives in vertebrate species (i.e., fish, birds and mammals) are required. The AOPWIN QSAR model in EPIWIN Suite was used for air half-lives and were estimated by reciprocally combining the AOPWIN hydroxyl radical and ozone half-lives. The default assumptions for radical concentrations (5×10^5 radicals/cm^3 and 7.0×10^{11} molecules/cm^3, respectively) and a 24-hour reaction period were assumed. Predictive models for estimating biodegradation from chemical structure included some QSARs[17] in EPIWIN's BIOWIN suite of models (US EPA 2011). However, as these models do not estimate half-lives, half-life extrapolation methods using BIOWIN output for water were used (Arnot et al. 2005; ARC 2018). Half-lives in soil and sediment followed the extrapolation ratio of 1:2:9 from water (i.e., water:soil:sediment) according to Aronson et al. (2006). For practical reasons, degradation reactions resulting from hydrolysis and photolysis were not considered, noting that these could be important degradation process for certain chemicals. The sediment compartment in which the organisms live is assumed to be primarily aerobic,

[17] BIOWIN 1, 3, 4 and 5 models were selected

therefore no anaerobic degradation was considered in the model. Fish biotransformation half-lives were obtained from various QSARs including EPI Suite (BCFBAFWIN), the Iterative Fragment Selection (IFS) model from Brown et al. (2012) and also the OPERA model (Mansouri et al. 2018). Mammalian biotransformation half-lives were calculated as the geometric mean of the biotransformation and total elimination half-lives for humans using the IFS QSARs (Arnot et al. 2014). Mammalian half-lives were assumed for avian species given the lack of available data for avian species. Finally, all biotransformation half-lives were scaled to a default organism mass using allometric relationships (ARC 2018).

The bioaccumulation model within RAIDAR v3.0 is largely based on the AQUAWEB model Arnot and Gobas (2004). However, greater chemical-specific partitioning for specific organism tissues can be entered, if available. Protein-water partition coefficients, carbohydrate-water partition coefficients, membrane-water partition coefficients, and storage lipid-water partition coefficients can now be entered to better understand the distribution in target model organisms (ARC 2018). For IOCs, partitioning information for neutral species and the charged species can be entered for in place of the neutral values. However, these were not available for ERC2 substances and consequently the above tissue partition coefficients for the neutral form of ERC2 for substances were used. This was considered a conservative assumption for subsequent bioaccumulation calculations.

Finally, model domain of application was considered when interpreting the quality of model output from RAIDAR v3.0. Fugacity and extreme property warnings were tracked and factored into the confidence scoring of model output used for exposure descriptors similarly to RAIDAR HAF calculations (see Appendix VI). The warnings indicate when physical-chemical property data used to parameterize the model may not be of good quality (extreme low and high values) and may result is mass-balance concentrations exceeding maximum solubility (fugacity warnings) in various media.

3.5.3 Exposure Descriptors

The following sections describe the exposure descriptors used to classify exposure concern in ERC2.

- Response time
- Mobility
- Emission pattern
- Food web exposure
- Margin of exposure

The majority of descriptor values were generated using the RAIDAR v3.0 model, except chemical quantity (mass) and use pattern. Appendices VIII-XI describes the logic rules used for exposure classification and confidence scoring.

3.5.3.1 Response Time

Response time, lag time or clearance time are measures of how long (e.g., years) a chemical can reside in any one or more environmental media from the time global emissions to the environment have ceased (e.g., due to regulatory action, product lifespan, economic viability). This is of particular concern for mobile chemicals, as exposure in remote sensitive areas can be delayed, resulting in potential risk to organisms well after regulatory actions are in place. The issue of response time has been presented by ECCC and Health Canada to the CMP Science Committee as a subtopic for improving fate and exposure assessment in Canada (Government of Canada 2019), where the departments noted a need to better consider this aspect of persistence in chemical evaluation.

Response time was used by Gouin and Wania (2007) in a global distribution model (GLOBOPOP) to estimate the "lag-time" between the point at which chemical emissions begin to decrease for 96 hypothetical chemicals globally and the beginning of decline in a chemical's Arctic Contamination Potential.[18] The authors concluded that slow oceanic transport of very persistent chemicals to the Arctic requires a 10-year half-life to reach a significant lag in time for exposure to Arctic species. The authors further conclude that exposure concentrations will only increase over time and that "swift" regulatory action is needed for substances with long response times. Similarly, "clearance time" has been advocated by Stroebe et al. (2004) using overall persistence (P_{ov}) because "*P_{ov} should represent the long-term clearance of a chemical from the environment after the stop of emissions and not the fate at steady state.*" Overall persistence must be calculated using a multimedia model. It is often used synonymously with "residence time" and does not account for losses from "advection" (e.g., sediment burial, transport out of the model environment) nor dilution. Its value is governed by two principle factors: the amount of substance residing in a medium (based on a default emission to the environment) determined by fugacity relationships and the reaction rate (i.e., degradation rate) of a chemical in a specific medium. Finally, Wegmann et al. (2008) use the "temporal remote state" concept as a measure of response time: "*The temporal remote state is the situation in which all releases of chemical have ceased, and overall loss is controlled by the slowest responding compartment, i.e. the compartment where the chemical's half-time for removal is longest.*"

Here, we define P_{ov} as the sum of all medium-specific half-lives (hours) weighted by the mass fraction of the chemical in a medium (based on Pennington 1997). P_{ov} is calculated as:

$$P_{ov} = 1/\left(\frac{f_a}{\tau_a} + \frac{f_w}{\tau_w} + \frac{f_s}{\tau_s} + \frac{f_{sed}}{\tau_{sed}}\right)$$

[18] An immediate and a long-term Arctic Contamination Potential (ACP) as defined in Wania (2003) is the fraction of the total amount in global surface media that is in the Arctic after 1 and 10 years of steady emissions with a generic zonal distribution, respectively.

Where,

P_{ov} = persistence, overall; f = mass fraction distributed in medium (%); τ = medium-specific half-life (hours or days); a = air, w = water, s = soil and sed = sediment.

The mode of entry into the environment (i.e., air, water, soil) has a significant impact on the mass-balance and fate of a chemical, including calculation of P_{ov}. This is because the receiving medium initially contains the largest mass of the chemical and influences the subsequent behaviour and distribution of the chemical. Webster et al. (1998) explain this concept in detail. Ideally, altering the emission rate according to known mode of entry into the environment (e.g., 20% water, 75% soil, 5% air) would represent a more realistic fate and mass-balance of the chemical in the environment and thus P_{ov}. Regional scale fate and environmental concentrations estimated by the model would thus reflect the known mode(s) of entry. However, detailed knowledge of use pattern (e.g., downstream use) is generally not available for DSL substances (Government of Canada 2019). Therefore, P_{ov} is estimated using RAIDAR v3.0 for a "mostly water emission" (previously described in section 3.4.3.4), given that most industrial emissions are to water, but losses from water to air during industrial processing involving water are also likely.

RAIDAR calculates P_{ov} as described above. Similar to Stroebe et al. (2004), response time in ERC2 is also connected to P_{ov} for relative chemical comparison purposes; it is simply calculated as the P_{ov} multiplied by five. After a time period of five half-lives or P_{ov}s, given first order rate decay theory, the remaining chemical mass fraction in any system is approximately 3%. Consequently, response time calculated using this method can also be regarded as an approximate measure of the full lifetime of the chemical in the total environment. Appendix VIII describes the rules used for exposure classification of P_{ov} and the rules for confidence scoring while Figure 19 below shows the relative distribution of ERC2 substances according to exposure classification outcomes for response time.

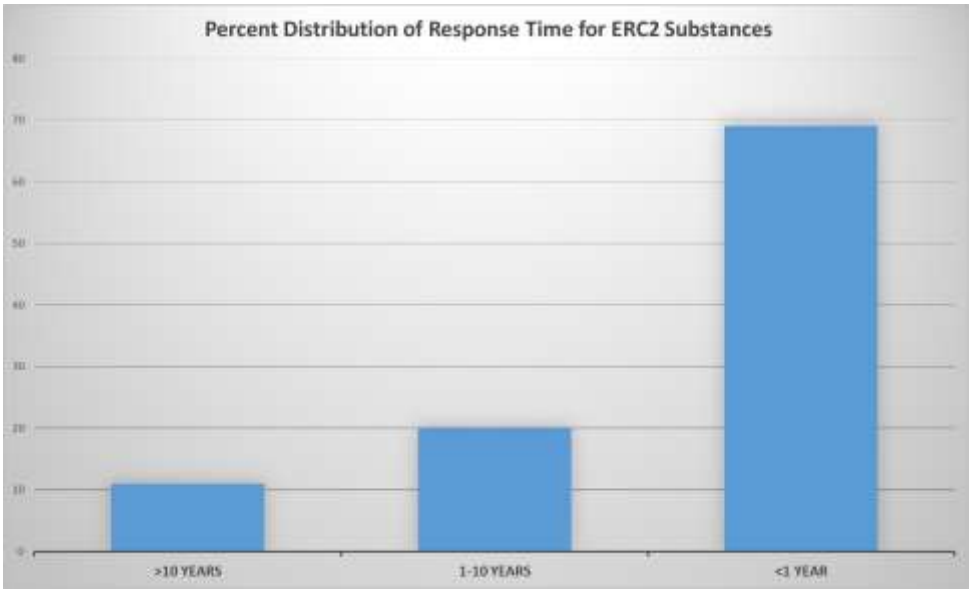


Figure 19: Percent distribution of response times calculated for 12 200 ERC2 substances according to the exposure classification rules for response time

3.5.3.2 Mobility

Mobility is also a key parameter determining substances of concern using the ERC approach. Evidence of chemical contamination in remote areas of the globe, both measured and predicted, has been a regulatory driver for prioritizing and assessing the risk of substances in Canada (e.g., categorization of the DSL) and internationally (e.g., UNEP POPRC, UNECE CLRTAP)[19] in the last few decades. More recently, the identification of persistent, mobile and toxic (PM and PMT) substances has received a great deal of attention from both academic and regulatory agencies as previously discussed in section 2.4. There are now European Union-wide projects specifically oriented to addressing PMT substances (e.g., European Commission Horizon 2020 PMT[20]). In the European Union context, mobility can include groundwater transport as well as surface water and air transport. ERC2 does not include consideration of groundwater transport, focusing instead on surface waters and air. Groundwater transport concerns are considered on a case-by-case basis in risk assessments conducted by the Government of Canada.

The ability of a chemical to persist in a mobile medium such as air or water and be transported over long distances via that medium is a key factor for determining the spatial extent of exposure in ERC2. However, dispersion of the chemical via the medium is equally important. As described in ECCC, HC

[19] United Nations Environment Program Persistent Organic Pollutant Review Committee. United Nations Economic Commission for Europe's (UNECE) Convention on Long-range Transboundary Air Pollution

[20] https://ec.europa.eu/programmes/horizon2020/en

(2019), chemical risk assessment often assumes that predicted environmental concentrations reflect steady-state condition emissions that will not change over time. Thus, they are a "snapshot" of exposure extrapolated over time and space. Close to the point of release, given a chemical that is continuously present, the persistence of the substance is largely irrelevant because of the short time from release to exposure. This can be interpreted to mean "*concentrations, exposures, and risk are constant with time, but variable spatially*" (Mackay et al. 2014). In instances where the time of exposure is long, the residence time determined by degradation, advection (transport) and dilution in the environment becomes a major determinant of exposure, particularly at distance from source. When these properties of a chemical are coupled with a toxicological profile that suggests a substance can cause adverse effects (particularly poorly or non-reversible effects) at very low levels of exposure, chemicals that threaten the planetary boundary emerge. The effect of changes in chemical emissions on far field concentrations as a function of residence time ("distant residence time") and mobility is discussed in detail by Mackay and Reid (2008), and Reid and Mackay (2008).

Mobility in ERC2 is determined using a multimedia model where "characteristic travel distance" (CTD) is calculated using the RAIDAR model. CTD is a "transport-oriented" metric, meaning there is no specific remote environment considered and calculations are based on a travel distance from source emissions (e.g., in kilometers). This is different from a "target-oriented" metric, where a remote area such as the Arctic is used a target environment in a model (e.g., Wania 2003). CTD is defined as the distance at which the concentration of a chemical in air or water is reduced to 37% from degradation and partitioning to other media (Bennett et al. 1998; Beyer et al. 2000). This means that slightly more than approximately one-third of the chemical mass will travel further than indicated by the CTD.

RAIDAR v3.0 uses the CTD method for air or water first proposed by Beyer et al. (2000):

$$CTD_{air} = P_{ov} \times F_{air} \times \text{wind speed}$$

$$CTD_{water} = P_{ov} \times F_{water} \times \text{water speed}$$

Where P_{ov} is the overall persistence of the chemical (as discussed in the previous section) and the chemical mass fraction (F) in air and water determined according to the mass-balance outcome of the RAIDAR model. CTD can be simplified this way because emissions are 100% to air (CTD_{air}) and 100% to water (CTD_{water}) using two separate modelling scenarios (equation 9 in Beyer et al. 2000). The default wind and water speeds used for CTD estimates were set at 14.4 and 0.072 km/h, respectively. The CTDs in ERC2 were calculated separately for water, based on 100% emission to water and air. Appendix IX describes the rules for exposure classification and confidence scoring of chemical mobility in water and air. Figure 20 below gives the relative distribution of CTDs in water and air according to exposure classification rules outlined in Appendix IX showing very few chemicals in ERC2 have <5 kilometers of travel distance (<1%).

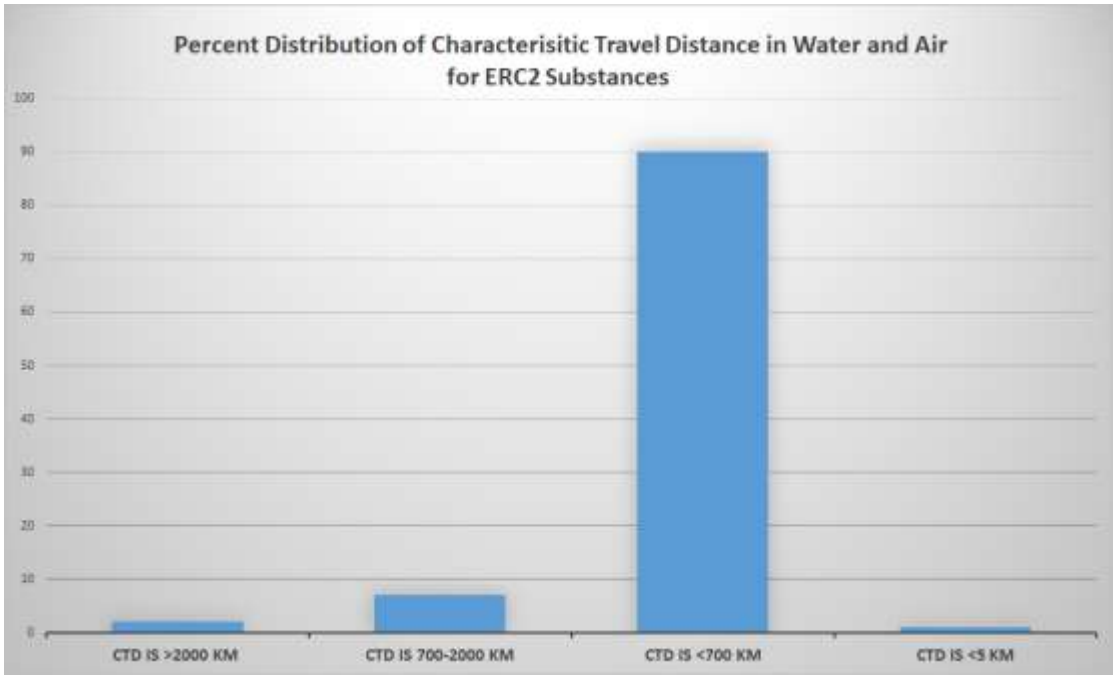


Figure 20: Distribution of characteristic travel distance (CTD) calculated for 12 200 ERC2 substances in water and air

3.5.3.3 Emission Pattern

In addition to persistence, the quantity (mass) of chemical entering the environment at any given time, the rate at which it enters (emission rate) and how it enters (mode of entry) are critical parameters for determining the likelihood of organism contact with a contaminant in the environment. These are the main elements considered by ERC2 under the emission pattern descriptor, the exposure classification and confidence rules, for which, are described in Appendix X.

Chemical quantity is determined using annual tonnage data (kilotonnes per year) from import and manufacturing survey data for the pure substance (i.e., not the quantity of products or articles in which it is contained) in Canada. However, ~97% of the substances in ERC2 have not been surveyed for annual tonnage since 1986 because these substances were previously not determined as priorities for assessment during categorization of the DSL from 1999-2006. Consequently, unless more recent data was available, chemical mass in ERC2 was estimated by extrapolating from 1986 DSL tonnage data. A scaling factor was determined based on a comparison of more recent CEPA section 71 surveys[21] tonnage data for ~1700 DSL substances with the corresponding mean annual tonnage reported in 1986.

[21] Survey conducted under Section 71 of CEPA: 2012, 2014, 2015, 2017

Attempts were also made to extrapolate tonnage information from more recent REACH and TSCA tonnage surveys using population data; however, little CAS RN overlap existed between inventories and no correlation with population density was apparent for substances that did overlap.

It is acknowledged that extrapolated estimates of chemical mass are perhaps the most uncertain descriptor in ERC2. Many of the ERC2 substances may in fact no longer be in commerce in Canada. Based on experience with information collected section 71 of CEPA from 2012-2017, this number can range upwards of 40% not in commerce. In addition, confidence with chemical quantity reported to ECCC several years ago reflect a snapshot in time and can also be said to have high uncertainty due to temporal trends. A probabilistic uncertainty analysis would better evaluate the degree of uncertainty with these data. However, for the purposes of prioritization using ERC2, the confidence score given to tonnage data reflects this uncertainty.

Little is known about the rate of emission (e.g., kg/h) for substances in ERC2. For relative chemical by chemical comparisons, 100% of the chemical mass (kt/yr) is emitted to the regional environment (i.e., homogeneously distributed over 100 000 km/sq.) as described by the RAIDAR model. This is a conservative regional scale emission pattern because it assumes no loss of the total chemical mass during substance use (e.g., formulation in products) or removal by sewage treatment before entering surface waters. This rate of emission and mode of entry best describes emissions to the environment from use patterns that release chemicals in a very dispersive pattern (e.g., from households or multiple sewage treatment plants) or from long-lived mobile chemicals that reach steady-state regional concentrations from fewer point source emissions. Local point source emission patterns are thus not directly included in ERC2's emission pattern. However, the margin of exposure descriptor outlined in section 3.5.3.5 was designed to address this difference in scale of emission pattern.

The mode of entry into the environment (i.e., air, water, soil) has a significant impact on the mass-balance and fate of a chemical because a receiving medium initially contains the largest mass of the chemical and influences the subsequent behaviour and distribution of the chemical. Webster et al. (1998) explains this concept in detail. Ideally, partitioning of the emission rate according to known mode of entry into the environment (e.g., 20% water, 75% soil, 5% air) would represent a more realistic fate and mass-balance of the chemical in the environment. Regional scale fate and environmental concentrations estimated by the RAIDAR model would ideally reflect the known mode(s) of entry. However, detailed knowledge of use pattern (e.g., downstream use) is generally not available in Canada for DSL substances. Except for CTD in air, mode of entry for calculation of all exposure descriptors and tissue residue concentrations using the RAIDAR model was thus set to "mostly water" and soil. "Mostly water" refers to the adjustment of the chemical mass emitted to water from losses to air during chemical processing and release to water treatment systems using the air-water partition coefficient (K_{aw}). Mass-balance simulations in soil were simulated using 100% of the chemical mass emitted to soil. Figure 21 shows the relative distribution of chemical quantity data based largely on the extrapolated values from the 1986 DSL quantity data.

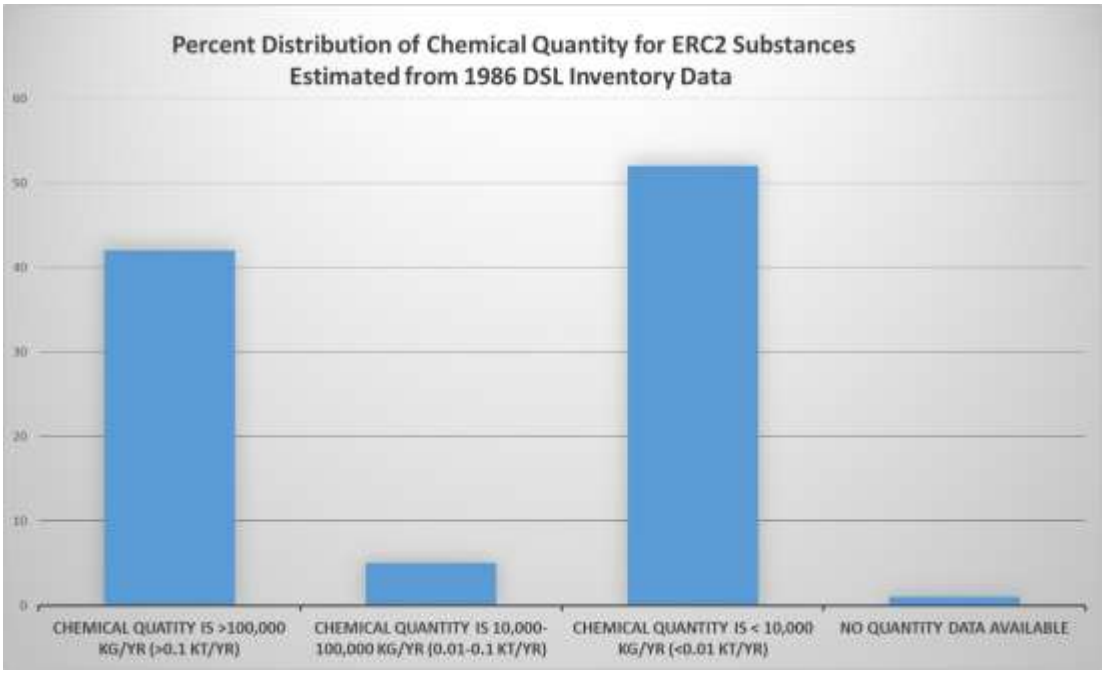


Figure 21: Chemical quantity distribution (kt/yr) for ~12 200 substances in ERC2 based on extrapolated quantity data from the DSL in 1986

Exposure classification of emission pattern relies on chemical quantity and P_{ov} for water or soil, depending on mode of entry to the environment. P_{ov} is used as a means of weighting chemical quantity such that high volume chemicals that are quickly degraded in the environment (e.g., short chain acids and alcohols, aliphatic hydrocarbons) will not be classified as high concern based on quantity alone while lower quantity substances with high P_{ov} will receive higher concern outcomes. This combination of variables was also used in ERC1 (ECCC 2016) for the same reasons.

3.5.3.4 Food Web Exposure

Food web exposure is determined by integrating the properties of persistence and bioaccumulation using the RAIDAR model environment. This descriptor estimates the degree of exposure from direct contact and food web transfer of contaminants in vulnerable food web species. This is accomplished using the RAIDAR exposure assessment factor (EAF), which is the concentration in sensitive food web species based on a default or unit emission rate to the default environment of the model (e.g., 1 kg/h). In other words, EAF "*quantifies the ability of the environment to deliver the specific chemical to the most vulnerable organism in the defined environment and food web*" (ARC 2018). The highest EAFs identify the species that are most vulnerable to contamination, but not necessarily vulnerable to toxic effects. Therefore, the EAF is similar to the RAIDAR HAF, but leaves out the toxicity component.

The EAF was included as a general food web exposure descriptor in addition to the HAF to relate effects evident from the ERC2 hazard profile not captured by the internal effects concentration (IEC) used to calculate the HAF. The combined fate and bioaccumulation models in RAIDAR calculate the EAF for each chemical and for each representative species in the model. The rules for classification and confidence scoring this descriptor are described further in Appendix XI. Figure 22 shows the distribution of RAIDAR EAF across ERC2 substances using the rules outlined in Appendix XI and generally shows a lower concern for food web exposure for the ERC2 DSL substances which were not categorized as P&B in 2006.

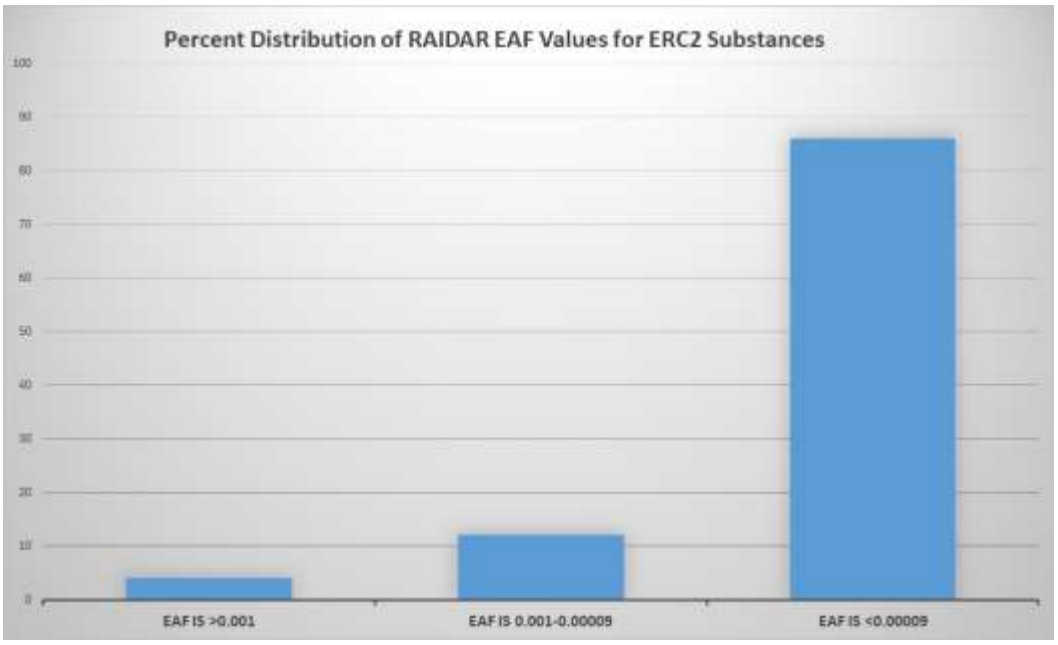


Figure 22: Percent distribution of RAIDAR EAF values for 12 200 ERC2 substances

3.5.3.5 Margin of Exposure

Margin of exposure (MoE) is a commonly used concept in human health risk assessment. It refers to the difference between a critical level of effect, such as a predicted no effect concentration (PNEC) and the measured or predicted concentration ecological receptors are exposed to in the environment (MEC or PEC). It is essentially a risk quotient and in the RAIDAR model is known as the risk assessment factor (RAF). In ERC2, MoE is calculated as the ratio of the critical rate of emission, E_C (kt/yr), and the actual rate of emission, E_A (kt/yr), estimated by ECCC using 1986 DSL data. The E_C is the rate of emission (to water or soil) that results in a concentration exceeding the mode of action-based IEC in the RAIDAR regional fate and food web model. E_C is calculated using the default unit rate of emission E_U (1 kg/h) in the RAIDAR model as described previously and is calculated for the most vulnerable species in aquatic or terrestrial food web.

MoE calculations using the E_C were based on an emission to water. MoE is used as a "check" mechanism of low exposure potential driven by low concern outcomes from other exposure descriptors. Importantly, the check mechanism can capture low quantity substances that are not highly mobile or persistent, yet still present some concern in aquatic food webs. MoE is calculated using a regional fate simulation; it may therefore not capture concerns at the local scale. A MoE of less than 1000 is used to flag substances that may present this concern.

During the development of ERC1 (ECCC 2016), ECCC conducted a comparison of outcomes of risk classification using the ERC approach which, as in ERC2, uses the RAIDAR regional-scale model, and a "local scale verification" using the traditional PNEC/PEC approach. The comparison showed a high degree (>90%) of correlation between the two approaches for identifying high and low concern risk classifications. The local scale verification used in ERC1 has therefore been incorporated into the MoE descriptor for ERC2. It is again emphasized that MoE values (or PNEC/PEC quotients if used) are ~99% reliant on extrapolated chemical quantity and therefore remain an uncertain descriptor in ERC2. Confidence assignments reflect the uncertainty associated with chemical quantity data (Appendix X) and extreme property of fugacity warnings from the RAIDAR model (e.g., Appendix IX). Currently, ~6% of ERC2 substances have been flagged for MoE concerns. Updates of chemical quantities from surveys conducted for ERC2 chemicals will help reduce this uncertainty.

Bioactivity-exposure ratios (BERs) can also be used as another method to determine MoE based on *in vitro* bioactivity data. BERs express the ratio between an estimated internal effects concentrations (IEC) (mg/kg or mmol/kg) using median bioactivity concentrations (AC50 mg/L or mmol/L), often extracted from the US EPA TOXCAST/TOX21 program (Becker et al. 2015). Toxicokinetic modelling is often required to perform this extrapolation to achieve concentrations on a whole body basis. The IEC is then compared to a whole body internal tissue residue (mg/kg or mmol/kg), also estimated using toxicokinetic modelling, based on expected exposure concentrations in the environment. The ratio between these two internal values becomes the margin of exposure or BER.

ECCC has been active in developing different approaches for calculating fish BERs according to fugacity-based (Mackay 2001; Arnot and Gobas 2004) and physiologically-based pharmacokinetic (Rowland et al. 2017; Stadnicka-Michalak and Schirmer 2019) modelling methods. Currently, however, it is difficult to estimate fish BERs for ~12 200 organic substances in ERC2 in the shorter term. Therefore, fish BERs are not currently used for MoE considerations in ERC2, but will be included as they become available in future updates to ERC2 and applied in a similar manner to MoEs calculated using the RAIDAR food web scenario.

3.5.3.6 Chemical Use Pattern

Information on the known and predicted uses of substance in ERC2 was gathered from three main sources:

- US EPA EXPOCAST Consumer Products Database (CPDat) (Dionisio et al. 2018)
- US EPA EXPOCAST Quantitative Structure-Use Relationship Modeling (QSUR) (Isaacs et al. 2016)

- ECCC manual survey of use patterns for selected substances on the DSL

Use pattern information is not used in ERC2 to classify the exposure concern; it is used as supporting information when forming chemical groups based on functional use (e.g., colorants, antioxidants, fragrance). Use pattern is also useful for understanding the expected emission pattern. For example, it can be expected that down-the-drain uses will result in a more dispersive release to the environment than chemicals used as intermediates in product formulations or are used as laboratory reagents. Information on use pattern from one or more sources was available for almost two-thirds of the ~12 200 substances included in ERC2. Both harmonized and curated use as well as predicted uses were gathered and generated for ECCC by EXPOCAST scientists as well as by ECCC (Table 1). There was considerable overlap of use pattern information between the available data sources likely reflecting the limited information publically available on use patterns.

Table 1: Percentage of ERC2 substances with use pattern data according to data source

Chemical Use Pattern Data Source	ERC2 Substance Coverage (%)
EXPOCAST Harmonized	37
EXPOCAST Curated	37
EXPOCAST Predicted (QSUR)	44
ECCC Literature Search	40

Abbreviations: QSUR, Structure-Use Relationship; ECCC, Environment and Climate Change Canada

4. Classification and Confidence Scoring

Rules for classification and confidence scoring of hazard and exposure descriptors is contained in Appendices III-XI. The following sections describe the routines used to determine *final* classification and *total* confidence scores. It is worth mentioning that the scoring approach used in ERC2 is based on a numerical value, the magnitude of which has been established based on existing empirical knowledge (e.g., potency scales, travel distances), where possible. Some classification scores and all confidence scoring, however, required expert judgement. Judgement was guided by the context of priority setting in Canada, according to the WoE principles discussed in section 2.2. The OECD principles and elements for establishing a weight of evidence for chemical evaluation (OECD 2019) cite that,

"*…numerical scores can be used to weigh lines of evidence, but 'those scores are numerical but not quantitative.' Thus, there is no advantage to a numerical system; in fact it may appear to be more 'rigorous' than is possible (Suter and Cormier 2011)*."

and that,

"*Determining a score is judgement based and context dependent and thus absolute rules or criteria to judge the level of reliability and relevance are not provided here and should be developed by individual agencies.*"

Accordingly, ERC2 has established a scoring scheme that involves a numerical value which was translated to categorical outcomes (e.g., very high, high, moderate, low, very low) for interpretation in a regulatory context. Scoring is described in detail below and in the associated appendices for transparency purposes, which is one of the most important principles of weight of evidence (OECD 2019).

4.1 Hazard

Classification of hazard relies largely on the potency of a substance given the interactions described in previous sections, the rules outlined in Appendices III-VII, and the hazard space defined for ERC2 (section 3.4.1). Figure 23 illustrates the logic workflow for hazard classification and confidence scoring. Classification scores can range from three to one, indicating the following levels of hazard concern:

- Class 3 hazard score indicates a high level of concern
- Class 2 hazard score indicates a moderate level of concern
- Class 1 hazard score indicates a low level of concern

Classification scores are generated for each hazard descriptor (Figure 23), except for cumulative toxicity which cannot be classified at this time (see section 3.4.3.5). Identical hazard descriptors are used to determine all levels of hazard classification. This process results in *descriptor level classifications* of hazard according to the rules in Appendices III-VII. Differing or common descriptor classification scores can result depending on the type of hazard interaction the chemical can have and the potency of this interaction. A class 3 (high) or class 2 (moderate) descriptor level outcome requires no more than one set of hazard descriptor rules (i.e., "any" in Figure 23) to be met, otherwise hazard classification defaults to class 1 (low) hazard. Similarly, *descriptor level confidence scores* (Table 2) are calculated for each hazard descriptor according to the confidence rules also described in Appendices III-VII. All hazard descriptors are accepted to be equally significant for prioritization purposes.

Final classifications of hazard are based on the highest score for one or more hazard descriptors as outlined above and shown in Figure 23 (final hazard outcomes). A *total hazard confidence score* is calculated for each final hazard classification based on the sum of each descriptor level confidence score when the descriptor classifications agree with the final hazard classification (Table 2). For example, if the final hazard classification is triggered only by a class 3 result for receptor-mediated interactions, then the total hazard confidence score is based only on the confidence score for receptor-mediated interactions. Conversely, if all descriptors result in class 3 hazard outcomes, then the total confidence score is based on the sum of all individual hazard descriptor scores. Lower confidence scores can result from lack of consensus between hazard descriptor classification outcomes as well as significant data gaps. Lack of descriptor consensus represents ~22% of low confidence scores compared with ~88% low confidence resulting for lack of data (mostly *in vivo* data). Finally,

Descriptor

hazard classification outcomes are used with exposure classification outcomes to determine the risk matrix (section 4.3). Total hazard confidence scores are added to exposure confidence scores to produce a final risk confidence score.

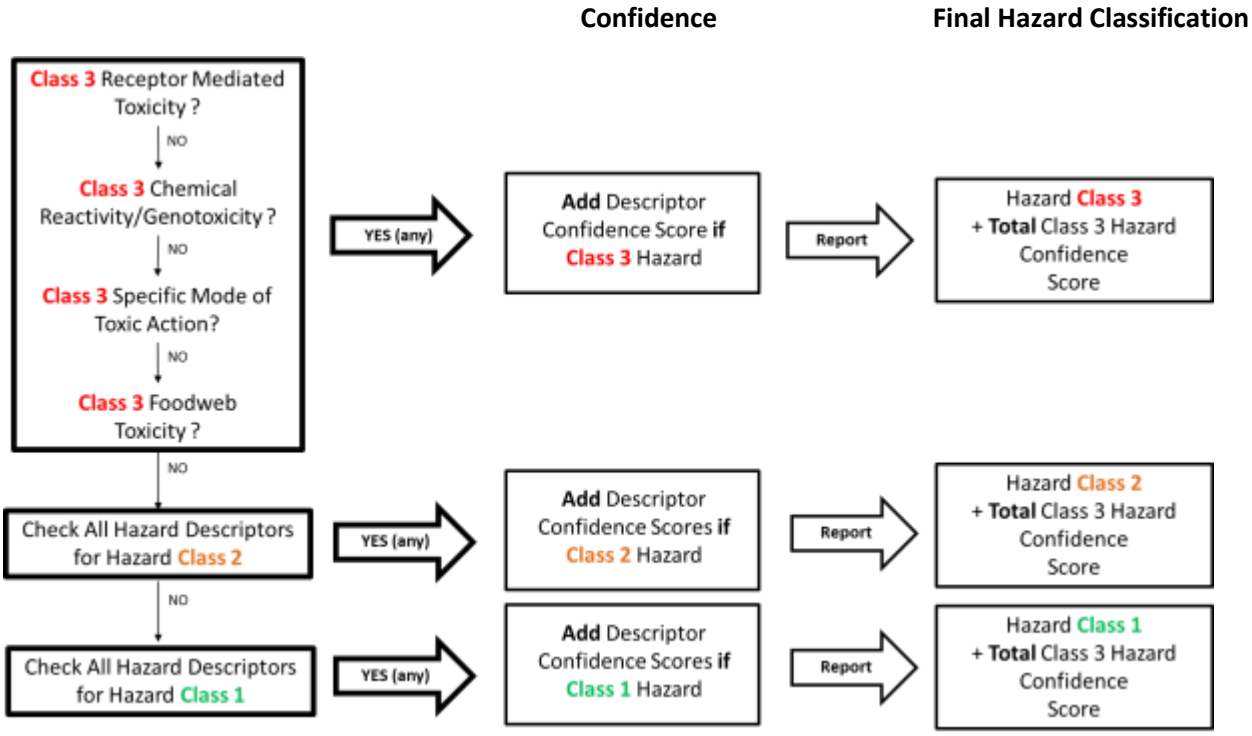


Figure 23: Logic workflow for determining all hazard classifications and hazard confidence scoring

Table 2: Descriptor level confidence assignments for hazard

Descriptor	Maximum Descriptor Confidence Score
Receptor-Mediated Toxicity	88
Chemical Reactivity/Genotoxicity	93
Mode of Toxic Action	15
Food Web Toxicity	10
Cumulative Toxicity	0 (not included in total hazard confidence score)

The descriptor confidence weighting was determined according to data origin/type and abundance. This impacts overall confidence weight assignments for each descriptor in Table 2. For example, for chemical reactivity/genotoxicity, five types of data are used (*in silico, in chemico, in vitro*, mammalian and aquatic *in vivo* data), where many sources of data are used to parameterize each data type and therefore greater classification consensus between all data sources must result for higher descriptor confidence scores. In contrast, for food web toxicity, a single *in silico* data source is used to derive output for classification and therefore is regarded as a single data source with no consensus possible within the

descriptor. The descriptors for receptor-mediated toxicity and chemical reactivity/genotoxicity assign confidence according to the data hierarchy described in section 3.4.2. To avoid low confidence assignments to *in vivo* data used as the only source of evidence for hazard classification due to lack of data consensus or data gaps, a default confidence score of 26 is assigned when multiple data are available or 21 when only single values are available. These values were selected to ensure a moderate level of confidence or better is given when the total hazard confidence score is summed and only in vivo data are used to classify hazard (i.e., due to lack of consensus with other data). In such cases, plausible mechanistic-causality relationships are "unconfirmed" and noted as such in final results along with all other plausible target interactions (Appendix XII).

A maximum total hazard confidence score of 206 can be achieved using the confidence weighting scheme in Table 2. Table 3 shows how the range of confidence scores are partitioned to assign a confidence category. The category thresholds were assigned based on examination of the relative distribution of substances within each category and considering the maximum possible confidence score of 206.

Table 3: Categorical assignment of total confidence score for hazard

Hazard Confidence Score	Confidence Category
1-10	Very Low
11-25	Low
26-60	Moderate
61-100	High
>100	Very High

4.2 Exposure

Exposure classification is conducted in a very similar manner to that described above for hazard. Classification of exposure is driven by the likelihood of organism contact with a contaminant over a varying spatial and temporal scales of exposure according to the classification rules outlined in Appendices VIII-XI and considering the exposure space defined for ERC2 (section 3.4.2). Figure 24 illustrates the logic workflow for exposure classification and confidence scoring. Classification scores can range from three to one, indicating the following degree of exposure concern:

- Class 3 exposure score indicates a high level of concern
- Class 2 exposure score indicates a moderate level of concern
- Class 1 exposure score indicates a low level of concern

Classification scores are generated for each exposure descriptor (Figure 24), except for margin of exposure which is used as a verification mechanism for low exposure concern outcomes from other exposure descriptors. This process results in *descriptor level* classifications of exposure according to the

rules in Appendix VIII-XI. Differing or common descriptor classification scores can result, depending on emission pattern differences and fate in the environment. A class 3 (high) or class 2 (moderate) descriptor level outcome requires no more than one set of exposure descriptor rules (i.e., “any” in Figure 24) to be met, otherwise exposure classification defaults to class 1 (low) exposure. Similarly, *descriptor level confidence scores* are calculated for each exposure descriptor according to the confidence rules also described in Appendices VIII-XI.

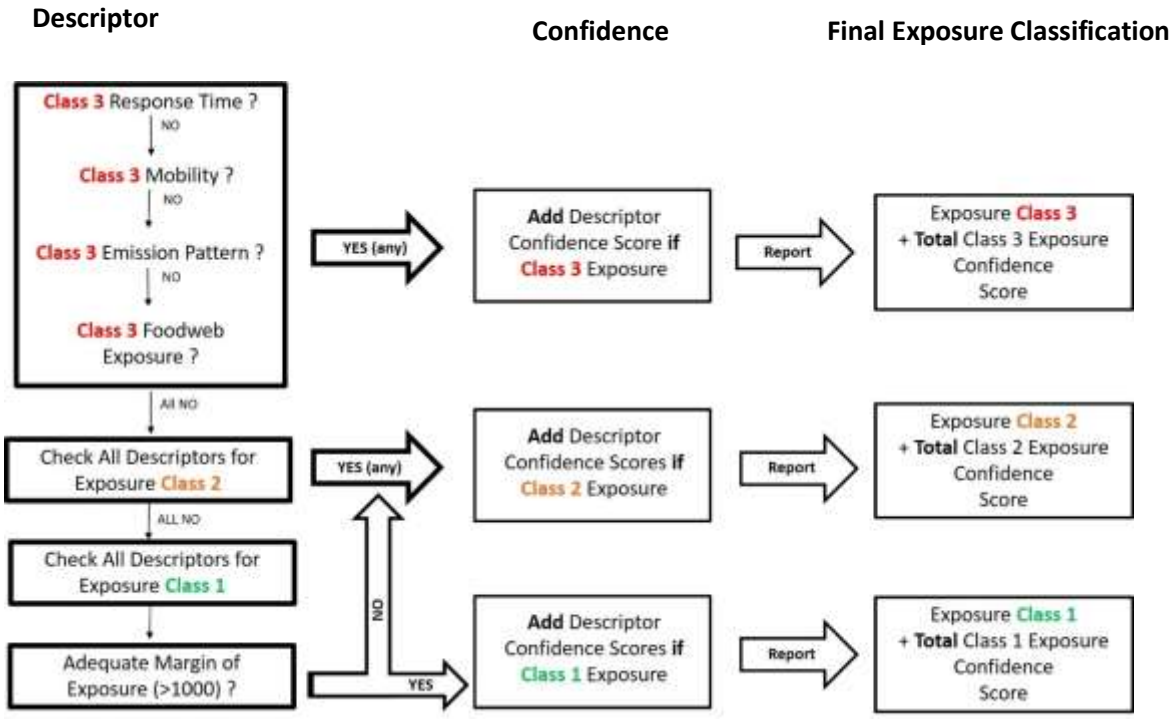


Figure 24: Logic workflow for determining all exposure classification and exposure confidence scoring

Identical to hazard, final classifications of exposure are based on the highest score for one or more exposure descriptors as outlined above and shown in Figure 24 (final exposure outcomes). A *total exposure confidence score* is calculated for each final exposure classification using the same procedure for hazard. However, the margin of exposure descriptor is used in a verification step to confirm low exposure concern in the environment and contributes to the total exposure confidence score only when triggered. Lower confidence scores can result from lack of consensus between exposure descriptor classification outcomes as well as significant data gaps. Descriptor data was available for ~92% of ERC2 substances and therefore data gaps were not a significant contributor to lower confidence scores. Lower confidence scores were driven by lack of descriptor concordance, noting that this is likely attributable to the greater sensitivity of exposure classification to fate and behaviour in the environment across the chemistries within ERC2 compared with hazard descriptors. Exposure classification outcomes are used with hazard classification outcomes to determine the risk matrix (section 4.3). Total exposure confidence scores are added to hazard confidence scores to produce a final risk confidence score.

Finally, it is emphasized that 97% of substances in ERC2 have not undergone a chemical quantity inventory update since 1986. As noted previously, results from inventory updates conducted for the years 2012-2017 and from IRAP updates suggest that 40-50% of substances surveyed may no longer be in commerce in Canada. Substances prioritized for future regulatory assessment activity based on ERC2 may require more recent chemical quantity data via inventory updates to confirm their exposure profile. Substances determined no longer to be in commerce in Canada by ECCC and Health Canada via these updates will see their ERC2 exposure (and risk) classification scores greatly lowered such that they may no longer be recommended as priorities for risk assessment.

A maximum total exposure confidence score of 50 can be achieved using the confidence weighting scheme in Table 4. The maximum score of 50 reflects an equal weighting of exposure descriptors because all descriptors make use of *in silico* data only for exposure classification and chemical quantity is currently extrapolated for 97% of ERC2 substances.

Table 4: Descriptor level confidence assignments for exposure

Descriptor	Maximum Descriptor Confidence Score
Response Time	10
Mobility	10
Emission Pattern	10
Food Web Exposure	10
Margin of Exposure	10 (only used when triggered)

Using the maximum total confidence score for exposure, the following categorical confidence assignments were selected in Table 5. The category thresholds were distributed equally across the range of exposure confidence scores.

Table 5: Categorical assignment of total confidence score for exposure

Exposure Confidence Score	Confidence Category
1-10	Very Low
11-20	Low
21-30	Moderate
31-40	High
41-50	Very High

4.3 Risk

Ecological risk classification outcomes are determined using a risk matrix by combining the classification scores for hazard and exposure. A similar approach was used in ERC1 (ECCC 2016). Table 6 shows the possible risk outcomes based on the hazard-exposure score combinations. The categorical designation of risk outcomes (i.e., high, moderate, low) according to the matrix reflect a weight of evidence scoring concept (OECD 2019), where there is equal probability of being assigned to one of the three categorical outcomes, reducing bias to any one risk category. The risk outcomes were selected based on the regulatory context and need for prioritizing ERC2 substances for assessment considering various types of concern (e.g., identification of endocrine active substances, regrettable substitution, chemicals of global concern). Therefore, risk outcomes reported according to the matrix balance the need to identify chemicals of concern, while minimizing both false negative and false positive conclusions to the degree that the evidence is available to do so (see risk confidence scoring below).

Table 6: Matrix used for determining risk outcomes based on hazard and exposure classification scores.

Risk Matrix		Hazard		
		1	2	3
Exposure	1	low	low	moderate
	2	low	moderate	High
	3	moderate	High	High

One interpretation of the risk matrix can show the impact of exposure on risk classification. A combination of high exposure (3) with a moderate or high hazard score (2 or 3) results in a high concern risk classification. However, low exposure (1) will produce at most a moderate risk outcome for high hazard (3) substances. A high exposure (3) and low (1) hazard score results in a moderate risk classification as well. This is a precautionary risk classification assignment given a high potential for exposure, and unknown and unforeseen hazards.

Risk classification will have ranging confidence scores based on the sum of hazard and exposure classification confidences. Confidence scores reveal the degree of data consensus as well as data gaps. Table 7 summarizes the scoring routines used for categorical risk classification and risk confidence outcomes, as well as the percentage of ERC2 substances for each classification outcome. In Table 7, 70% of all ERC2 substances were scored with a moderate or high risk outcome according to the risk matrix above (Table 6) and risk score in Table 7. This result does not account for risk confidence or risk severity,

which are factored into the overall outcome for future regulatory activities (section 6.0). Table 7 also shows that 67% of all ~12 200 ERC2 substances have low or very low confidence outcomes.

Table 7: Summary of risk classification and risk confidence outcomes for ~12 200 ERC2 substances

Risk Classification Scoring			
Rule Description	**Risk Score**	**Risk Classification**	**Percentage of ERC2**
Risk Score is:	5-6	High	39
Risk Score is:	4	Moderate	31
Risk Score is:	1-3	Low	21
Insufficient data for scoring	NA	NA	9
TOTAL			100
Risk Confidence Scoring			
Rule Description	**Confidence Score**	**Confidence Category**	**Percentage of ERC2**
Confidence score is:	>150	Very High	<<1
Confidence score is:	101-150	High	4
Confidence score is:	51-100	Moderate	20
Confidence score is:	26-50	Low	33
Confidence score is:	1-25	Very Low	34
Insufficient data for scoring		NA	9
TOTAL			~100

Abbreviation: NA, not available

5. Severity

Chemicals can simultaneously cause effects via more than one mechanism or pathway and organisms can be exposed at different temporal and spatial scales, from different routes of exposure and different emission patterns. In situations where higher volume substances are capable of more than one toxicity interaction and more than one route of exposure at different spatial and temporal scales, the potential impact to organisms in the environment from such substances can be said to be more “severe”. Because moderate and high risk outcomes can be triggered by any one hazard and exposure descriptor, severity is used as a method to weight the possible risk outcomes in the risk matrix and can be regarded as a measure of the scale of risk. The weighting is performed using severity scores (Table 8). Severity scores are first calculated separately for hazard and exposure. Risk severity is then a function of summing the hazard and exposure severity scores. Table 8 below describes the rules and scoring routine used for severity scoring for all ~12 200 substances in ERC2 with results.

Table 8: Rules and scoring routine used to assess substance severity

	Hazard and Exposure Severity		
Rule No.	**Rule Description**	**Score**	**Severity Assessment**
1	Class 1 (low) hazard or exposure outcomes	0	Very low
2	Class 2 (moderate) hazard or exposure outcome	1	Low
3	A single Class 3 (high) hazard or exposure outcome	2	Moderate
4	Two or three Class 3 (high) hazard or exposure outcomes	3	High
5	Greater than three Class 3 (high) hazard or exposure outcomes	4	Very high
	Risk Severity		
Rule No.	**Rule Description**	**Score**	**Severity Assessment**
1	Sum of hazard and exposure severity scores	0 to 8	Hazard + exposure

The scoring routine in Table 8 is weighted to give greater insight into high risk outcomes such that a higher severity score is achieved when multiple Class 3 (high) hazard or exposure classifications result for a substance. This allows ECCC to parse out risk outcomes according to various regulatory needs and drivers. Severity is often correlated with confidence, where higher confidence scores often can result in higher severity scores and vice versa. But this is not always the case because a chemical can have a low confidence score (e.g., relies mainly on *in silico* results) yet achieve a higher severity score due to hazard and exposure descriptor classification consensus. Therefore, confidence is a measure of certainty with classification outcomes whereas severity is a measure weighted risk scale. Both of these ERC2 metrics can be used to integrate risk results for transparent communication of potential future regulatory activities. This is described in the next section of this document.

6. Results

The following two sections provide examples of recommendations for regulatory assessment activity based on risk outcomes and targeting different regulatory priorities (e.g., potential endocrine active substances) that are of current concern to ECCC or forecasted to be of future concern and which may not always be risk-based. The recommendations provided in the examples below are not considered to be definitive for the regulatory program and may be adjusted based on additional considerations outside of the ERC2 approach. The targeted examples do not rely on risk matrix outcomes from ERC2, but rather they rely on the outcomes from combining specific descriptors classifications and confidence scores. In both situations, regulatory activities are related to the level of acceptable uncertainty as described in section 2.2.2 when weighing evidence to inform regulatory decision-making. ERC2 results for all 13 162 compounds is summarized in an MS Excel® spreadsheet, which accompanies this document as Appendix XII. Finally, Appendix XIII provides example summary profile information and

classification outcomes for three substances that represent possible ERC2 risk outcomes as well as possible future regulatory activities for these substances.

6.1 Results Using Risk Matrix

Using a matrix of risk confidence and the sum of hazard and exposure severity scores, various risk-based outcomes can be examined for ERC2 substances with hazard and exposure results (Table 9). Three possible ERC2 outcomes for recommended regulatory assessment activity are envisioned, the results of which can be examined by hazard, exposure or risk, independent of each other using Table 9. This allows ECCC to target key areas of further work according to hazard or exposure confidence and severity scale.

- **No Further Action (NFA) at this Time (green):** Severity is very low in all cases, suggesting none of the classification rules for moderate or high concern have been activated. Confidence ranges from very low to very high. This percentage of ERC2 substances are currently not recommended as a priority for data collection or risk assessment.
- **Data Collection (beige):** Severity and confidence range in scale from low to very high, suggesting additional hazard and/or exposure data should be collected (e.g., via research and monitoring or surveys) to refine ERC2 outcomes.
- **Further Assessment (orange):** Severity and confidence range from moderate to very high, suggesting that this percentage of ERC2 substances are a priority for further risk assessment based on risk classification. If not recently obtained, chemical quantity data should be collected as a priority for these substances in order to confirm/ increase confidence in their exposure profile.

Table 9: Confidence-severity matrix for risk for ~12 200 ERC2 substances

Risk Severity Score	Hazard / Exposure Severity Category	Very Low	Low	Moderate	High	Very High
0	Very Low / Very Low	No Further Action	No Further Action	No Further Action	No Further Action	No Further Action
1	Low / Very Low	No Further Action	No Further Action	No Further Action	No Further Action	No Further Action
1	Very Low / Low	No Further Action	No Further Action	No Further Action	No Further Action	No Further Action
2	Low / Low	Data Collection	Data Collection	Data Collection	Data Collection	Data Collection
2	Moderate / Very Low	Data Collection	Data Collection	Data Collection	Data Collection	Data Collection
2	Very Low / Moderate	Data Collection	Data Collection	Data Collection	Data Collection	Data Collection
3	High / Very Low	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
3	Low / Moderate	Data Collection	Data Collection	Data Collection	Data Collection	Data Collection
3	Moderate / Low	Data Collection	Data Collection	Data Collection	Data Collection	Data Collection
3	Very Low / High	Data Collection	Data Collection	Data Collection	Data Collection	Data Collection
4	High / Low	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
4	Low / High	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
4	Moderate / Moderate	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
5	High / Moderate	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
5	Moderate / High	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
5	Very High / Low	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
6	High / High	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
6	Moderate / Very High	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
6	Very High / Moderate	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
7	High / Very High	Data Collection	Data Collection	Further Assess.	Further Assess.	Further Assess.
NA	NA	Not Available	Not Available	Not Available	Not Available	Not Available

Abbreviation: NA, not available

Table 9 allows ECCC to parse out possible future regulatory assessment activities based on hazard or exposure severity or both, and in overall risk confidence. No further action (NFA) at this time (see green cells in Table 9) is recommended for substances having very low severity outcomes (i.e., low concern risk classifications for both hazard and exposure across all descriptors) regardless of confidence score, which equates to approximately 22% of substances in ERC2 (Figure 25). Data collection (beige cells) is suggested for 63% of all substances in ERC2 (Figure 25), where the severity score is three or lower and emphasizes substances with very low or low confidence scores. Data collection is suggested to refine risk classification results from the current outcomes. Evaluation (further assessment), shown by the orange cells in Table 9, is suggested as the follow-up regulatory activity for approximately 7% of ERC substances (Figure 25). This was determined by setting risk confidence as moderate or higher and risk severity score as four or higher. An exception is made to this general rule to include a small number of substances (~1%) where hazard severity score is three (high) and exposure severity score is zero (very low) and confidence is moderate or higher (Table 9). Finally, 7% of substances had insufficient data for classification under ERC2. These are predominantly UVCB biologicals from natural products (e.g., waxes, oils), where a representative structure was not available for ERC2 computation.

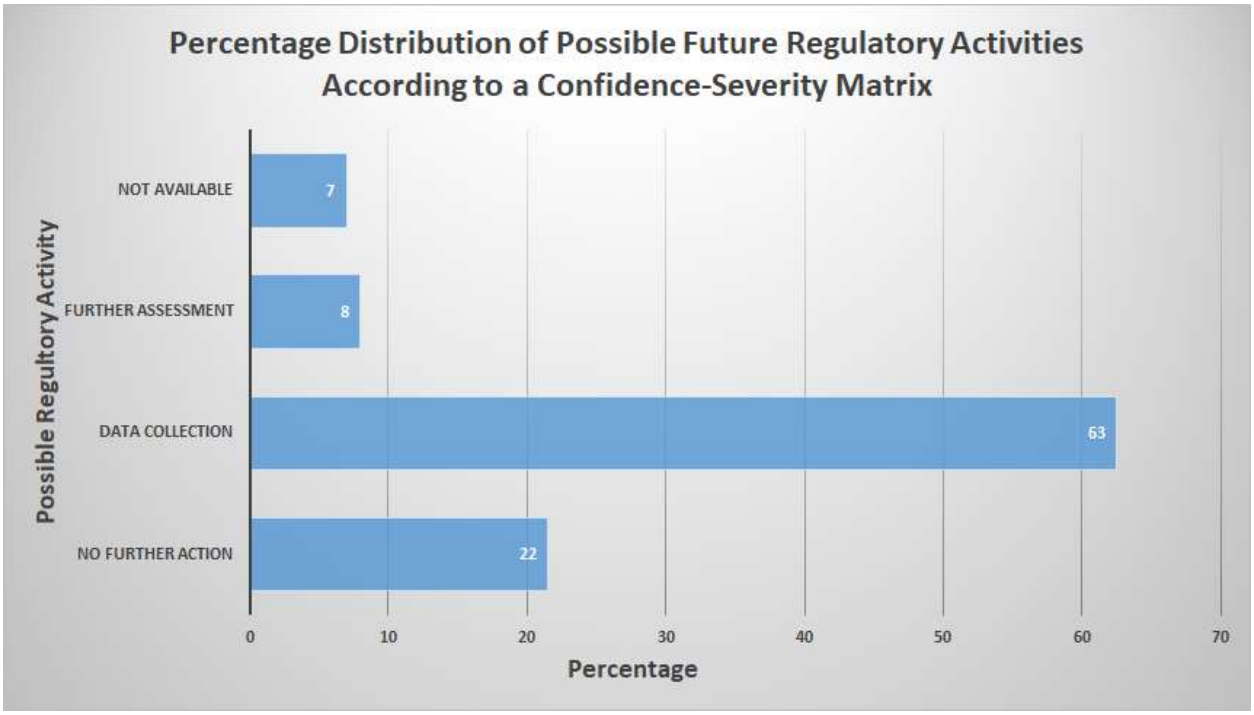


Figure 25: Percentage distribution of ERC2 substances assigned to a future regulatory activity using a confidence-severity matrix for risk outcomes risk

6.2 Results by Targeted Assessment Activity

Combinations of classification and confidence scores from both hazard and exposure descriptors can also be used to identify substances that could be targeted for assessment activities in relation to specific concerns. The following are examples of types of concerns that could be targeted by subsequent assessment activities, along with some estimates of the number of ERC2 substances that could be associated with each concern.

6.2.1 Endocrine Active Substances

Substances that have a higher confidence of being endocrine active according to the potency rules and consensus between the various data types used to describe endocrine activity (i.e., *in silico*, *in vitro* and *in vivo* mammalian and aquatic data) can be isolated using the target interaction from the *in silico* and *in vitro* mechanistic data. Unconfirmed target interactions are not included in this analysis. Table 10 summarizes the percentage of potential endocrine active substances by hazard classification and confidence scores for receptor-mediated descriptor.

Table 10: Distribution of potential endocrine active substances in ERC2 according to confidence score

Potential Endocrine Active Substances (%)			Descriptor Confidence				
			1-5	6-25	26-40	41-60	>60
			Very Low	**Low**	**Moderate**	**High**	**Very High**
Hazard Classification	1	**Low**	33.1	56.0	1.8	0.5	0.3
	2	**Moderate**	2.8	1.2	2.9	1.3	<0.1
	3	**High**	2.4	1.7	4.5	0.9	<0.1

Legend for outcomes: Coloured cells indicate recommended regulatory assessment activities (orange = further assessment; beige = data collection; green = no further action)

Table 10 reveals that just under 18% of all ERC2 substances have potential endocrine activity (moderate and high hazard classifications combined) identified by ERC2's receptor-mediated descriptor regardless of confidence score. Approximately two-thirds of these have low or very low confidence scores, suggesting significant *in vivo* data gaps and lack of potency consensus between data types. Various regulatory activities can be proposed based on the combination of hazard potency (i.e., hazard classification) and confidence as shown in Table 10. Importantly, assessment of a small percentage (10%) of all ERC2 substances with receptor-mediated descriptor results can be recommended for further assessment based on hazard classification for this descriptor regardless of confidence outcome. The data selected to determine moderate and high hazard classification with very low and low confidence score is predominantly driven by *in vitro* and *in silico* results (~69% of the time) with *in vivo* data forming the remaining ~31% of the data selected in this scenario. More than half of the time (~64%) there was lack of consensus among these data types to explain low confidence outcomes (i.e. endocrine activity classification was driven by a single *in silico* or *in vitro* result).

6.2.2 Regrettable Substitution

A "regrettable substitution" occurs when an alternative substance is equally or more harmful to humans or the environment than the substance being replaced. While it remains very difficult for the Government of Canada to identify functionally suitable alternatives due to the lack of intimate knowledge of chemical use processes, ERC2 can identify structurally similar substances with high hazard potential as potentially harmful alternatives.

In Table 11, high hazard substances are identified using the confidence and severity matrix for hazard. This uses all of the ERC2 hazard descriptors contributing to classification results. Chemical quantity is not used in this instance due to the uncertainty in current and future use of the substance as an alternative. However, known and predicted use patterns can be used to group these high hazard substances according to a gross level functional use in order to prove some context for potential common commercial use (Appendix XII).

Table 11 shows the distribution of ERC2 substances according to the confidence-classification matrix for hazard.

Table 11: Confidence-classification matrix for ERC2 hazard outcomes (%) excluding NA results

Hazard Classification		Total Hazard Confidence					
		Very Low	Low	Moderate	High	Very High	NA
1	Low	0.3	1.9	9.4	0.8	0.6	0.0
2	Moderate	18.1	2.7	2.5	0.5	0.5	0.0
3	High	47.2	7.7	5.2	1.7	0.9	0.0
NA	NA						6.8

Abbreviation: NA, not available

The above table suggests that when a hazard classification of moderate and high is used as the primary method for identifying substances that can be regarded as regrettable substitutes, approximately 76% have very low or low confidence outcomes resulting in a recommendation for further hazard data collection to help confirm that these substances are indeed potentially harmful alternatives. If moderate or better confidence with classification is used as a threshold, ~12% of ERC2 substances with hazard classification results could be considered as regrettable substitutes. However, a more rigorous approach considers hazard severity by focusing on substances with a higher probability of adverse effects occurring from one or more toxicological targets, indicated in Table 12 by higher severity scores. Table 12 shows that high hazard substances denoted by further assessment (orange cells) are those associated with low or better hazard confidence scores and moderate or better hazard severity scores. This accounts for ~14% of ERC2 substances with hazard results which can considered as more probable to be regrettable alternatives.

Table 12: Future regulatory activities suggested using a hazard confidence-severity matrix

Final Hazard			Confidence				
			1-10	11-25	26-60	61-100	>100
			Very Low	**Low**	**Moderate**	**High**	**Very High**
Severity	0	**Very Low**	NFA	NFA	NFA	NFA	NFA
	1	**Low**	Data Collection	Data Collection	Data Collection	NFA	NFA
	2	**Moderate**	Data Collection	Further Assess.	Further Assess.	Further Assess.	Further Assess.
	3	**High**	Data Collection	Further Assess.	Further Assess.	Further Assess.	Further Assess.
	4	**Very High**	Data Collection	Further Assess.	Further Assess.	Further Assess.	Further Assess.

Abbreviation: NFA, No further action

6.2.3 Cumulative Risk in Vulnerable Species Populations

ERC2 can be used to identify candidates for cumulative risk assessment (CRA). Use pattern information provides a rationale for forming a functional chemical group (category) for targeted mixtures during assessment using ERC2 output. In the following theoretical example, it is assumed that substances in the targeted mixture co-occur in the environment.

In section 3.4.3.5, the ERC2 approach to CRA was discussed as an early indication for conducting cumulative approaches for targeted mixtures. The approach uses an internal effects concentration (IEC, mmol/kg) and an assessment factor to derive the tissue residue associated with no effects (TR_{NOEC}) and the aggregated tissue residue from the targeted mixture (TR_A) accordingly:

$$CRA_{TARGETED} = TR_{NOEC} / TR_A$$

Using the high hazard substances identified as requiring further assessment from Table 12 above, 133 substances were identified as having a fragrance use pattern using the EXPOCAST curated use pattern information (Appendix XII). Given that the hazard severity index indicates more than one type of toxicity interaction is plausible, more than one mode of action was identified: endocrine active/receptor docking, types of chemical reactivity and narcosis. To determine a cumulative toxicity threshold, an IEC by mode of action is required for each of the four modes of toxic action identified (see section 3.4.3.5). These become "internal eco-thresholds of toxicological concern" or eco-TTC_i (ECCC 2016; Ellison et al. 2021). Lower range CBR50 values from Escher et al. (2011) or consensus TR50s calculated using fish toxicity (Appendix IV) can be used for this purpose. In this example, chemical reactivity is used as the mode of action, which represents 19 of the 133 substances. An IEC of 0.01 mmol/kg for chemical reactivity from Escher et al. (2011) was selected. An assessment factor of 10 was arbitrarily applied to the IEC to derive the TR_{NOEC} for the purposes of this example. If acute fish toxicity values (mg/L) are used to estimate the IEC for the same mode of action, it is suggested that a 95th percentile be used.

Using an emission to water, aquatic mammals have been identified as the most sensitive species based on food web toxicity (Table 13). Chemical fate in the environment and the organism affect the tissue residue concentration, hence mammals were not always identified as the most sensitive species (e.g., also rodents, fish and birds). However, for 84% of the chemicals in this example, aquatic mammals were identified as most sensitive and can therefore be regarded as the most vulnerable species from the standpoint of the most exposed organism in the food web and sensitivity to the mixture's mode of toxic action. Assuming co-occurrence, the aggregate tissue residue concentration (mmol/kg) or TR_A can be calculated as the sum of the tissue residue concentrations for aquatic mammals for the example fragrance use pattern (Table 13).

Tissue residue concentrations for all 19 chemicals were estimated even when aquatic mammals were not the most sensitive species (Chemical No. 5, 12, 15) in order to calculate the TR_A in aquatic mammals from exposure to the entire fragrance mixture. Tissues residues were calculated using the "actual emission" rate (E_A) in the RAIDAR model. E_A was calculated using extrapolated chemical quantities from

1986 data to illustrate this example. Table 13 summarizes the IECs and tissue residues calculated and summed for 19 chemicals with RAIDAR model output.

Table 13: Example cumulative risk estimate for selected ERC2 substances used as fragrances

Chemical No.	TR_{NOEC} (mmol/kg)	TR_A (mmol/kg)	Vulnerable Species	CRA Risk?
1	0.001	3.52519E-09	Aquatic mammal	FALSE
2	0.001	1.67204E-07	Aquatic mammal	FALSE
3	0.001	4.80759E-08	Aquatic mammal	FALSE
4	0.001	1.05111E-07	Aquatic mammal	FALSE
5	0.001	0.009621 (aq. mammal)	Small rodent	TRUE
6	0.001	1.20764E-05	Aquatic mammal	FALSE
7	0.001	2.11294E-08	Aquatic mammal	FALSE
8	0.001	1.29533E-11	Aquatic mammal	FALSE
9	0.001	3.45945E-06	Aquatic mammal	FALSE
10	0.001	1.23832E-05	Aquatic mammal	FALSE
11	0.001	0.003709872	Aquatic mammal	TRUE
12	0.001	0.00038 (aq. mammal)	Piscivorous fish	FALSE
13	0.001	1.54429E-06	Aquatic mammal	FALSE
14	0.001	1.96115E-07	Aquatic mammal	FALSE
15	0.001	0.00073 (aq. mammal)	Avian passerine	FALSE
16	0.001	9.14465E-07	Aquatic mammal	FALSE
17	0.001	9.78419E-08	Aquatic mammal	FALSE
18	0.001	4.48224E-09	Aquatic mammal	FALSE
19	0.001	2.70404E-08	Aquatic mammal	FALSE
All	**TR_{NOEC} = 0.001**	**TR_A = ~0.014**	**Aquatic mammal**	**TRUE**

Abbreviations: NA, not available; TR_{NOEC}, tissue residue associated with no effects; TR_A, aggregated tissue residue; CRA, cumulative risk assessment

Table 13 shows that the cumulative TR_A for aquatic mammals exceeds the TR_{NOEC} by just over one order of magnitude, suggesting that if exposure to the 19 fragrances above occurs simultaneously, then this functional targeted mixture can be considered a candidate for a future CRA. The TR_A is 14-fold higher than the TR_{NOEC} allowing for uncertainty inherent with effects/exposure ratios (risk quotients). However, currently ECCC has no knowledge of the co-occurrence of these fragrances and they are used solely as an example.

6.2.4 Chemicals That Can Affect the Planetary Boundary

Section 2.4 discussed chemicals of concern in 2021 from both an ECCC and international perspective. One of these concerns relates to chemicals that can affect the planetary boundary as described by MacLeod et al. (2014) as a threat from persistent and mobile substances (and their transformation products) that are capable of causing effects which may also be expressed in subsequent generations (i.e., epigenetic inheritance). This type of effect is known to be "irreversible or poorly reversible" and result in a "regime shift" in populations (MacLeod et al. 2014).

To isolate such chemicals among ERC2 substances, specific descriptors from hazard and exposure can be selected. By combining the hazard descriptor for chemical reactivity/genotoxicity and the exposure descriptors for response time and mobility, a focused persistent, mobile and toxic (PMT) approach prevails where toxicity is genetic damage leading to development and reproductive effects and persistence and mobility are examined at various temporal and spatial scales. Table 14 combines results for a sample (n=20) of the high concern (Class 3) outcomes for these three descriptors with their respective confidence and targets for toxicity. In this example, ~2% of substances in ERC2 with results have a profile that suggests that they could be considered candidates for threats to the planetary boundary specifically from a regime shift in populations (MacLeod et al. 2014). If chemical quantity (Q) is also considered and is set to high volume chemicals at equal to or greater than 0.1 kt/yr (i.e., ~PMTQ approach where Q represents chemical quantity), then <1% of ERC2 substances are implicated in this example.

The scale of exposure in this example reaches up to the global level predominantly from water transport, suggesting that the temporal and spatial distribution of exposure to substances having the potential to cause irreversible or poorly reversible adverse effects from genetic damage is wide-scale. The confidence scores for each of these descriptors is, however, generally low suggesting that most of these outcomes are based on *in silico* profiling. The total severity score from just these three descriptors would result in a high severity category outcome. If the previous confidence-severity matrix is used, further evaluation would be recommended for all results noting that a key data collection need is current reports of chemical quantity. The results can also be used to better target environmental media and biomonitoring.

Table 14: Example table of 20 ERC2 substances profiled to be potential threats to the planetary boundary

ERC SUBSTANCE	HAZARD			EXPOSURE							
	Chemical Reactivity/Genotoxicity			Response Time (YR)			Mobility (KM)				
No.	Classification	Confidence	Target	Classification	Confidence	Response Time	Classification	Confidence	CTD	Transport Medium	Exposure Scale
1	3	13	DNA+Protein	3	5	26	3	5	3163	Water	Global
2	3	2	DNA	3	5	18	3	5	2308	Water	Global
3	3	4	DNA	3	5	25	3	5	3163	Water	Global
4	3	4	DNA+Protein	3	5	24	3	5	2118	Water	Global
5	3	3	DNA	3	5	16	3	5	2063	Water	Global
6	3	3	DNA+Protein	3	5	18	3	5	2246	Water	Global
7	3	3	DNA	3	5	17	3	5	2124	Water	Global
8	3	3	DNA+Protein	3	5	20	3	5	2542	Water	Global
9	3	3	DNA	3	5	24	3	5	2992	Water	Global
10	3	3	DNA+Protein	3	5	19	3	5	2385	Water	Global
11	3	3	DNA+Protein	3	5	17	3	5	2200	Water	Global
12	3	3	DNA+Protein	3	5	26	3	5	3305	Water	Global
13	3	3	DNA+Protein	3	5	16	3	5	2004	Water	Global
14	3	3	DNA+Protein	3	5	19	3	5	2361	Water	Global
15	3	2	Protein	3	5	17	3	5	2200	Water	Global
16	3	3	DNA+Protein	3	5	30	3	5	3750	Water	Global
17	3	3	DNA+Protein	3	5	25	3	5	3136	Water	Global
18	3	3	DNA+Protein	3	5	26	3	5	3287	Water	Global
19	3	3	DNA+Protein	3	5	25	3	5	3087	Water	Global
20	3	3.5	DNA+Protein	3	5	32	3	5	3657	Water	Global

Abbreviations: YR, year; CTD, characteristic travel distance

7. Performance and Conclusions

While the first version of ERC (ERC1) was a considerable step forward for targeting chemicals of concern, it can nonetheless be considered a first effort or proof of concept by ECCC to incorporate 21st-century science and methods to advance risk-based ecological prioritization approaches and science. ERC2 represents a considerable evolution from ERC1; it has not only maintained the evidence-driven principles of the earlier version, but also improves on the toxicological and exposure space while giving more weight to cross-species susceptibility to identify chemicals of concern over a wide range of temporal and spatial scales. The incorporation of consensus approaches using the AOP and other concepts increases the transparency of evidence weighting, so that the notion of "*best available science for evidence-based decision-making*" is maintained and used to provide direction for possible future regulatory activities.

ERC2 was designed with a focus on 13 162 chemicals on the Canadian DSL not categorized as a priority in 2006; however, the question remains regarding its performance outside of this chemical space. A form of chemical benchmarking may be helpful in answering this. Figure 26 provides risk-based results for known chemical classes of emerging concern: perfluorinated acids, BPA-like structures, and organic phosphate ester chemicals possibly used as replacement flame retardants (total n=77). Figure 27 provides results for the same analysis using only hazard classification.

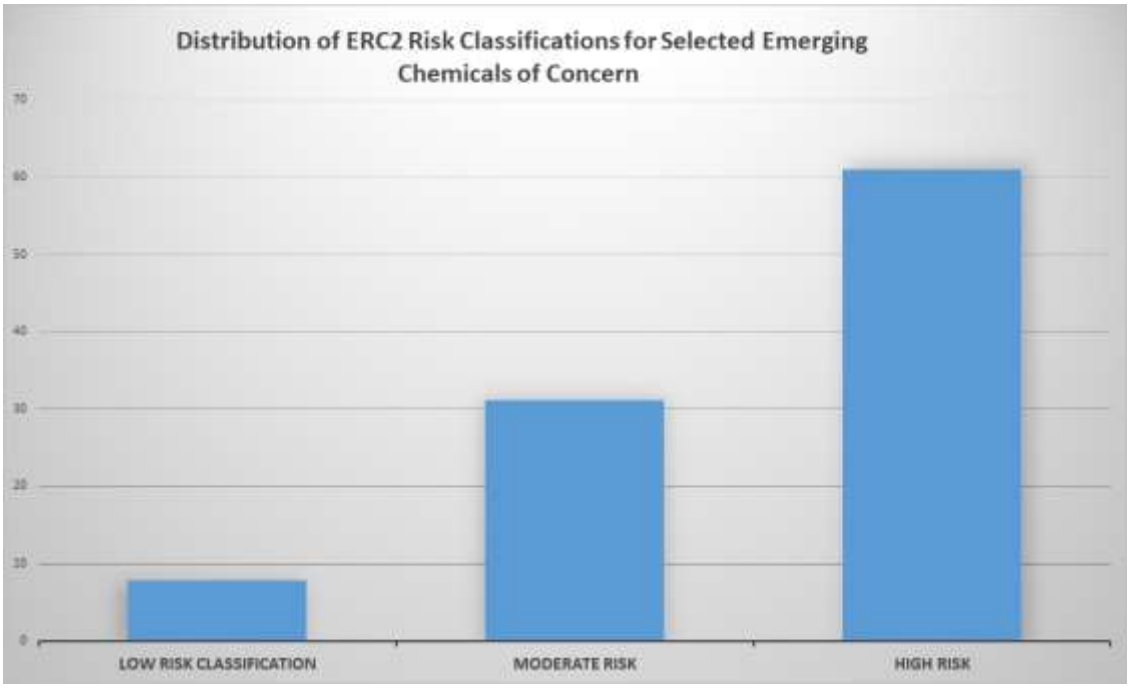


Figure 26: Distribution (%) of ERC risk classification outcomes for 77 selected emerging chemicals of concern

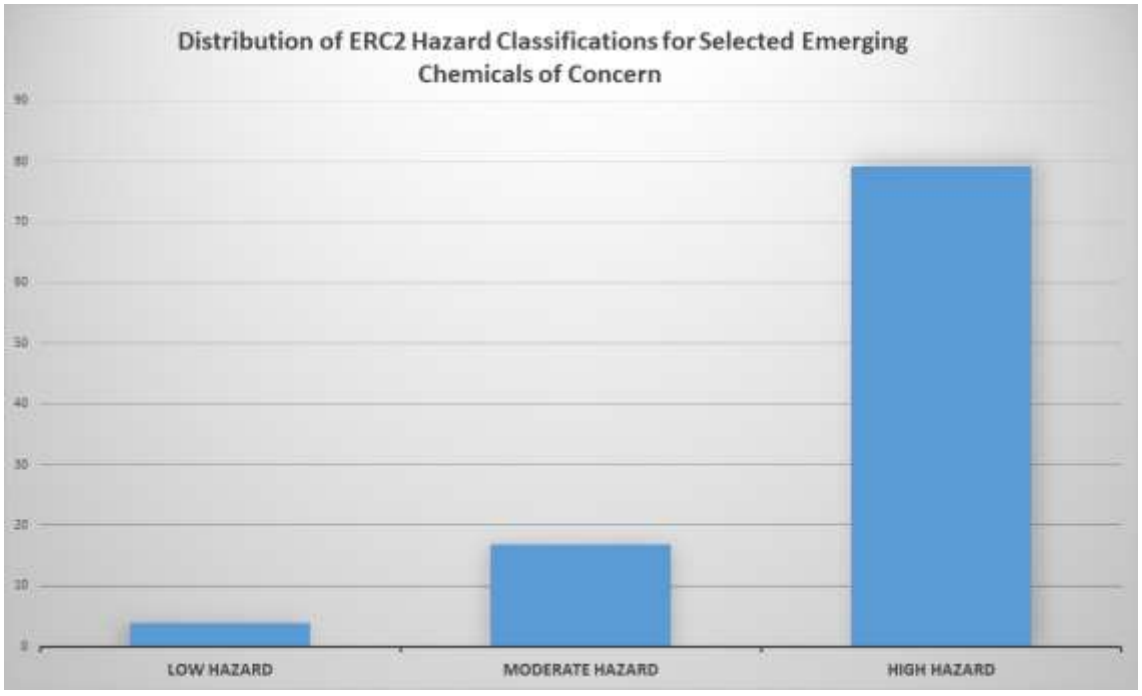


Figure 27: Distribution (%) of ERC hazard classification outcomes for 77 selected emerging chemicals of concern

It can be concluded from the above two figures that ERC2 is sensitive to capturing selected emerging chemicals (>90%) of concern if moderate and high risk classification outcomes are grouped and risk

outcomes are considered (Figure 26). A high degree of concordance with hazard classification can be seen in Figure 27 at over 90%, again combining moderate and high hazard outcomes. The risk and hazard confidence scores vary across the 77 substances (from low to very high), indicating some of these chemicals should receive greater data collection.

Figure 28 compares hazard classification outcomes for 29 REACH "Substances of Very High Concern" (SVHC) candidates for authorization from January 2021[22] contained in ERC2.

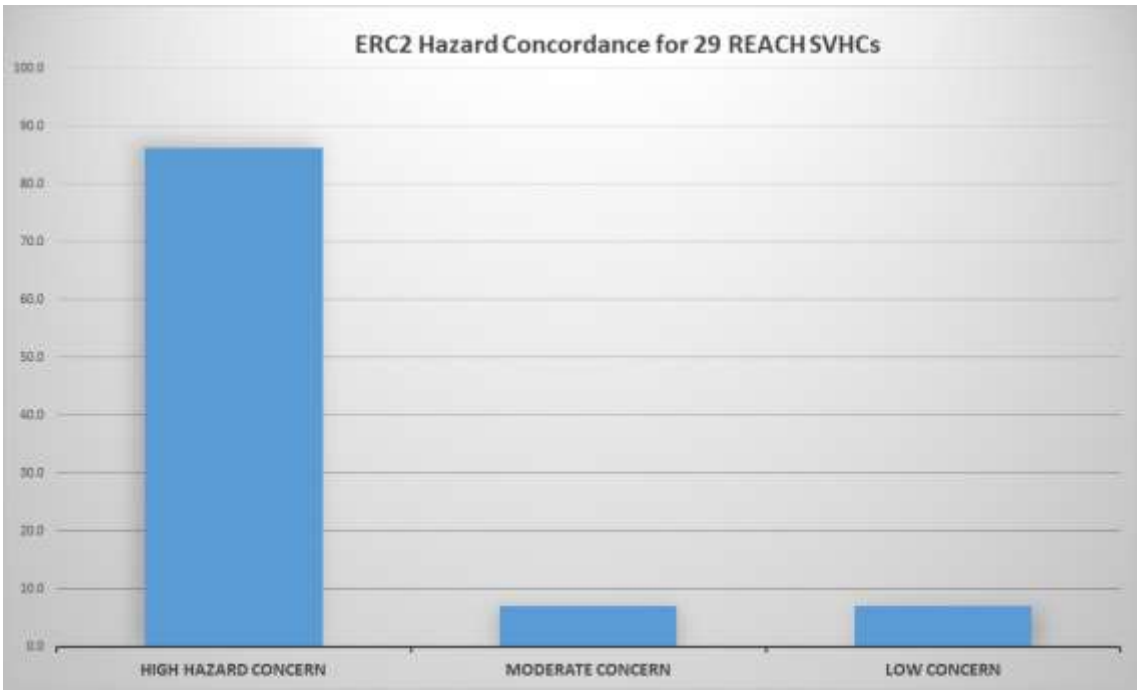


Figure 28: Percentage ERC2 hazard classification concordance for 29 substances listed as candidate SVHCs in the EU

Results in Figure 28 show that regardless of the human or ecological basis for SVHC nomination, ERC2 hazard classification gives almost 90% concordance with REACH SVHC candidates when moderate and high hazard classifications are merged, supporting the "one toxicology" concept advocated in ERC2 as well as regulatory alignment.

Benchmarking substances of low concern is difficult because few scientific journals or regulatory programs publically report lists of low risk outcomes that are easily accessed for benchmarking purposes. Besides the Government of Canada's lists of low priority substances from previous Rapid Screening or ERC1 activities, which are not contained in ERC2, the US EPA's list of low priority chemicals not subject to further evaluation under the Toxic Substances Control Act (TSCA)[23] can be used. TSCA low priority designations are risk-based, therefore ERC2 risk classification results for 18 of the 20 TSCA low priority chemicals are used for the comparison (not all 20 substances are contained in ERC2). Table 15

[22] Candidate List of substances of very high concern for Authorisation - ECHA (europa.eu)
[23] Low-Priority Substances under TSCA | Assessing and Managing Chemicals under TSCA | US EPA

shows the results of this comparison. ERC2 would consider five of the 18 substances from the TSCA low priority list to be of low risk classification. All 18 substances have moderate or low exposure potential in ERC2, which is likely comparable to the TSCA risk assessment.

The lack of concordance between ERC2 and the 18 TSCA low priority substances is the mainly result of ERC2's relatively conservative logic rule for hazard and exposure classification. That is, low exposure classification associated with a high hazard classifications results in a moderate risk classification according to ERC2's risk matrix. This explains ~31% of the outcomes that are different (Table 15). However, five substances have evaluation flags for highly ionized chemicals flags suggesting that ERC2's moderate and high risk outcomes are not without uncertainty. Just over 50% of the substances also had margin of exposure flags resulting in an upgrade from low to moderate exposure classification according to the rules described under the margin of exposure descriptor and final exposure classification.

Table 15: Comparison of TSCA low priority designations with ERC2 risk classifications

CAS	Chemical Name	ERC2 Risk Classification	ERC2 Hazard Classification	ERC2 Exposure Classification	Evaluation Flags	Margin of Exposure Fl
90802	);gluconic acid, .delt	Low	Low	Low		
105533	iethyl ester;diethyl	High	High	Moderate		MoE
108598	ester;memalonat;pr	Moderate	Moderate	Moderate		MoE
109433	canedioate;sebacic	Moderate	Moderate	Moderate		MoE
110985	propanol, 1,1'-oxybi	Low	Low	Low		
111013	23-hexamethyl-;2,6,	High	Moderate	High		MoE
112925	ctadecanol;octadeca	High	High	Moderate		MoE
299274	s,4r,5r)-2,3,4,5,6-per	Moderate	High	Low	>=99% fraction ionized	
299285	bis[(2r,3s,4r,5r)-2,3	Moderate	High	Low	>=99% fraction ionized	
526954	cid, d-;gluconic acid	High	High	Moderate	>=99% fraction ionized	MoE
527071	-2,3,4,5,6-pentahydr	High	High	Moderate	>=99% fraction ionized	MoE
629969	osan-1-ol;eicosanol	Moderate	Moderate	Moderate		MoE
661198	docosanol;docosan-1	Low	Moderate	Low		
4435534	3-methoxybutyl acet	Low	Low	Low		
24800440	nethyl-1,2-ethanedi	Low	Low	Moderate		MoE
25265718	-2(methyl-2) oxybis	Moderate	Moderate	Moderate		MoE
31138655	non-preferred name	Moderate	High	Low	>=99% fraction ionized	
88917220	-(2-methoxymethyl	Moderate	High	Low		

8. Remaining Uncertainties

The approaches used in ERC2 consider weighted lines of evidence, where uncertainty is factored into priority setting using confidence scoring and consensus approaches. Nonetheless, uncertainty is inherent in ecological assessment given the lack of species specificity, and the quality and amount of ecological data. Uncertainty therefore remains in some key areas of ERC2. First off, the formation of plausible mechanistic-causality relationships is based on a defined hazard space and key relationships using the hazard descriptors. When adverse effects cannot be explained via plausible mechanistic-causality relationships, observed *in vivo* effects are flagged as "unconfirmed" and are potentially the result of an unknown mechanism not included among those that define the toxicological space for hazard profiling. This applies primarily to receptor-mediated and chemical reactivity/genotoxicity descriptors, but also the MoA descriptor. *In vivo*-level data (predicted and observed) are used to determine specific vs. non-specific modes of action, noting that the selected specific MoA may not necessary explain the *in vivo* IEC_{50} potency in the consensus approach for MoA.

Another source of uncertainty remains with UVCBs profiled in ERC2. Currently, almost 1000 UVCBs with a chemical structure have results that are not available (NA) in ERC2. These are true unknowns where currently no knowledge of their fate, behaviour and toxicity are easily obtainable for prioritization purposes. As noted in section 3.2, UVCB results were profiled in ERC2 using a representative component approach. This can be regarded as a practical solution for dealing with thousands of UVCBs at a prioritization stage, but it should be re-emphasized that a single component, even if selected as a reasonable worst-case, may not reflect the behaviour of the UVCB as a complex mixture.

ERC2 contains a small percentage of ionizing organic chemicals (IOCs) that have a significant fraction of the mass existing in the ionized form at a relevant internal and environmental pH (pH 7.4). There are approximately 25% of IOCs with <1% neutral fraction. There is also a small percentage of ERC2 substances that are expected to be distributed to blood plasma in organisms (~1%), many being permanently charged chemicals such as quaternary ammonium chemicals and strong acids. ERC2 results for these substances have greater uncertainty than those with a higher fraction in the neutral form since toxicokinetics are often driven by the neutral form of a substance (e.g., alkyl amines). While some *in silico* approaches in ERC2 are able to deal with IOCs (e.g., RAIDAR v3.0), many QSAR-type models require the input for the neutral form. Results for IOCs with low or no neutral fraction are consequently regarded as generally conservative for anionic forms, but may approximate the behaviour of the cationic form based on comparison of physical-chemical properties using neutral vs. cationic species. The percentage of IOCs, here defined as permanently charged IOCs and those having greater than or equal to 99% ionizing form (pH 7.4), are flagged in Appendix XII.

In section 3.3 and Appendix II, rules are described for ADME outcomes suggesting low permeability/bioavailability and plasma distributed chemicals. Similar to highly ionized chemicals, uncertainty exists with ERC2 results for substances receiving these flags listed in the main table of results (Appendix XII) and as noted in the TSCA-ERC2 comparison previously. While ERC2 results are not affected by these flags directly, substances identified for further evaluation and data collection will

consider the impact of these flags under the suggested future regulatory activity. This may result in reclassification of ERC2 outcomes for these substances at a future date.

Efforts have been made to produce a system for prioritization that utilizes many sources of data to inform evidence-based decision-making (Appendix I). Nonetheless, an extensive substance-by-substance search of hazard and exposure literature could not practically be performed for the 13 162 chemicals. Information contained in databases having batch search capabilities and *in silico* tools able to batch run thousands of chemicals provided the main source of information for ERC2. It is therefore acknowledged that some information for individual substances will not be present in ERC2. These data can be added during future updates to ERC2 and may impact current ERC2 outcomes.

Finally, chemical quantity data has been extrapolated from reported tonnages in 1986 for 97% of substances in ERC2. The extrapolated chemical quantities are, as a whole, considered to be moderate to highly uncertain with a significant percentage (40-50%) of substances in ERC2 likely no longer in commerce in Canada. Based on ECCC's chemical quantity analysis described in section 3.5.3.3, the currently extrapolated tonnages are expected to be a factor of five or lower than any updated quantity data 80% of the time and within a factor of 10, 95% of the time. Use of the extrapolated quantities directly affects the emission pattern and margin of exposure descriptors. These descriptors have had their confidence score weighted lower as a result. This, however, has little impact on the total exposure confidence results since most of the exposure confidence scores are low to begin with (reflecting the reliance on *in silico* data). Therefore, for the examples used in section 6.0, where evaluation is suggested as the future regulatory activity, a first step for the evaluation activity will require data collection of chemical quantity and possibly use pattern to refine these outcomes.

9. References

Ames Test ISSSTY [database of *in vitro* mutagenicity results (Ames test) results]. 2011. Ver. V4b. Rome (IT): Istituto Superiore di Sanità (ISS); Liebefeld (CH): Federal Office of Public Health. Benigni R, Battistelli CL, Bossa C, Tcheremenskaia O, Crettaz P. 2013. New perspectives in toxicological information management, and the role of ISSTOX databases in assessing chemical mutagenicity and carcinogenicity. Mutagenesis. 28(4):401-409].

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, et al. 2010. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. Environ Toxicol Chem. 29(3):730-41.

Ankley GT, Gray LE. 2013. Cross-species conservation of endocrine pathways: a critical analysis of tier 1 fish and rat screening assays with 12 model chemicals. Environ Toxicol Chem. 32(5):1084-1087.

[ARC] ARC Arnot Research & Consulting. 2014. Parameterization and application of the RAIDAR model to support prioritization and assessment of substances. Gatineau (QC): Environment Canada.

[ARC] ARC Arnot Research & Consulting. 2018. Generation of physical-chemical property data and the application of models for estimating fate and transport and exposure and risk potential for organic substances on the Canadian DSL. Gatineau (QC): Environment and Climate Change Canada.

Armitage J, Arnot JA, Bonnell M. 2018. Comparing mode of action (MOA) classification using body residues, membrane concentrations and chemical activity for chemical prioritization. Abstracts, 39th Annual Meeting, Society of Environmental Toxicology and Chemistry Sacramento. November 4–8. Sacramento, CA United States of America.

Arnot JA. 2011. Updating the RAIDAR and FHX models to aid in the prioritization and assessments of chemicals including ionisable substances. Draft Technical Report. Ottawa (ON): Health Canada. p. 39.

Arnot JA, Brown TN, Wania F. 2014. Estimating screening-level organic chemical half-lives in humans. Environ Sci Technol. 48:723-730.

Arnot JA, Gobas FAPC. 2004. A food web bioaccumulation model for organic chemicals in aquatic ecosystems. Environ Toxicol Chem. 23(10):2343-2355.

Arnot JA, Gouin T, Mackay D (Canadian Environmental Modelling Network (CEMN, Trent University, Peterborough, ON). 2005. Development and fate of models of chemical fate in Canada. Practical methods for estimating environmental biodegradation rates. Gatineau (QC): Environment Canada. Report No. 200503. 50 p.

Arnot JA, Mackay D. 2008. Policies for chemical hazard and risk priority setting: can persistence, bioaccumulation, toxicity and quantity information be combined? Environ Sci Technol. 42(13):4648-4654.

Arnot JA, Mackay D, Webster E, Southwood J. 2006. Screening level risk assessment model for chemical fate and effects in the environment. Environ Sci Technol. 40:2316-2323.

Aronson D, Boethling R, Howard P, Stiteler W. 2006. Estimating biodegradation half-lives for use in chemical screening. Chemosphere. 63(11):1953-1960.

Arthur K, Limpus C, Balazs GH, Capper A, Udy, J, Shaw G, Keuper-Bennett U, Bennett P. 2008. The exposure of green turtles (Chelonia mydas) to tumour promoting chemicals produced by the cyanobacterium Lyngbya majuscula and their potential role in the aetiology of fibropapillomatosis. Harmful Algae. 7(1):114–125.

Backhaus T, Faust M. 2012. Predictive environmental risk assessment of chemical mixtures: a conceptual framework. Environ Sci Technol. 46:2564-2573.

Baron MG, Lilavois CR, Martin TM. 2015. MOATox: A comprehensive mode of action and acute toxicity database for predictive model development. Aquat Toxicol. 161:102-107.

Bauer F, Thomas PC, Fouchard SY, Neunlist SJM. 2018. A new classification algorithm based on mechanisms of action. Comput Toxicol. 5:8-15.

Becker RA, Friedman KP, Simon TW, Marty MS, Patlewicz G, Rowlands JC. 2015. An exposure activity profiling method for interpreting high-throughput screening data for estrogenic activity—Proof of concept. Regul Toxicol Pharmacol. 71(3):398-408.

Belanger S, Sanderson H, Embry M, DeZwart D, Farr B, Gutsell S, Haider M, Sternberg R, Wilson P. 2015. It is time to develop ecological thresholds of toxicological concern to assist environmental hazard assessment. Environ Toxicol Chem. 34(12):2864-2869.

Bennett DH, McKone TE, Matthies M, Kastenberg WE. 1998. General formulation of characteristic travel distance for semivolatile organic chemicals in a multimedia environment. Environ Sci Tech. 32(24):4023-4030.

Beyer A, Mackay D, Matthies M, Wania F, Webster E. 2000. Assessing long-range transport potential of persistent organic pollutants. Environ Sci Tech. 34:699-703.

Brown TN, Arnot JA, Wania F. 2012. Iterative fragment selection: A group contribution approach to predicting fish biotransformation half-lives. Environ Sci Technol. 46:8253-8260.

Canada. 1999. *Canadian Environmental Protection Act*, 1999. S.C. 1999, c.33. Canada Gazette Part III, vol. 22, no. 3. http://laws-lois.justice.gc.ca/eng/acts/C-15.31/.

Canada. 2000. Canadian Environmental Protection Act, 1999: Persistence and Bioaccumulation Regulations. P.C. 2000-348, 29 March, 2000, SOR/2000-107. http://laws-lois.justice.gc.ca/eng/regulations/SOR-2000-107/page-1.html.

Connors KA, Beasley A, Barron MG, Belanger SE, Bonnell M, Brill JL, de Zwart D, Kienzler A, Krailler J, Otter R, et al. 2019. Creation of a curated aquatic toxicology database: EnviroTox. Environ Toxicol Chem. 38:1062-1073.

Cotterill JV, Palazzolo L, Ridgway C, Price N, Rorije E, Moretto A, Peijnenburg A, Eberini I. 2019. Predicting estrogen receptor binding of chemicals using a suite of *in silico* methods – Complementary approaches of (Q)SAR, molecular docking and molecular dynamics. Toxicol App Pharmacol. 378:1-9.

Cousins IT, DeWitt JC, Glüge J, Goldenman G, Herzke D, Lohmann R, Ng CA, Scheringer M, Wang Z. 2020. The high persistence of PFAS is sufficient for their management as a chemical class. Environ Sci Process Impacts. 22(12):2307-2312. doi: 10.1039/d0em00355g.

de Wolf W, Siebel-Sauer A, Lecloux A, Koch V, Holt M, Feijtel T, Comber M, Boeije G. 2005. Mode of action and aquatic exposure thresholds of no concern. Environ Toxicol Chem. 24(2):479-485.
Diamond ML, de Wit CA, Molander S, Scheringer M, Backhaus T, Lohmann R, Arvidsson R, Bergman Å, Hauschild M, Holoubek I, et al. 2015. Exploring the planetary boundary for chemical pollution. Environ Internat. 78:8-15.

Dimitrov, S, Detroyer A, Piroird C, Gomes C, Eilstein J, Pauloin T, Kuseva C, Ivanova H, Popova I, Karakolev Y, et al. 2016. Accounting for data variability, a key factor in *in vivo*/*in vitro* relationships: application to the skin sensitization potency (*in vivo* LLNA versus *in vitro* DPRA) example. J Appl Toxicol. 36(12):1568-1578.

Dyer S, Warne MSJ, Meyer JS, Leslie HA, Escher BI. 2010. Tissue residue approach for chemical mixtures. Integ Environ Assess Manage. 7(1):99-115.

[EC, HC] Environment Canada, Health Canada. 2015. Proposed Approach for Cumulative Risk Assessment of Certain Phthalates under the Chemicals Management Plan. Ottawa (ON): Government of Canada.

[ECCC] Environment and Climate Change Canada. 2016. Science approach document: ecological risk classification of organic substances. Ottawa (ON): Government of Canada.
Risk management objective for Dinoseb.

[ECCC, HC] Environment and Climate Change Canada, Health Canada. [modified 2017 Mar 12]. Categorization. Ottawa (ON): Government of Canada.

[ECCC, HC] Environment and Climate Change Canada, Health Canada. 2019. New approaches for integrating chemical fate and spatial and temporal scale in exposure assessment. Objectives paper - June 12–13, 2019 - Chemicals Management Plan Science Committee. Ottawa (ON): Government of Canada.

[ECCC, HC] Environment and Climate Change Canada, Health Canada. 2019. Dinoseb - Canada.ca. Ottawa (ON): Government of Canada.

[ECHA] European Chemicals Agency. 2020. Candidate list of substances of very high concern for authorization. Helsinki (FI): ECHA.

Ellison EA, Api A-M, Becker RA., Efremenko AY, Gadhia S, Hack CE, Hewitt NJ, Varcin M, Schepky. 2021. Internal Threshold of Toxicological Concern (iTTC): Where We Are Today and What Is Possible in the Near Future. Frontiers in Toxicology. 2: 621541.

Endo S, Escher BI, Goss KU. 2011. Capacities of membrane lipids to accumulate neutral organic chemicals. Environ Sci Technol. 45:5912-5921.

Escher BI, Ashauer R, Dyer S, Hermens JLM, Lee J-H, Leslie HA, Mayer P, Meador JP, Warne MSJ. 2011. Crucial role of mechanisms and modes of toxic action for understanding tissue residue toxicity and internal effect concentrations of organic chemicals. Integr Environ Assess Manag. 7(1):28-49.

Franco A, Trapp S. 2010. A multimedia activity model for ionizable chemicals: Validation study with 2,4-dichlorophenoxyacetic acid, aniline, and trimethoprim. Environ Toxicol Chem. 29(4):789-799.

Gallagher SS, Rice GE, Scarano LJ, Teuschler LK, Bollweg G, Martin L. 2015. Cumulative risk assessment lessons learned: a review of Case studies and issue papers. Chemosphere. 120:697-705.

Gerberick GF, Vassallo JD, Bailey RE, Chaney JG, Morrall SW, Lepoittevin JP. 2004. Development of a peptide reactivity assay for screening contact allergens. Toxicol Sci. 81:332-343.

Gouin T, Wania F. 2007. Time trends of Arctic contamination in relation to emission history and chemical persistence and partitioning properties. Environ Sci Technol. 41(17):5986-92.

Jaworska J, Natsch A, Ryan C, Strickland J, Ashikaga T, Miyazawa M. 2015. Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: a decision support system for quantitative weight of evidence and adaptive testing strategy. Arch Toxicol. 89(12):2355-2383.

Karickhoff SW. 1981. Semiempirical estimation of sorption of hydrophobic pollutants on natural sediments and soils. Chemosphere. 10:833-849.

Kienzler A, Connors KA, Bonnell M, Barron MG, Beasley A, Inglis CG, Norberg-King TJ, Martin T, Sanderson H, Vallotton N, et al. 2019. Mode of action classifications in the EnviroTox database: Development and implementation of a consensus MOA classification. Environ Toxicol Chem. 38:2294-2304.

Kirkland D. 1998. Chromosome aberration testing in genetic toxicology—past, present and future. Mutat Res-Fund Mol M. 404(2):173-185.

Książek P, Bryl K. 2015. Molecular docking reveals binding features of estrogen receptor beta selective ligands. Curr Comput Aided Drug Des. 11(2):137-51.

LaLone CA, Villeneuve DL, Lyons D, Helgen HW, Robinson SL, Swintek JA, Saari TW, Ankley GT. 2016. Editor's highlight: sequence alignment to predict across species susceptibility (SeqAPASS): A web-based tool for addressing the challenges of cross-species extrapolation of chemical toxicity. Toxicol Sci. 153(2):228-45.

Leeman WR, Krul L, Houben GF. 2014. Reevaluation of the Munro dataset to derive more specific TTC thresholds. Regul Toxicol Pharmacol. 69(2):273-8.

Mackay D. 2001. Multimedia Environmental Models - The Fugacity Approach. 2nd Ed. Boca Raton (FL): CRC Press. 272 p.

Mackay D, Arnot JA, Petkova EP, Wallace KB, Call DJ, Brooke LT, Veith GD. 2009. The physicochemical basis of QSARs for baseline toxicity. SAR QSAR Environ Res. 20(4):393-414.

Mackay D, Di Guardo A, Paterson S, Cowan C. 1996. Evaluating the environmental fate of a variety of types of chemicals using the EQC model. Environ Toxicol Chem. 15:1627-1637.

Mackay D, Hughes DM, Romano M, Bonnell M. 2014. The role of persistence in chemical evaluations. Integr Environ Assess Manag. 10(4):588–594.

Mackay D, Reid LK. 2008. Local and distant residence times of contaminants in multi-compartment models. Part I: A review of the theoretical basis. Environ Pollut. 156:1196-1203.

MacLeod M, Breitholtz M, Cousins IT, de Wit CA, Persson LM, Rudén C, McLachlan MS. 2014. Identifying chemicals that are planetary boundary threats. Environ Sci Technol. 48:11057–11063.

Maeder V, Escher BI, Scheringer M, Hungerbühler K. 2004. Toxic ratio as an indicator of the intrinsic toxicity in the assessment of persistent, bioaccumulative, and toxic chemicals. Environ Sci Technol. 38(13):3659-3666.

Mansouri K, Grulke CM, Judson RS, Williams AJ. 2018. OPERA models for predicting physicochemical properties and environmental fate endpoints. J Cheminform. 10(1):10.

Martin TM, Young DM, Lilavois CR, Barron MG. 2015. Comparison of global and mode of action–based models for aquatic toxicity. SAR QSAR Environ Res. 26:245–262.

Matthies M, Solomon K, Vighi M, Gilman A, Tarazona JV (2016). The origin and evolution of assessment criteria for persistent, bioaccumulative and toxic (PBT) chemicals and persistent organic pollutants (POPs). Environ Sci Process Impacts. 18:1114–1128.

Matthiessen P, Ankley GT, Biever RC, Bjerregaard P, Borgert C, Brugger K, Blankinship A, Chambers J, Coady KK, Constantine L, et al. 2017. Recommended approaches to the scientific evaluation of ecotoxicological hazards and risks of endocrine-active substances. Integr Environ Assess Manag. 13:267-279.

Mayer P, Reichenberg F. 2006. Can highly hydrophobic organic substances cause aquatic baseline toxicity and can they contribute to mixture toxicity? Environ Toxicol Chem. 25:2639-2644.

McCarty LS, Arnot JA, Mackay D. 2013. Evaluation of critical body residue for acute narcosis in aquatic organisms. Environ Toxicol Chem. 32(10):2301-2314.

McCarty LS, Mackay D. 1993. Enhancing ecotoxicological modeling and assessment: critical body residues and modes of toxic action. Environ Sci Technol. 27(9):1719-1728.

McLachlan MS, Kierkegaard A, Radke M, Sobek A, Malmvärn A, Alsberg T, Arnot JA, Brown TN, Wania F, Breivik K, et al. 2014. Using model-based screening to help discover unknown environmental contaminants. Environ Sci Technol. 48:7264–7271.

Mekenyan O, Dimitrov S, Serafimova R, Thompson E, Kotov S, Dimitrova N, Walker J. 2004. Identification of the structural requirements for mutagenicity by incorporating molecular flexibility and metabolic activation of chemicals I: TA100. Chem Res Toxicol. 17:753-766.

Mekenyan O, Serafimova R. 2009. Endocrine Disruption Modeling. Boca Raton (FL): CRC Press. Chapter, Mechanism-based modeling of estrogen receptor binding affinity a COREPA Implementation. p. 259-293.

Mekenyan O, Todorov M, Serafimova R, Stoeva S, Aptula A, Finking R, Jacob E. 2007. Identifying the structural requirements for chromosomal aberration by incorporating molecular flexibility and metabolic activation of chemicals. Chem Res Toxicol. 20:1927−1941.

Micronucleus ISSMIC [database of *in vivo* micronucleus mutagenicity assay results]. 2011. Ver. v4b. Rome (IT): Istituto Superiore di Sanità (ISS); Liebefeld (CH): Federal Office of Public Health. [Benigni R, Battistelli CL, Bossa C, Tcheremenskaia O, Crettaz P. 2013. New perspectives in toxicological information management, and the role of ISSTOX databases in assessing chemical mutagenicity and carcinogenicity. Mutagenesis. 28(4):401-409].

Moore DRJ, Breton RL, MacDonald DB. 2003. A comparison of model performance for six quantitative structure-activity relationship packages that predict acute toxicity to fish. Environ Toxicol Chem. 22:1799-1809.

Muir DCG, Howard PH. 2006. Are there other persistent organic pollutants? A challenge for environmental chemists. Environ Sci Technol. 40:7157–7166.

Natsch A, Emter R, Gfeller H, Haupt T, Ellis G. 2015. Predicting skin sensitizer potency based on *in vitro* data from KeratinoSens and kinetic peptide binding: global versus domain-based assessment. Toxicol Sci. 143(2):319-332.

Natsch A, Gfeller H. 2008. LC-MS-based characterization of the peptide reactivity of chemicals to improve the *in vitro* prediction of the skin sensitisation potential. Toxicol Sci. 106:464-478.

Nendza M, Müller M, Wenzel A. 2014. Discriminating toxicant classes by mode of action: 4. Baseline and excess toxicity. SAR QSAR Environ Res. 25(5):393-405.

[OECD] Organization for Economic Co-operation and Development. 2018. Prioritisation of chemicals using the integrated approaches for testing and assessment (IATA)-based ecological risk classification. Paris (FR): OECD, Environment Directorate. (Series on Testing and Assessment No. 291; Report No.: ENV/JM/MONO(2018)27, JT03435853).

[OECD] Organization for Economic Co-operation and Development. 2019. Guiding principles and key elements for establish a weight of evidence for chemical assessment. Paris (FR): OECD, Environment Directorate. (Series on Testing and Assessment No. 311; Report No.: ENV/JM/MONO(2019)31, JT03453231).

Pennington DW. 2001. An evaluation of chemical persistence screening approaches. Chemosphere. 44:1589-1601.

Princz J, Bonnell M, Ritchie E, Velicogna J, Robidoux P-Y, Scroggins R. 2014. Estimation of the bioaccumulation potential of a non-chlorinated bisphenol and an ionogenic xanthene dye to Eisenia andrei in field-collected soils, in conjunction with predictive *in silico* profiling. Environ Toxicol Chem. 33(2):308-316.

Reid LK, Mackay D. 2008. Local and distant residence times of contaminants in multi-compartment models. Part II: Application to assessing environmental mobility and long-range atmospheric transport. Environ Pollut. 156:1182-1189.

Reppas-Chrysovitsinos E, Sobek A, MacLeod M. 2018. *In silico* screening-level Pprioritization of 8468 chemicals produced in OECD countries to identify potential planetary boundary threats. Bull Environ Contam Toxicol. 100(1):134-146.

Richarz AN, Schultz TW, Cronin MTD, Enoch SJ. 2014. Experimental verification of structural alerts for the protein binding of sulfur-containing chemicals. SAR QSAR Environ Res. 25(4):325-341.

Roberts DW, Schultz TW, Wolf EM, Aptula AO. 2010. Experimental reactivity parameters for toxicity modeling: application to the acute aquatic toxicity of S2 electrophiles to Tetrahymena pyriformis. Chem Res Toxicol. 23(1):228–234.

Rockstrom J, Steffen W, Noone K, Persson A, Chapin FS, Lambin EF, Lenton TM, Scheffer M, Folke C, Schellnhuber HJ, et al. 2009. Planetary boundaries: Exploring the safe operating space for humanity. Ecol Soc. 14(2):32.

Rowland MA, Perkins EJ, Mayo ML. 2017. Physiological fidelity or model parsimony? The relative performance of reverse-toxicokinetic modeling approaches. BMC Syst Biol. 11:35.

Rüdel H, Körner W, Letzel T, Neumann M, Nödler K, Reemtsma T. 2020. Persistent, mobile and toxic substances in the environment: a spotlight on current research and regulatory activities. Environ Sci Eur. 32:5.

Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE, Drummond RA. 1997. Predicting modes of toxic action from chemical structure: acute toxicity in the fathead minnow (Pimephales promelas). Environ Toxicol Chem. 16(5):948−967.

Salvito D, Fernandez M, Jenner K, Lyon DY, de Knecht J, Mayer P, MacLeod M, Eisenreich K, Leonards P, Cesnaitis R, et al. 2020. Improving the Environmental Risk Assessment of Substances of Unknown or Variable Composition, Complex Reaction Products, or Biological Materials. Environ Toxicol Chem. 39: 2097-2108.

Sapounidou M, Ebbrell DJ, Bonnell M, Campos B, Firman WB, Gutsell S, Hodges G, Roberts J, Cronin MTD. 2021. Development of an enhanced mechanistically-driven mode of action classification scheme for adverse effects in environmental species. Environ Sci Technol. 55(3):1897-1907.

Sappington KG, Bridges TS, Bradbury SP, Erickson RJ, Hendriks KA, Lanno RP, Meador JP, Mount DR, Salazar MH, Spry DJ. 2010. Application of the tissue residue approach in ecological risk assessment. Integr Environ Assess Manag. 7(1):116–140.

Schmidt SN, Armitage JM, Arnot J, Kusk K, Mayer P. 2016. Linking algal growth inhibition to chemical activity: a tool for identifying excess toxicity [presentation]. Society of Environmental Toxicology and Chemistry (SETAC) Europe; 2016 May 22-26; Nantes, France.

Schmidt SN, Mayer P. 2015. Linking algal growth inhibition to chemical activity: baseline toxicity required 1% of saturation. Chemosphere. 120:305-308.

Schultz TW, Sparfkin CL, Aptula AO. 2010. Reactivity-based toxicity modelling of five-membered heterocyclic chemicals: application to Tetrahymena pyriformis. SAR QSAR Environ Res. 21(7-8):681–691.

Serafimova R, Todorov M, Nedelcheva D, Pavlov T, Akahori Y, Nakai M, Mekenyan O. 2007. QSAR and mechanistic interpretation of estrogen receptor binding. SAR QSAR Environ Res. 18(3-4):1-33.

Serafimova R, Todorov, M, Pavlov T, Kotov S, Jacob E, Aptula A, Mekenyan O. 2007. Identification of the structural requirements for mutagencitiy, by incorporating molecular flexibility and metabolic activation of chemicals. II. General Ames mutagenicity model. Chem Res Toxicol. 20:662-676.

Seth R, Mackay D, Muncke J. 1999. estimating the organic carbon partition coefficient and its variability for hydrophobic chemicals. Environ Sci Technol. 33(14):2390-2394.

Stadnicka-Michalak J, Schirmer K. 2019. *In vitro-in vivo* extrapolation to predict bioaccumulation and toxicity of chemicals in fish using physiologically based toxicokinetic models. Methods Pharmacol Toxicol. 44:1-30.

Stahl CH, Cimorelli AJ. 2013. A demonstration of the necessity and feasibility of using a clumsy decision analytic approach on wicked environmental problems. Integr Environ Assess Manag. 9(1):17-30.

Stroebe M, Scheringer M, Hungerbuhler K. 2004. Measures of overall persistence and the temporal remote state. Environ Sci Technol. 38(21):5665-5673.

Struijs J, Stoltenkamp J, Van de Meent D. 1991. A spreadsheet-based box model to predict the fate of xenobiotics from a municipal wastewater treatment plant. Wat Res. 25(7):891-900.

Suter II GW and Cormier SM. 2011. Why and how to combine evidence in environmental assessments: Weighing evidence and building cases. Sci Total Environ. 409: 1406-1417.
Tagliabue SG, Faber SC, Motta S, Denison MS, Bonati L. 2019. Modeling the binding of diverse ligands within the Ah receptor ligand binding domain. Sci Rep. 9:10693.

[TIMES] TIssue MEtabolism Simulator [prediction module]. 2016. Ver. 2.27.19. Bourgas (BG): University "Prof. Dr. Assen Zlatarov", Laboratory of Mathematical Chemistry.

Todorov M, Mombelli E, Ait-Aissa S, Mekenyan O. 2011. Androgen receptor binding affinity: A QSAR evaluation. SAR QSAR Environ Res. 22(3-4):265-291.

Trapp S, Franco A, Mackay D. 2010. Activity-based concept for transport and partitioning of ionizing organics. Environ Sci Technol. 44(16):6123-6129.

[EPI Suite] Estimation Program Interface Suite for Microsoft Windows [estimation model]. c2000-2012. Ver. 4.11. Washington (DC): US Environmental Protection Agency, Office of Pollution Prevention and Toxics; Syracuse (NY): Syracuse Research Corporation.

Urbisch D, Mehling A, Guth K, Ramirez T, Honarvar N, Kolle S, Landsiedel R, Jaworska J, Kern P, Gerberick F, et al. 2015. Assessing skin sensitization hazard in mice and men using non-animal test methods. Regulat Toxicol Pharmacol. 71:337-351.

Van Leeuwen CJ, Vermeire TG, editors. 2007. Risk assessment of chemicals: an introduction. 2nd ed. The Netherlands: Springer Dordrecht. 686 p.

Verhaar HJM, Solbe J, Speksnijder J, Van Leeuwen CJ and Hermens JLM. 2000. Classifying environmental pollutants: Part 3. External validation of the classification system. Chemosphere. 40:875-883.

Verhaar HJM, Van Leeuwen CJ, Hermens JLM. 1992. Classifying environmental pollutants. 1. Structure-activity relationships for prediction of aquatic toxicity. Chemosphere. 25:471-491.

Wang Z, Li Y, Ai C, Wang Y. 2010. *In silico* prediction of estrogen receptor subtype binding affinity and selectivity using statistical methods and molecular docking with 2-arylnaphthalenes and 2-arylquinolines. Int J Mol Sci. 11(9):3434-3458.

Wania F. 2003. Assessing the potential of persistent organic chemicals for long-range transport and accumulation in polar regions. Environ Sci Technol. 37:1344-1351.

Webster E, Mackay D, Wania F. 1998. Evaluating environmental persistence. Environ Toxicol Chem. 17(11):2148-2158.

Wegmann F, Cavin L, MacLeod M, Scheringer M, Hungerbuler K. 2009. The OECD software tool for screening chemicals for persistence and long range transport potential. Environ Model Software. 24(2):228–237.

Williams ES, Berninger JP, Brooks BW. 2011. Application of chemical toxicity distributions to ecotoxicology data requirements under REACH. Environ Toxicol Chem. 30(8):1943-1954.

Wu S, Blackburn K, Amburgey J, Jaworska J, Federla T. 2010. A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. Regulat Toxicol Pharmacol. 56(1):67-81.

Zhang X, Sühring R, Serodio D, Bonnell M, Sundin N, Diamond M. 2016. Novel flame retardants: Estimating the physical-chemical properties and environmental fate of 94 halogenated and organophosphate PBDE replacements. Chemosphere. 144:2401-2407.

10. APPENDIX I: Summary of *in silico* Tools and Empirical Data Used in ERC2

Table 1: List of *in silico, in vitro* and *in vivo* data sources according to ERC2 descriptor

ERC2 Descriptor/Use	*In silico* Model	Empirical Data Source/Database	Model Version/Date	Use in ERC2
Physical-Chemical Properties and Half-life Estimation	EPIWIN, TEST; ACD Labs; Arnot et al. 2005; ARC 2018	OECD QSAR Toolbox 4.3/4.4	4.11, 4.2.1, 2020.1.2	Input for multimedia fate, behaviour and exposure modelling; model domain boundaries for fish acute toxicity; ADME
ADME	Advanced Chemistry Development (ACD)/ Percepta		2020.1.2	Volume of Distribution (Vd); Protein plasma binding (PPB); Rate of Intestinal Absorbtion (Ka); Acid dissociation constant (pKa); Fraction ionized; Cytochrome P450 Inhibition; Melting Point; Molecular Weight
	Molecular Orbital PACkage (MOPAC)		2.29.1/2.30.1	Molecular cross-section (maximum diameter [D_{max}], effective diameter [D_{eff}])
Receptor-Mediated Toxicity	OECD QSAR Toolbox: OASIS Estrogen Receptor Binding		4.4	Estrogen Receptor (ER) interactions
	Tissue Metabolism Simulator (TIMES) Estrogen Receptor Binding		2.29.1/2.30.1	ER interactions
	ACD Percepta Estrogen Receptor Binding		2020.1.2	ER interactions
	OECD QSAR Toolbox: US EPA rtER Expert System		4.4	ER interactions
	Collaborative Estrogen Receptor Activity Prediction Project (CERAPP)		2.3	ER interactions

ERC2 Descriptor/Use	*In silico* Model	Empirical Data Source/Database	Model Version/Date	Use in ERC2
	TIMES Androgen Receptor Binding		2.29.1	Androgen Receptor (AR) Interactions
	Collaborative Modeling Project for Androgen Receptor Activity (CoMPARA)		2.3	AR Interactions
	Danish QSAR Database: Leadscope Thyroid Peroxidase (TPO) Inhibition		3.5	Thyroid Peroxidase Inhibition
	TIMES aromatase inhibition		2.30.1	Steroidogenesis
	TIMES Aryl Hydrocarbon Receptor Binding		2.29.1/2.30.1	Profiling Aryl Hydrocarbon Receptor Binding
		OECD QSAR Toolbox *in vitro* data for ER binding	4.4	*In vitro* ER binding
		CERAPP observed *in vitro* ER (binding, agonist, antagonist)	Not applicable (N/A)	*In vitro* ER binding, agonist antagonist
		CoMPARA observed *in vitro* AR (binding, agonist, antagonist)	N/A	*In vitro* AR binding, agonist antagonist
		CERI Japan observed *in vitro* ER/AR binding	N/A	*In vitro* ER and AR binding
		USFDA Endocrine Disruptor Knowledge Base (EDKB)	N/A	*In vitro* ER binding

ERC2 Descriptor/Use	*In silico* Model	Empirical Data Source/Database	Model Version/Date	Use in ERC2
		Thyroid Peroxidase (TPO) Inhibition observed	N/A	*In vitro* TPO Inhibition
Chemical Reactivity and Genotoxicity	OECD QSAR Toolbox: DNA binding structural alert from OECD		4.3/4.4	Chemical reactivity (endpoint agnostic)
	OECD QSAR Toolbox: DNA binding structural alert for CA from OASIS		4.3/4.4	Mutagenicity
	OECD QSAR Toolbox: *in vitro* mutagenicity (Ames) by ISS		4.3/4.4	Mutagenicity
	OECD QSAR Toolbox: DNA Alerts for AMES from OASIS		4.3/4.4	Mutagenicity
	OECD QSAR Toolbox: OASIS Micronucleus test (MNT)		4.3/4.4	Mutagenicity
	OECD QSAR Toolbox: *in vivo* mutagenicity (Micronucleus) alerts by ISS		4.3/4.4	Mutagenicity
	TIMES Transgenic Rodent (TGR) Mutation Assay		2.30.1	Mutagenicity
	TIMES Comet Assay		2.30.1	Mutagenicity
	TIMES Chromosomal aberration		2.29.1	Mutagenicity
	TIMES Micronucleus test (MNT)		2.29.1	Mutagenicity
	OECD QSAR Toolbox: Protein binding structural alert from OASIS		4.3/4.4	Chemical reactivity (endpoint agnostic)

ERC2 Descriptor/Use	*In silico* Model	Empirical Data Source/Database	Model Version/Date	Use in ERC2
	OECD QSAR Toolbox: Protein binding alerts for skin sensitization according to GSH		4.3/4.4	Chemical reactivity (endpoint agnostic)
	OECD QSAR Toolbox: Protein binding alerts for skin sensitization by OASIS		4.3/4.4	Chemical reactivity (endpoint agnostic)
	OECD QSAR Toolbox: Protein binding by OASIS		4.3/4.4	Chemical reactivity (endpoint agnostic)
	OECD QSAR Toolbox: Protein binding by OECD		4.3/4.4	Chemical reactivity (endpoint agnostic)
	OECD QSAR Toolbox: Protein binding potency Cys (DPRA 13%)		4.3/4.4	Chemical reactivity (endpoint agnostic)
	OECD QSAR Toolbox: Protein binding potency Lys (DPRA 13%)		4.3/4.4	Chemical reactivity (endpoint agnostic)
	OECD QSAR Toolbox: Protein binding potency GSH		4.3/4.4	Chemical reactivity (endpoint agnostic)
	Development and Reproductive Toxicity (DART) Scheme		1.4	DART

ERC2 Descriptor/Use	*In silico* Model	Empirical Data Source/Database	Model Version/Date	Use in ERC2
		OECD QSAR Toolbox *in vitro* observed databases: ECHA REACH; Genotoxicity OASIS; Bacterial mutagenicity ISSSTY; Toxicity Japan MHLW; Genotoxicity & Carcinogenicity ECVAM; Genotoxicity pesticides EFSA	4.4	IUCLID6 Picklist T143 v2.0 Gene mutation I Chromosome aberration II (Japan MHLW) Mutagenicity I (ECVAM) Chromosome aberration I (OASIS) Chromosome aberration V (ECVAM) Micronucleus I Chromosome aberration IV (EFSA) Gene mutation II
Mode of Toxic Action	Acute aquatic toxicity classification by Verhaar (modified)		3.2	Determination of mode of action
	Acute aquatic toxicity mode of action by OASIS		3.3	
	US EPA Assessment Tool for Evaluating Risk (ASTER) - Mode of Action		2.0	
	Toxicity Estimation Software Tool (TEST) - Predicted Hazard Class		4.2	
	iSafeRat Mechanism of Action		1.0	
	Uncouplers (MITOTOX)		1.0	
Food Web Toxicity	Risk Assessment IDentification and Ranking (RAIDAR) Model - Hazard Assessment Factor		3.0	
	Iterative Fragment Selection (IFS) QSAR: Fish and Mammal		B.0	Metabolism Rate

ERC2 Descriptor/Use	*In silico* Model	Empirical Data Source/Database	Model Version/Date	Use in ERC2
***In vivo* Mammalian Toxicity (via OECD QSAR Toolbox)**		Food TOX Hazard EFSA ECHA REACH MUNRO non-cancer EFSA Repeated Dose Toxicity HESS ECOTOX Developmental toxicity ILSI Genotox and Carcinogenicity ECVAM Genotoxicity pesticides EFSA Micronucleus ISSMC Micronucleas OASIS Toxicity Japan MHLW Transgenic Rodent Database ToxRefDB US-EPA Rep Dose Fraunhofer ITEM	OECD QSAR Toolbox 4.4	DART, neurotoxicity
***In vivo* Aquatic Toxicity (via OECD QSAR Toolbox)**		ECOTOX ECHA REACH Aquatic ECETOC Aquatic Japan MoE Aquatic OASIS Food TOX Hazard EFSA	OECD QSAR Toolbox 4.4	Acute and chronic aquatic toxicity (lethal and non-lethal endpoints)
Predicted Acute Fish Toxicity (LC50)	Artificial Intelligence Expert Predictive System (AIEPS); Toxicity Estimation Software Tool (TEST) MoA and Consensus QSARs; ECOSAR; ASTER; TIMES OASIS Fathead Minnow; TOPKAT; ACD Labs		3.0, 4.2.1; 4.11; ASTER 2012; 0.05; 5.01; 2020.1.2	Fish LC50 for MoA tissue residue methods
Response Time	Risk Assessment IDentification and Ranking (RAIDAR)		3.0	Calculation of overall persistence (P_{ov})

ERC2 Descriptor/Use	*In silico* Model	Empirical Data Source/Database	Model Version/Date	Use in ERC2
Mobility	Risk Assessment IDentification and Ranking (RAIDAR)		3.0	Calculation of characteristic travel distance
Tonnage		Canadian DSL tonnage data 1986	n/a	Calculation of mean annual tonnage (kt/yr); predicted emission rate (kt/hr)
Food Web Exposure	Risk Assessment IDentification and Ranking (RAIDAR)		3.0	Calculation of food web tissue residues
Margin of Exposure	Risk Assessment IDentification and Ranking (RAIDAR)		3.0	Calculation of critical emission rate
Use Pattern	US EPA EXPOCAST QSUR (quantitative structure-use relationship) model	US EPA Chemical and Products Database (CPDat); ECCC use pattern analysis for DSL chemicals	2016; 2018; 2015	Known and predicted substance use pattern

References

ACD/Percepta. c1997-2020. Toronto (ON): Advanced Chemistry Development, Inc.

[AIEPS] Artificial Intelligence Expert Predictive System. c2010-2012. Ver. 3.0. Gatineau (QC): Environment Canada. Model developed by Stephen Niculescu.

Akahori Y, Nakai M, Yamasaki K, Takatsuki M, Shimohigashi Y, Ohtaki M. 2008. Relationship between the results of *in vitro* receptor binding assay to human estrogen receptor α and *in vivo* uterotrophic assay: Comparative study with 65 selected chemicals. Toxicol *In vitro*. 22:225-231.

Ames Test ISSSTY [database of *in vitro* mutagenicity results (Ames test) results]. 2011. Ver. V4b. Rome (IT): Istituto Superiore di Sanità (ISS); Liebefeld (CH): Federal Office of Public Health. [Benigni R, Battistelli CL, Bossa C, Tcheremenskaia O, Crettaz P. 2013. New perspectives in toxicological information management, and the role of ISSTOX databases in assessing chemical mutagenicity and carcinogenicity. Mutagenesis. 28(4):401-409].

[ARC] ARC Arnot Research & Consulting. 2018. Generation of physical-chemical property data and the application of models for estimating fate and transport and exposure and risk potential for organic substances on the Canadian DSL. Gatineau (QC): Environment and Climate Change Canada.

Arnot JA, Gouin T, Mackay D (Canadian Environmental Modelling Network (CEMN, Trent University, Peterborough, ON). 2005. Development and fate of models of chemical fate in Canada. Practical methods for estimating environmental biodegradation rates. Gatineau (QC): Environment Canada. Report No. 200503. 50 p.

[ASTER] Assessment Tools for the Evaluation of Risk. 2012. Ver 2.0. Duluth (MN): US Environmental Protection Agency, Mid-Continent Ecology Division. [restricted access].

Bassan A, Fioravanzo E, Pavan M, Stocchero M. 2011. Applicability of physicochemical data, QSARs and read-across in Threshold of Toxicological Concern assessment. Parma (IT): European Food Safety Authority.

CATALOGIC [environmental fate and ecotoxicity model]. 2014. Ver. 5.11.15. Bourgas (BG): University "Prof. Dr. Assen Zlatarov", Laboratory of Mathematical Chemistry.

[CERAPP] Collaborative estrogen receptor activity prediction project. c2020. San Francisco (CA): OPEn structure–activity/property Relationship App (OPERA), Github, Inc. [Model described in Mansouri K, Abdelaziz A, Rybacka A, Roncaglioni A, Tropsha A, Varnek A, Zakharov A, Worth A, Richard AM, Grulke CM, et al. 2016. CERAPP: Collaborative estrogen receptor activity predication project. Environ Health Persp. 124(7):1023-1033.].

[CHRIP] Japanese Chemical Risk Information Platform Aquatic MoE. [database]. 2008. Ver. 1.3. Tokyo (JP): Ministry of Environment, National Institute of Technology and Evaluation (NITE).

[CoMPARA] Collaborative modeling project for androgen receptor activity. c2020. San Francisco (CA): OPEn structure–activity/property Relationship App (OPERA), Github, Inc. [Mansouri K, Kleinstreuer N, Grulke C, Richard A, Shah I, Williams AJ, Judson RS. 2018. Virtual screening of chemicals for endocrine disrupting activity through CERAPP and CoMPARA projects [poster]. Society of Toxicology; 2018 March 11-15; San Antonio, TX.].

[DART] Developmental And Reproductive Toxicology Scheme. 2020. Ver. 1.4. Cincinnati (OH): Proctor & Gamble; Bourgas (BG): Laboratory of Mathematical Chemistry.

Ding D, Xu L, Fang H, Hong H, Perkins R, Harris S, Bearden ED, Shi L, Tong W. 2010. The EDKB: an established knowledge base for endocrine disrupting chemicals. BMC Bioinformatics. 11 Suppl 6:S5.

[ECETOC] European Center of Ecotoxicology and Toxicology. 2003. Aquatic hazard assessment II. Technical report. Brussels (BG): European Center of Ecotoxicology and Toxicology. Report No.: 91.

[ECHA] European Chemicals Agency. c2007-2020. Registered substances database, Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Helsinki (FI): ECHA. [updated 2020 November 9.

[ECOSAR] ECOlogical Structure Activity Relationships Class Program [estimation model]. 2012. Ver. 1.11. Washington (DC): US Environmental Protection Agency, Office of Pollution Prevention and Toxics; Syracuse (NY): Syracuse Research Corporation.

[ECOTOX] ECOTOXicology database [database]. 2020. Ver. 5.3. Washington (DC): US Environmental Protection Agency, Office of Research and Development; National Health and Environmental Effects Research Laboratory; Mid-Continent Ecology Division. [updated 2020 December 15].

[EFSA] European Food Safety Authority Genotoxicity Pesticides. [database]. 2017. Ver. 1.0. Parma (IT): European Food Safety Authority. [Metruccio F, Castelli I, Civitella C, Galbusera C, Galimberti F, Tosti L, Moretto A. 2017. Compilation of a database, specific for the pesticide active substance and their metabolites, comprising the main genotoxicity endpoints. EFSA supporting publication 2017. 125 p.]

[EFSA] European Food Safety Authority OpenFoodTox. [database]. 2020. Parma (IT): European Food Safety Authority.

Environment Canada. 1988. Data relating to the Domestic Substances List (DSL) 1984-1986, collected under CEPA, 1988, s.25(1). Based on Reporting for the Domestic Substances List [guide] 1988. Data prepared by: Environment Canada.

[EPI Suite] Estimation Program Interface Suite for Microsoft Windows [estimation model]. c2000-2012. Ver. 4.11. Washington (DC): US Environmental Protection Agency, Office of Pollution Prevention and Toxics; Syracuse (NY): Syracuse Research Corporation.

[EURL EVCAM] EU Reference Laboratory for Alternatives to Animal Testing. [database]. 2015. Ispra (IT): European Commission, Joint Research Centre.

[HESS] Hazard Evaluation Support System. [database]. 2010. Tokyo (JP): Safety Assessment Division, Chemical Management Center, National Institute of Technology and Evaluation (NITE). [Sakuratani Y, Zhang HQ, Nishikawa S, Yamazaki K, Yamada T, Yamada J, Gerova K, Chankov G, Mekenyan O, Hayashi M. 2013. Hazard Evaluation Support System (HESS) for predicting repeated dose toxicity using toxicological categories. SAR QSAR Environ Res. 24(5):351-363].

[IFSAPP] Iterative Fragment Selection – Fish. 2018. Toronto (ON): ARC Arnot Research.

[iSafeRat®] *in silico* Algorithms For Environmental Risk And Toxicity. c2019. Ver. 1.0. L'Isle d'Abeau (FR): KREATiS. [restricted access].

[ISLI DevToxDB] International Life Sciences Institute Developmental Toxicity Database. [database]. 2012. Washington (DC): International Life Sciences Institute.

[JECDB] Japan Existing Chemical Database. [database]. 2018. Japan: Ministry of Health, Labour and Welfare, National Institute of Health Sciences, Division of Risk Assessment.

Lambert IB, Singer TM, Boucher SE, Douglas GR. 2005. Detailed review of transgenic rodent mutation assays. Mutat Res. 590:1-280.

Leadscope Enterprise [data management and decision support platform]. 2016. Ver. 3.5. Columbus (OH): Leadscope, Inc. [restricted access].

Micronucleus ISSMIC [database of *in vivo* micronucleus mutagenicity assay results]. 2011. Ver. v4b. Rome (IT): Istituto Superiore di Sanità (ISS); Liebefeld (CH): Federal Office of Public Health. [Benigni R, Battistelli CL, Bossa C, Tcheremenskaia O, Crettaz P. 2013. New perspectives in toxicological information management, and the role of ISSTOX databases in assessing chemical mutagenicity and carcinogenicity. Mutagenesis. 28(4):401-409].

[MOPAC2016] Molecular Orbital PACkage. 2016. Ver. 2.30.1. Colorado Springs (CO): Stewart Computational Chemistry.

OECD QSAR Toolbox [read-across tool]. 2020. Ver. 4.4. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry.

OECD QSAR Toolbox: Acute aquatic toxicity classification by Verhaar (Modified) [endpoint-specific model]. 2010. Ver. 3.2. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry. [Verharr HJM, van Leeuwen CJ, Hermens JLM. 1992. Classifying environmental pollutants. Chemosphere. 25(4):471-491.].

OECD QSAR Toolbox: Acute aquatic toxicity MOA by OASIS [endpoint specific model]. 2010. Ver. 3.3. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry.

OECD QSAR Toolbox: DNA alerts for AMES, CA and MNT by OASIS [endpoint specific model]. 2013. Ver. 2.7. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry.

OECD QSAR Toolbox: DNA binding by OASIS [mechanistic model]. 2013. Ver. 1.7. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry.

OECD QSAR Toolbox: DNA binding by OECD [mechanistic model]. 2010. Ver. 2.3. Paris (FR): Organisation for Economic Co-operation and Development; Liverpool (UK): School of Pharmacy and Chemistry, Liverpool John Moores University.

OECD QSAR Toolbox: OASIS Estrogen Receptor Binding [mechanistic model]. 2011. Ver. 2.2. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry.

OECD QSAR Toolbox: Protein binding alerts for skin sensitization according to GSH [endpoint-specific model]. 2017. Ver. 1.1. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry.

OECD QSAR Toolbox: Protein binding by OASIS [mechanistic model]. 2010. Ver. 2.7. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry.

OECD QSAR Toolbox: Protein binding by OECD [mechanistic model]. 2011. Ver. 2.3. Paris (FR): Organisation for Economic Co-operation and Development; Liverpool (UK): School of Pharmacy and Chemistry, Liverpool John Moores University.

OECD QSAR Toolbox: Protein binding potency Cys (DPRA 13%) [mechanistic model]. 2012. Ver. 1.0. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry.

OECD QSAR Toolbox: Protein binding potency GSH [mechanistic model]. 2011. Ver. 3.5. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry.

OECD QSAR Toolbox: Protein binding potency Lys (DPRA 13%) [mechanistic model]. 2012. Ver. 1.0. Paris (FR): Organisation for Economic Co-operation and Development, Laboratory of Mathematical Chemistry.

[OPERA] OPEn structure–activity/property Relationship App. c2020. San Francisco (CA): Github, Inc. [Mansouri K, Grulke CM, Judson RS, Williams AJ. 2018. OPERA models for predicting physicochemical properties and environmental fate endpoints. J Cheminform. 10:10.].

[RAIDAR] Risk Assessment IDentification And Ranking [fugacity-based model, including for ionizing chemicals]. 2019. Version 2.99 (beta). Ottawa (ON): Environment and Climate Change Canada (ECCC), Health Canada (HC). Programmed by Jon Arnot (Arnot Research and Consulting) for ECCC and HC.

RepDose. [database]. 2014. Hannover (DE): Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM).

[rtER] Estrogen Receptor Expert System [estimation model]. 2010. Washington (DC): US Environmental Protection Agency, Office of Pollution Prevention and Toxics; Syracuse (NY): Syracuse Research Corporation.

[TEST] Toxicity Estimation Software Tool. 2016. Ver. 4.2. Washington (DC): US Environmental Protection Agency.

[TIMES] TIssue MEtabolism Simulator [prediction module]. 2018. Ver. 2.29.1. Bourgas (BG): University "Prof. Dr. Assen Zlatarov", Laboratory of Mathematical Chemistry.

[TOPKAT] TOxicity Prediction by Komputer Assisted Technology [prediction module]. c2005-2009. Ver. 5.01. San Diego (CA): BIOVIA, Discovery Studio.

[US EPA] US Environmental Protection Agency. 2010. Toxicity Reference Database (ToxRefDB). Washington (DC): US EPA, Office of Research and Development.

[US EPA] US Environmental Protection Agency. 2018. The Chemicals and Products Database (CPDat). Durham (NC): US EPA, Office of Science Information Management.

Wedebye EB, Dybdahl M, Reffstrup TK, Abildgaard Rosenberg S, Løfstedt M, Nikolov NG. 2016. The new Danish (Q)SAR database: A freely available tool with predictions for > 600,000 substances. Toxicol Lett. 258:118-118.

11. APPENDIX II: ERC2 Rules for ADME

Table 1: ADME rules for bioavailability

Rule Number	Rule Description	Bioavailability Assignment	Confidence Category
1	5 Model Consensus: [MW is >750 g/mol] and [Dmax is >2.9nm] and [Deff is >1.4nm] and [multispecies dietary absorption efficiency is <5%] and [MP>300]	Low Bioavailability	Very High
2	Any 4 models agree on low permeability	Low Bioavailability	High
3	Any 3 models agree on low permeability	Low Bioavailability	Moderate
4	Any 2 models agree on low permeability	Low Bioavailability	Low
5	Single model suggests low permeability	Low Bioavailability	Very Low
6	5 Model Consensus: [MW is <750 g/mol] and [Dmax is <2.9nm] and [Deff is <1.4nm] and [multispecies dietary absorption efficiency is >5%] and [MP<300]	High Bioavailability	Very High

Where:

MW = molecular weight (g/mol)

Multispecies dietary absorption efficiency = the percentage absorbed from the diet by invertebrates and vertebrates species represented in the foodweb of version 3.0 of the RAIDAR model (based on Kelly et al. 2004)

D_{max} = average maximum cross-sectional diameter of a chemical substance (nm)

D_{eff} = average effective diameter of a chemical substance (nm)

MP = melting point (^{o}C)

The combination of the above rules is based on evidence that suggests that molecular dimensions (steric hindrance) as well as equilibrium partitioning resistance of super hydrophobic molecules can significantly reduce the rate of tissue permeation in the gut as well as other tissues such as gills and skin (Gobas et al. 1986; Arnot and Gobas 2006; Dimitrov et al. 2002; Dimitrov et al. 2003; Dimitrov et al. 2005; Sakuratani 2008). Confidence was assigned according to the degree of model consensus as outlined in Table 1. These rules do not suggest that there is an apparent strict threshold for permeation as noted in Arnot et al. (2009), but rather that the rate of permeation is significantly restricted such that other ADME processes reduce body burdens resulting in a lower potential for critical toxicity. The rules also isolate substances with low environmental availability due to being irreversibly bound to solid

particles that even upon ingestion are not expected to partition into the organism (including other non-passive routes of uptake). The combination of the molecular descriptors used in the rules in Table 1 also isolate many solid chemicals which have very low solubility in lipids, carbon and water and are expected to have low internal and external bioavailability.

Table 2: ADME consensus rules for plasma distribution

Rule Description	Rule Number	Confidence	Approach
3 Model Consensus: [PPB >80%] and [logK_a HSA >=4] and Vd [<3]	1	High	Model Consensus
Not suspected to be plasma bound	2	High	Model Consensus

Where:

PPB = PPB (%) values represent the probability of overall fraction of substance bound in human plasma, (i.e. accounts for interactions with different proteins: albumin, alpha1-acid glycoprotein, lipoproteins, SHBG, transcortin etc. %PPB is calculated as (1 – fu) * 100%, where fu is fraction of free (unbound) drug in plasma ranging from 0 to 1.)

LogK_a HSA = a substance's affinity constant to human serum albumin – the major carrier protein in plasma and is calculated: log ([LA]/([L][A])), where [LA] is concentration of ligand bound to albumin, [L] – that of free ligand, and [A] – concentration of free albumin which is estimated at ~0.6 mM in human plasma

Vd = the apparent Volume of Distribution (L/kg). This is a measure of how much a substance must be distributed throughout the whole body (human) to provide the same concentration in blood plasma. It is calculated as the ratio of the drug in the body divided by the plasma concentration

The rules in Table 2 above combine measures of the probability (>80%) of blood protein plasma binding (PPB), binding probability to serum albumin in blood (LogK_a HSA >=4), the major carrier of protein plasma, and volume of distribution (Vd < 3 L/kg). The rules were established based on domain of application information from ACD 2019[24] largely from pharmaceuticals. A PPB of >80% represents a high probability of PPB given the range of values in the training set. Reliability index for PPB must be >0.7 or the result was considered out of domain and not used in ERC2. Similarly, a logK_a HSA value of >=4 was selected based on the upper range of binding within the experimental data used to develop the model.

[24] http://perceptahelp.acdlabs.com/help_v2019/index.php/Protein_Binding. http://perceptahelp.acdlabs.com/help_v2019/index.php/Vd
"Experimental data that were utilized to build predictive models were collected from drug prescription information, reference pharmacokinetic tabulations and many original articles. The main sources of Vd data were well-known pharmacokinetic books: "Therapeutic Drugs" (ed. by C. dollery), and Goodman & Gilman's "The Pharmacological Basis of Therapeutics", while albumin affinity constants were collected mainly from original articles by Valko K. et al. J Pharm Sci. 2003;92(11):2236-48., and Kratochwil N.A. et al. Biochem Pharmacol. 2002;64(9):1355-74. [2] The compiled data sets contain %PPB data for almost 1500 chemicals, about 340 albumin affinity constants and almost 800 Vd values."

Reliability Index for $LogK_a$ HSA was not used due to low training set coverage, but good agreement with PPB and Vd outcomes. A Vd (L/kg) value <3 was selected to represent the range below which the majority fraction of a substance's distribution in the body will be associated with plasma (e.g., Vd values of >200 L/kg are required for distribution to include target lipids). Confidence with plasma distribution determinations was assigned high when all three measures of plasma distribution the PPB reliability index agreed (Table 2), else a substance was not suspected to be plasma bound.

References

ACD/Percepta [prediction module]. 2019. Toronto (ON): Advanced Chemistry Development, Inc.

Arnot J, Gobas FAPC. 2006. A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in fish. Environ Rev. 14:257-297.

Arnot JA, Arnot M, Mackay D, Couillard Y, MacDonald D, Bonnell M, Doyle P. 2009. Molecular size cut-off criteria for screening bioaccumulation potential: fact or fiction? Integr Environ Assess Manage. 6(2):210-224.

Dimitrov SD, Dimitrova NC, Walker JD, Veith GD, Mekenyan OG. 2002. Predicting bioconcentration factors of highly hydrophobic chemicals. Effects of molecular size. Pure Appl Chem. 74:1823–1830.

Dimitrov SD, Dimitrova NC, Walker JD, Veith GD, Mekenyan OG. 2003. Bioconcentration potential predictions based on molecular attributes—An early warning approach for chemicals found in humans, birds, fish and wildlife. QSAR Combinatorial Sci. 22:58–68.

Dimitrov S, Dimitrova N, Parkerton TF, Comber M, Bonnell M, Mekenyan O. 2005. Base-line model for identifying the bioaccumulation potential of chemicals. SAR QSAR Environ Res. 16:531–554.

Gobas FAPC, Opperhuizen A, Hutzinger O. 1986. Bioconcentration of hydrophobic chemicals in fish: Relationship with membrane permeation. Environ Toxicol Chem. 5:637–646.

Kelly, B.C., Gobas, F.A.P.C. and McLachlan, M.S. 2004. Intestinal absorption and biomagnification of organic contaminants in fish, wildlife, and humans. Environmental Toxicology and Chemistry, 23: 2324-2336.

Kelly BC, Ikonomou MG, Blair JD, Morin AE, Gobas FA. 2007. Food web-specific biomagnification of persistent organic pollutants. Science. 317(5835): 236-9.

Sakuratani Y, Noguchi Y, Kobayashi K, Yamada J, Nishihara T. 2008. Molecular size as a limiting characteristic for bioconcentration in fish. J Environ Biol. 29:89–92.

12. APPENDIX III: Rules for Classification and Confidence Scoring of Receptor-Mediated Interactions

1. In silico:

The prediction of endocrine activity was profiled using a combination of the best model and consensus model approaches based on the in domain output of various *in silico* tools. Table 1 below describes the potency-based hazard classification rules for receptor-mediated interactions both from a qualitative perspective (weak, moderate, strong, positive, negative, etc.) and relative binding affinity (RBA) relative to beta-estradiol (estrogen receptor [ER]) and testosterone (androgen receptor [AR]) depending on the type of model output. In some cases, both types of output are available (e.g., TIMES model). Table 2 lists the *in silico* model-consensus rules dictating confidence assignment to hazard classification based on *in silico* results only. Note that both the parent molecule and metabolites (via S9 activation) are profiled and considered for activity in the TIMES ER model. The CERAPP and CoMPARA[25] models were given the highest weighting in the confidence scoring rule structure outlined in Table 2 (rule number 1). These models represent the most advanced *in silico* tools available and were developed as a consensus of over 40 models for ER and ~90 model for AR (i.e., multi-model consensus), with both models showing high prediction reliabilities across the model evaluation sets used (Mansouri et al. 2016; Mansouri et al. 2020). CERAPP and CoMPARA estrogen and androgen receptor interactions (binding, agonist, antagonist) results were previously generated by the model developer for ECCC outside of the now available OPERA model (Mansouri et al. 2018). If results were not available from CERAPP or CoMPARA (<15%) or models suggest very weak or no activity for ER or AR interactions, and to account for thyroid and aryl hydrocarbon receptor (AhR) activity as well as S9 metabolites, subsequent model consensus rules were triggered (Table 2: Rules 2-7). Positive results of binding for thyroid peroxidase (TPO) and AhR dictated hazard classification only if very weak, inactive or no binding potential resulted ER or AR data. Otherwise, TPO and AhR were added to possible target interactions alongside ER and AR.

Table 1: Potency based hazard classification rules for *in silico* prediction of receptor-mediated effects

Rule Description	Relative Binding Affinity (RBA)	Hazard Classification
"Very strong, strong, or moderate" receptor binding to ER or AR	RBA >= 0.01	3
"Weak" ER or AR binding	RBA >0.001-0.01	2
"Very weak", "inactive", "no" ER or AR binding	RBA <0.001	1

[25] CERAPP: Collaborative Estrogen Receptor Activity Prediction Project, CoMPARA. COMPARA: Collaborative Modeling Project for Androgen Receptor Activity

“Positive_in” TPO inhibition in either or both of two models	Positive	3
“Negative_in” TPO inhibition in either or both of two models	Negative	1
"Active" AhR binding	Positive	3
"Not active" AhR binding	Negative	1

Abbreviations: RBA, relative binding affinity; TPO, thyroid peroxidase; ER, estrogen receptor; AR, androgen receptor; AhR, aryl hydrocarbon receptor

Where available, domain of applicability results were taken into account according to individual model domain boundaries. The following model domain considerations were taken:

- Where available (OPERA output only), results from CERAPP and CoMPARA were accepted if applicability domain index was >=0.6 and confidence index was >=0.5
- ACD LogRBA ER reliability index (RI) is >= 0.3 and probability of LogRBA is >0.5, else considered non-binder or out of domain. ACD ER model does not use the same potency scale as other ER/AR models, thus strong binder outcomes had RBA ER set to 0.5 and weak binder RBA ER set to 0.005
- TIMES model results must be in parameter domain (phys-chem) and have correct structural fragments >=60%. If result is reported “active”, RBA ER/AR must be >=0.001
- Thyroid peroxidase inhibition results from the Danish QSAR database must be in domain (POS_IN or NEG_IN)
- Domain boundaries not available for thyroid alpha and beta receptor binding at time of computation. Results not used.

Table 2: Model Consensus Rules for *in silico* confidence scoring of hazard classification outcomes

Rule Description	Rule Number	Confidence Score	Model Reference
Multi-model Consensus: CERAPP (ER) and/or CoMPARA (AR) for binding, agonist or antagonist result is weak binding or stronger	1	5	Mansouri et al. 2016; Mansouri et al. 2018; Mansouri et al. 2020
4 Model Consensus: [TIMES ER parent or metabolite] and [OECD Toolbox ER] and [ACD logRBA ER] and [Toolbox rtER has ‘mechanism identified’]	2	5	TIMES 2018; Serafimova et al. 2007; Mekenyan and Serafimova 2009; ACD Labs Percepta 2018; OECD QSAR Toolbox v4.3/4.4 2018-2020; Denny et al. 2015; Schmieder et al. 2016

3 Model Consensus: [TIMES ER parent or metabolite] and/or [OECD Toolbox ER] and/or [ACD logRBA] and/or [Toolbox rtER has 'mechanism identified']	3	4	TIMES 2018; Serafimova et al. 2007; Mekenyan and Serafimova 2009; ACD Labs Percepta 2018; OECD QSAR Toolbox v4.3/4.4 2018-2020; Denny et al. 2015; Schmieder et al. 2016
2 Model Consensus: [TIMES ER parent or metabolite] and/or [OECD Toolbox ER] and/or [ACD logRBA ER] and/or [Toolbox rtER has 'mechanism identified']]	4	3	TIMES 2018; Serafimova et al. 2007; Mekenyan and Serafimova 2009; ACD Labs Percepta 2018; OECD QSAR Toolbox v4.3/4.4 2018-2020; Denny et al. 2015; Schmieder et al. 2016
Single Model ER: [TIMES ER parent or metabolite or [OECD Toolbox ER] or [ACD logRBA ER]or [Toolbox rtER has 'mechanism identified']	5	2	TIMES 2018; Serafimova et al. 2007; Mekenyan and Serafimova 2009; ACD Labs Percepta 2018; OECD QSAR Toolbox v4.3/4.4 2018-2020; Denny et al. 2015; Schmieder et al. 2016
Single Model AR: TIMES Model AR parent is active (must be NA in Rule 1-4)	5	2	TIMES 2018; Todorov et al. 2011
Single Model AhR: TIMES AhR parent Active	6	2	TIMES 2018; Todorov et al. 2011
Single Model Thyroid: Danish TPO is in domain and is positive	7	2	Rosenberg et al. 2017

Abbreviations: RBA, relative binding affinity; TPO, thyroid peroxidase; ER, estrogen receptor; AR, androgen receptor; AhR, aryl hydrocarbon receptor

2. In vitro

In vitro data for ER and AR and thyroid were available from several databases and model training sets (knowledge bases). The database sources of *in vitro* data used in ERC2 for receptor-mediated interactions are listed in Appendix 1. The rule structure for hazard classification and confidence scoring is similar to those developed for *in silico* profiling (*in silico* models are trained using *in vitro* data). Hazard classification is also potency driven using RBA scales as well as qualitative categorical; therefore, the hazard classification rules for *in vitro* data are the same as those used for *in silico* data in Table 1. However, positive results of binding for TPO and AhR dictated hazard classification only if very weak, inactive or no binding potential resulted in ER or AR data. Table 3 below summarizes the data consensus rules used for scoring classification confidence.

Table 3: Rules for confidence scoring of *in vitro* hazard classification for receptor-mediated interactions

Rule No.	*In vitro* Confidence Scoring Rules	Confidence Class
1	CERAPP and/or CoMPARA consensus observed value (binding and/or agonist and/or antagonist)	5
2	Single Observed value (ER)	3
3	Single Observed value (AR)	3
4	Single observed value (THY)	3
5	Single observed value (AhR)	3

Abbreviations: TPO, thyroid peroxidase; ER, estrogen receptor; AR, androgen receptor; AhR, aryl hydrocarbon receptor; THY, ????

The CERAPP and CoMPARA observed evaluation databases (including TOXCAST AC50 bioactivity values normalized to potency scales) used for *in silico* model development contained the highest number of values of any database used in this analysis (~81%). They have both undergone consensus analysis, curation and potency binning by the authors. The data largely come from the US EPA's NCCT collected and curated PubChem data (64 sources) (Mansouri et al. 2016; Mansouri et al. 2020). Similar to the *in silico* rules, these databases received high confidence scoring in Table 3 and dictated most of the *in vitro* hazard classification outcomes for estrogen and androgen receptor interactions. When substances in ERC2 were not available in the CERAPP and CoMPARA databases, single observed values from other databases or *in silico* QSAR training sets were selected (e.g., TIMES ER and AR).

3. In vivo

Mammalian or aquatic effects data were available for approximately 6% of the substances in ERC2. Final classification of receptor-mediated interactions was therefore not heavily influenced by these data. *In vivo* data were collected using the OECD QSAR Toolbox v4.3/4.4. The Toolbox contains a high degree of toxicity data that are publically obtainable and was a practical tool for extracting data from several databases for the ~12 200 substances. The databases searched by the Toolbox for *in vivo* data are listed in Appendix I. Table 4-6 outline the rules used for parsing and screening the endpoints and species considered both for mammalian and acute and chronic aquatic organisms, noting that for mammalian data, effects that might be considered "adaptive" by some agencies were also included given the varying opinions on the definition of "adaptive" and to take precaution from excluding effects data. Table 7(a,b,c) lists the rules used *in vivo* hazard classification and confidence scoring for mammalian and avian, chronic and acute receptors, for both quantitative and categorical data (i.e., DART data), respectively.

Table 4: Summary of data selection rules for vertebrate data

Rule No.	Description	Endpoint Targets	Test organisms (species)
1	LOEL used preferentially	Developmental Toxicity / Teratogenicity Repeated Dose Toxicity Toxicity to Reproduction	Rat; Dog; Rabbit; Mouse; Mus musculus; Rattus norvegicus; Peromyscus maniculatus;

			Monkey; Mallard duck; Mesocricetus auratus; Bos taurus; Oryctolagus cuniculus; Bobwhite quail Ochotona rufescens ssp. Rufescens; Sus scrofa Capra hircus; Bubalus bubalis Felis catus; Octodon degus Ovis aries; Japanese quail Common quail; Microtus pinetorum; Cavia porcellus Oryctolagus; Microtus socialis Cat; Primate
2	If LOEL =no data, then NOEL used		
3	mg/kg or equivalent selected		
4	Vertebrates only		
5	>24 hr duration only		
6	Points of departure = LOEL, LOAEL, LOAEC, LOEC, NOEL, NOAEL, NOAEC, NOEC		
7	All endpoints accepted for preliminary analysis except mortality		
8	No unbounded values for LOEL, LOAEL, LOAEC, LOEC; no < or </= for NOEL, NOAEL, NOAEC, NOEC,		
9	Undefined dose data removed		
10	5th percentile of data taken for all points of departure when multiple values available for single CAS RN; else single value used		

Table 5: Summary of data selection rules for chronic aquatic data

Rule No.	Rule Description	Endpoints	Phyla
1	LOEC or equivalent (e.g., >EC20 selected preferentially; if no LOEC, then NOEC or	Reproduction; Immobilization Behaviour; Development Growth; Feeding Behaviour Morphology; Avoidance Physiology; Mobility	Crustacea; Echinozoa; Vertebrata; Hexapoda; Rhabditophora; Neodermata; Medusozoa; Anthozoa Various invertebrates and vertebrate species in QSAR Toolbox associated with

	equivalent selected (e.g., EC10-EC20)		endpoints selected (n=485) (OECD QSAR Toolbox v4.4.1)
2	No unbounded values accepted		
3	mg/L, mg/kg or equivalent		
4	Exposure duration in hours, days, weeks, months only		
5	Invertebrates and vertebrates only (no bacteria, no plants/algae)		

Table 6: Summary of data selection rules for acute aquatic data

Rule No.	Rule Description	Endpoints	Species: Invertebrates/ Vertebrates
1	5th percentile selected when multiple data for single CAS RN	Development; Physiology; Growth; Morphology Population; Growth Rate; Reproduction; Number Hatched; Weight; Immunological; Immobilisation; Frond Number	Various in vertebrate and vertebrates species (n=124)
2	Acute defined as >=24 hours, but <= 7days		
3	No unbounded values		
4	mg/L or equivalent only		
5	Point of departures: >= EC25 or IC25, LOEC, MATC		
6	WAF (water accommodated fraction) studies accepted		
7	No algae, bacteria or protozoa		

Table 7a: Summary of hazard classification and confidence rules for mammalian *in vivo* toxicity

Numerical and Categorical Mammalian Hazard Classification Rules		
Rule No.	**Rule Description**	**Hazard Classification**
1	Value is <=10 mg/kg bw or DART result is positive	3
2	Value is >10 to <= 500 mg/kg bw	2
3	Value is >500 mg/kg bw or DART result is negative	1
Numerical and Categorical Data Confidence Rules		
Rule No.	**Rule Description**	**Confidence Score**
1	5^{th} percentile of numerical data **OR** "pos" or "neg" categorical data with no equivocal results	26
2	single numerical value or **OR** "pos" or "neg" categorical data with equivocal results	21

Note: When both numerical and categorical data are both available for a substance, numerical data were selected, preferentially.

7b: Summary of hazard classification and confidence rules for chronic *in vivo* aquatic toxicity

Numerical Chronic Aquatic Hazard Classification Rules		
Rule No.	**Rule Description**	**Hazard Classification**
1	Value is <=10 mg/L	3
2	Value is >10 to <100 mg/L	2
3	Value is >=100 mg/L	1
Confidence Scoring Rules		
Rule No.	**Rule Description**	**Confidence Score**
1	5th percentile of data	26
2	Single value	21

7c: Summary of hazard classification and confidence rules for acute *in vivo* aquatic toxicity

Numerical Acute Aquatic Hazard Classification Rules		
Rule No.	**Rule Description**	**Hazard Classification**
1	Lowest EC50 <=0.1 mg/L	3
2	Lowest EC50 >0.1 to <=1.0 mg/L	2
3	Lowest EC50 >1 mg/L	1

Confidence Scoring Rules		
Rule No.	**Rule Description**	**Confidence Score**
1	5th percentile	26
2	Single value	21

The numerical toxicity rules in Table 7a were established specifically for ERC2, but are based on regulatory hazard toxicity classes commonly used for mammalian lethal oral dose data (e.g., Hodge and Sterner scale; Gosselin, Smith and Hodge)[26]. The numerical chronic and acute toxicity rules in Table 7b and 7c for aquatic receptors were established specifically for ERC2. The chronic toxicity thresholds were selected to approximate the numerical toxicity thresholds. The acute toxicity thresholds in Table 7c are generally much lower than those selected in past Canadian prioritization schemes such as the 2006 categorization of the DSL. These thresholds were selected to capture chemicals that even under acute exposure durations, often one to 4 days, developmental or reproductive effects or lethality are observed. That is, acute *in vivo* hazard classification prioritizes chemicals with potencies greater than baseline narcosis. Finally, the confidence rules in all tables above reflect the amount of toxicological data available for specific endpoints, where single values are treated as being less certain of reflecting the variability of toxicological data and are thus down weighted.

References

ACD/Percepta [prediction module]. c1997-2020. Toronto (ON): Advanced Chemistry Development, Inc.

Denny J, Kolanczyk R, Hornung M, Sheedy B, Serrano J, Schmieder P. 2015. Estrogen receptor expert system overview and examples [presentation]. Mid West Society of Environmental Toxicology and Chemisty (MW SETAC) Meeting; 2015 March 9-11; Duluth, Minnesota.

[ECCC, HC] Environment and Climate Change Canada, Health Canada. [modified 2017 Mar 12]. Categorization. Ottawa (ON): Government of Canada.

Mansouri K, Kleinstreuer N, Abdelaziz AM, Alberga D, Alves VM, Andersson PL, Andrade CH, Bai F, Balabin I, Ballabio D, et al. 2020. CoMPARA: Collaborative modeling project for androgen receptor activity. Environ Health Perspect. 128:2.

Mansouri K, Abdelaziz A, Rybacka A, Roncaglioni A, Tropsha A, Varnek A, Zakharov A, Worth A, Richard AM, Grulke CM. 2016. CERAPP: Collaborative estrogen receptor activity prediction project. Environ Health Perspect. 124:1023-1033.

[26] https://www.ccohs.ca/oshanswers/chemicals/ld50.html

Mansouri K, Grulke CM, Judson RS, Williams AJ. 2018. OPERA models for predicting physicochemical properties and environmental fate endpoints. J Cheminform. 10:10.

Mekenyan O, Serafimova R. 2009. Mechanism-based modeling of estrogen receptor binding affinity a COREPA implementation. Boca Raton (FL): CRC Press. p. 259-293.

Rosenberg SA, Watt ED, Judson RS, Simmons SO, Friedman KP, Dybdahl M, Nikolov NG, Wedebye EB. 2017. QSAR models for thyroperoxidase inhibition and screening of U.S. and EU chemical inventories. Computat Toxicol. 4:11-21.

Schmieder P, Kolanczyk R, Hornung M, Tapper M, Denny J, Veith G. US EPA Estrogen Receptor Expert System (ERES). 2016. Duluth (MN): US EPA, Office of Research and Development.

Serafimova R, Todorov M, Nedelcheva D, Pavlov T, Akahori Y, Nakai M, Mekenyan O. 2007. QSAR and mechanistic interpretation of estrogen receptor binding. SAR QSAR Environ Res. 18(3-4):1-33.

Todorov M, Mombelli E, Ait-Aissa S, Mekenyan O. 2011. Androgen receptor binding affinity: A QSAR evaluation. SAR QSAR Environ Res. 22(3-4):265-291.

[TIMES] TIssue MEtabolism Simulator [prediction module]. 2018. Ver. 2.27.19. Bourgas (BG): University "Prof. Dr. Assen Zlatarov", Laboratory of Mathematical Chemistry.

13. APPENDIX IV: Rules for Classification and Confidence Scoring Chemical Reactivity and Genotoxicity

1. In silico:

The *in silico* prediction of genotoxicity was determined using a consensus of mechanistic profiling approaches largely contained in the OECD QSAR Toolbox v4.2-4.4 as well as mechanistic QSARs for genotoxicity (e.g., DNA damage, chromosomal aberrations, gene mutation). Most of these interactions are not potency-based. Positive interactions are indicated as 'structural alerts' by the Toolbox profilers when a substance contains a functional group(s) of concern known for initiating key molecular events associated with genotoxicity (e.g., protein or DNA binding, micronucleus development, DNA damage). Table 1 below outlines the hazard classification rules for genotoxicity. The classification scheme in Table 1 reflects cascading consensus rules for various DNA and protein interactions (molecular events) using 14 different targets (listed in Table 2). Greater confidence of a mutagenic outcome is associated when both DNA and protein interactions are possible. Potency-based *in silico* approaches are also available for protein binding. Binding potency was therefore used as an additional descriptor with evidence of a positive protein interaction to classify hazard (as high). However, given that most of the *in silico* profiling approaches for genotoxicity are not potency-based (i.e., are categorical), hazard classification in Table 1 generally reflects the degree of consensus that describes the plausibility of DNA and protein interactions.

Table 2 summarizes the rules used for confidence scoring based on model consensus. Short model descriptions are listed in Appendix I. Some of the approaches in Table 2 are used in ERC2 as endpoint agnostic (e.g., protein and DNA binding and potency). Some of these approaches are based on skin sensitization in humans (e.g., h-LAT). Nonetheless, they indicate a potential chemical reactivity, largely from covalent interactions. Moderate and strong interactions for protein potency for GSH interactions are categorical, but for h-CLAT and DPRA (Cysteine and Lysine) are given quantitatively based on *in chemico* studies. Therefore, 9-21% reactivity and >21% reactivity refers to moderate and strong interactions for both *in silico* and *in chemico* interactions (Table 3), respectively.

Table 1: Hazard *in silico* classification rules for genotoxicity

Rule No.	Rule Description	Hazard Classification
1	Positive protein **AND** DNA interactions	3
2	Protein interaction is positive **AND** protein potency is moderate or strong	3
3	Positive protein **OR** DNA interactions **OR** has positive DART mechanism	2
4	No DNA or protein interactions for any targets	1

Table 2: Hazard classification rules for *in silico* prediction of genotoxicity from interactions with proteins and DNA

Rule No.	Rule Description	Model Consensus	Confidence Score	Reference of Model or Approach
	Rules for Protein Binding Potency			
1	[GSH] AND [h-CLAT] AND [DPRA Cys] AND [DPRA Lys]	4 Model Consensus	5	OECD Toolbox v4.4; Gerberik et al. 2014; Natsch et al. 2008; Natsch et al. 2015; Jaworska et al. 2015; Dimitrov et al. 2016
2	[GSH] AND/OR [h-CLAT] AND/OR [DPRA Cys] AND/OR [DPRA Lys]	3 Model Consensus	4	OECD Toolbox v4.4; Gerberik et al. 2014; Natsch et al. 2008; Natsch et al. 2015; Jaworska et al. 2015; Dimitrov et al. 2016
3	[GSH] AND/OR [h-CLAT] AND/OR [DPRA Cys] AND/OR [DPRA Lys]	2 Model Consensus	3	OECD Toolbox v4.4; Gerberik et al. 2014; Natsch et al. 2008; Natsch et al. 2015; Jaworska et al. 2015; Dimitrov et al. 2016
4	[GSH] OR [h-CLAT] OR [DPRA Cys] OR [DPRA Lys]	Single Model Alert	2	OECD Toolbox v4.4; Gerberik et al. 2014; Natsch et al. 2008; Natsch et al. 2015; Jaworska et al. 2015; Dimitrov et al. 2016
5	No alerts in any models (includes slightly reactive and suspect protein potency)	4 Model Consensus	5	OECD Toolbox v4.4; Gerberik et al. 2014; Natsch et al. 2008; Natsch et al. 2015; Jaworska et al. 2015; Dimitrov et al. 2016
5a	No alerts in 3 models, NA in 2 models		3	OECD Toolbox v4.4; Gerberik et al. 2014; Natsch et al. 2008; Natsch et al. 2015; Jaworska et al. 2015; Dimitrov et al. 2016
6	No prediction possible		NA	
	Rules for Protein Binding			
1	[OASIS Protein] AND [OECD Protein] AND [OASIS CA or TIMES CA] AND [OASIS Skin Sensitization]	4 Model Consensus	5	OECD Toolbox v4.4; TIMES v2.03; Mekenyan et al. 2007
2	[OASIS Protein] AND/OR [OECD Protein] AND/OR [OASIS CA or TIMES CA] AND/OR [OASIS Skin Sensitization]	3 Model Consensus	4	OECD Toolbox v4.4; Mekenyan et al. 2007
3	[OASIS Protein] AND/OR [OECD Protein] AND/OR [OASIS CA or TIMES CA]	2 Model Consensus	3	OECD Toolbox v4.4; Mekenyan et al. 2007

	AND/OR [OASIS Skin Sensitization]			
4	[OASIS Protein] OR [OECD Protein] OR [OASIS CA or TIMES CA] OR [OASIS Skin Sensitization]	Single Model Alert	2	OECD Toolbox v4.4; Mekenyan et al. 2007
5	No alerts in any models	4 Model Consensus	5	OECD Toolbox v4.4; Mekenyan et al. 2007
	Rules for DNA Binding			
1	[TIMES COMET] AND [OASIS DNA or TIMES TGR] AND [ISS *in vivo* micronucleus] AND [ISS *in vitro* AMES]	4 Model Consensus	5	OECD Toolbox v4.4; Benigni and Bossa 2012a,b; Mekeyan et al. 2004; Serafimova et al. 2007
2	[TIMES COMET] AND/OR [OASIS DNA or TIMES TGR] AND/OR [ISS *in vivo* micronucleus] AND/OR [ISS *in vitro* AMES]	3 Model Consensus	4	OECD Toolbox v4.4; Benigni and Bossa 2012a,b; TIMES v2.30; Laboratory of Mathematical Chemistry
3	[TIMES COMET] AND/OR [OASIS DNA or TIMES TGR] AND/OR [ISS *in vivo* micronucleus] AND/OR [ISS *in vitro* AMES]	2 Model Consensus	3	OECD Toolbox v4.4; Benigni and Bossa 2012a,b; TIMES v2.30; Laboratory of Mathematical Chemistry
4	[TIMES COMET] OR [OASIS DNA or TIMES TGR] OR [ISS *in vivo* micronucleus] OR [ISS *in vitro* AMES]	Single Model Alert	2	OECD Toolbox v4.4; Benigni and Bossa 2012a,b; TIMES v2.30; Laboratory of Mathematical Chemistry
5a	No alerts in 3 models, NA in 2	3 model consensus	3	OECD Toolbox v4.4; Benigni and Bossa 2012a,b; TIMES v2.30; Laboratory of Mathematical Chemistry
5	No alerts in any models	4 Model Consensus	5	OECD Toolbox v4.4; Benigni and Bossa 2012a,b; TIMES v2.30; Laboratory of Mathematical Chemistry
	Rule for Developmental and Reproductive Toxicity (DART)			
1	[DART] profiler has positive mechanism	Single Model Alert	2	OECD Toolbox v4.4

Where:

GSH = OECD QSAR Toolbox v4.2-4.4 mechanistic profiler for covalent binding with the thiol group of glutathione

h-CLAT = OECD QSAR Toolbox v4.2-4.4 mechanistic profiler for human cell line activation (h-CLAT) *in vitro* assay for skin sensitization

DPRA Cys/Lys = Direct Peptide Reactivity Assay (DPRA) in cysteine (Cys) and lysine (Lys) greater than or less than 13% peptide depletion

CA = Chromosomal aberration

OASIS = Laboratory of Mathematical Chemistry OASIS Modeling Suite

TIMES = TIssue MEtabolism Simulator suite of QSAR models from Laboratory of Mathematical Chemistry

COMET = an *in vitro* assay that measures DNA damage in individual eukaryotic cells (Olive and Banáth 2006)

TGR = Transgenic rodent gene mutation assay (Lambert et al. 2005)

ISS = Istituto Superiore di Sanità, Rome, Italy

DART = Developmental and reproductive toxicity

NA = Not available

2. *In chemico*

Hazard classification of *in chemico* data (protein binding) was based on the current set of rules from the OECD Toolbox v4.4.1 *in silico* profilers and can be consulted for each of the *in chemico* endpoints in Table 3 below. No potency rules were available for LCMS adduct formation data. Rules for DPRA were therefore applied to LCMS data.

Table 3: Rules for hazard classification and confidence scoring for *in chemico* reaction assay data

Rule No.	Rule Description	Hazard Class	Confidence Score	Reference
1	DPRA >21%	3	5	OECD Toolbox v4.4.1 RC50 database; DPRA profiler; Roberts et al. 2010; Schultz et al. 2005,2006a,2006b,2007,2009
1	DPRA 9-21%	2	5	OECD Toolbox v4.4.1 RC50 database; DPRA profiler; Roberts et al. 2010; Schultz et al. 2005,2006a,2006b,2007,2009
1	DPRA <9%	1	5	OECD Toolbox v4.4.1 RC50 database; DPRA profiler; Roberts et al. 2010; Schultz et al. 2005,2006a,2006b,2007,2009

2	GSH <=15 mmol/L	3	5	OECD Toolbox v4.4.1 RC50 database; GSH profiler; Roberts et al. 2010; Schultz et al. 2005,2006a,2006b 2007,2009
2	GSH 16-70 mmol/L	2	5	OECD Toolbox v4.4.1 RC50 database; GSH profiler; Roberts et al. 2010; Schultz et al. 2005,2006a,2006b,2007,2009
2	GSH >70 mmol/ or non active	1	5	OECD Toolbox v4.4.1 RC50 database; GSH profiler; Roberts et al. 2010; Schultz et al. 2005,2006a,2006b,2007,2009
3	LCMS>21%	3	5	OECD Toolbox v4.4.1 RC50 database; DPRA profiler; Roberts et al. 2010; Schultz et al. 2005,2006a,2006b,2007,2009
	LCMS 9-21%	2	5	OECD Toolbox v4.4.1 RC50 database; DPRA profiler; Roberts et al. 2010; Schultz et al. 2005,2006a,2006b,2007,2009
	LCMS <9%	1	5	OECD Toolbox v4.4.1 RC50 database; DPRA profiler; Roberts et al. 2010; Schultz et al. 2005,2006a,2006b,2007,2009

Where:

GSH = OECD QSAR Toolbox v4.2-4.4 mechanistic profiler for covalent binding with the thiol group of glutathione (relative concentration for 50% depletion: RC50)

DPRA Cys/Lys = Direct Peptide Reactivity Assay (DPRA) in Cysteine and Lysine (% depletion)

LCMS = liquid chromatography tandem mass spectrometry (% depletion)

3. *In vitro*

Table 4: Rules for hazard classification and confidence scoring for *in vitro* genotoxicity

Rule No.	Rule Description	Hazard Classification	Confidence Score	Endpoints
1	Positive result in any single *in vitro* test, no equivocal results	3	5	T143; Gene mutation I; Chromosome aberration II (Japan MHLW); Mutagenicity I (ECVAM); Chromosome aberration I (Oasis); Chromosome aberration V (ECVAM); Micronucleus I Chromosome aberration IV (EFSA) Gene mutation II

2	Negative result in all *in vitro* tests, no equivocal results	1	5	
3	Positive or negative *in vitro* results containing equivocal results	3 or 1	3	

4. *In vivo*

Identical rules for selection, classification of hazard and scoring of confidence for *in vivo* data cited in Appendix III for receptor-mediated toxicity are also used for chemical reactivity and genotoxicity, given that any of these interactions can lead to developmental and reproductive effects. However, acute aquatic toxicity data includes mortality data to account for effects due to chemical reactivity. Also, categorical data for genotoxicity includes the following endpoints that are absent in receptor-mediated categorical data:

- T143 (strain)
- Chromosome aberration V (ECVAM)
- Micronucleus I
- Micronucleus II
- Chromosome aberration IV (EFSA)
- Gene mutation II
- Gene mutation I
- Mutagenicity I (ECVAM)

Numerical and Categorical Mammalian Hazard Classification Rules		
Rule No.	**Rule Description**	**Hazard Classification**
1	Value is <=10 mg/kg bw **OR** is positive (pos) in any one *in vivo* test	3
2	Value is >10 to <= 500 mg/kg bw	2
3	Value is >500 mg/kg bw **OR** is negative (neg) or inactive in all tests	1
Numerical and Categorical Data Confidence Rules		
Rule No.	**Rule Description**	**Confidence Score**
1	5th percentile of numerical data **OR** “pos” or “neg” categorical data with no equivocal results	26
2	Single numerical value **OR** “pos” or “neg” categorical data with equivocal results	21

Note: when both numerical and categorical data are both available for a substance, numerical data were selected, preferentially

The above categorical endpoints are similar to those for *in vitro* genotoxicity data.

References

Accounting for data variability, a key factor in *in vivo*/*in vitro* relationships: application to the skin sensitization potency (*in vivo* LLNA versus *in vitro* DPRA) example. J Appl Toxicol. 36(12):1568-1578.

Enoch SJ, Cronin MTD, Schultz TW, Madden JC. 2008. Quantitative and mechanistic read-across for predicting skin sensitization potential of alkenes acting via Michael addition. Chem Res Toxicol. 21:513-520.

Gerberick GF, Vassallo JD, Bailey RE, Chaney JG, Morrall SW, Lepoittevin JP. 2004. Development of a peptide reactivity assay for screening contact allergens. Toxicol Sci. 81:332-343.

Jaworska J, Natsch A, Ryan C, Strickland J, Ashikaga T, Miyazawa M. 2015. Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: a decision support system for quantitative weight of evidence and adaptive testing strategy. Arch Toxicol. 89(12):2355-2383.

Lambert IB, Singer TM, Boucher SE, Douglas GR. 2005. Detailed review of transgenic rodent mutation assays. Mutat Res. 590(1-3):1-280.

Mekenyan O, Dimitrov S, Serafimova R, Thompson E, Kotov S, Dimitrova N, Walker J. 2004. Identification of the structural requirements for mutagenicity by incorporating molecular flexibility and metabolic activation of chemicals I: TA100. Chem Res Toxicol. 17(6):753-766.

Mekenyan O, Todorov M, Serafimova R, Stoeva S, Aptula A, Finking R, Jacob E. 2007. Identifying the structural requirements for chromosomal aberration by incorporating molecular flexibility and metabolic activation of chemicals. Chem Res Toxicol. 20(12):1927-1941.

Micronucleus ISSMIC [database of *in vivo* micronucleus mutagenicity assay results]. 2011. Ver. v4b. Rome (IT): Istituto Superiore di Sanità (ISS); Liebefeld (CH): Federal Office of Public Health. [Benigni R, Battistelli CL, Bossa C, Tcheremenskaia O, Crettaz P. 2013. New perspectives in toxicological information management, and the role of ISSTOX databases in assessing chemical mutagenicity and carcinogenicity. Mutagenesis. 28(4):401-409].

Natsch A, Gfeller H. 2008. LC-MS-based characterization of the peptide reactivity of chemicals to improve the *in vitro* prediction of the skin sensitisation potential. Toxicol Sci. 106:464-478.

Natsch A, Emter R, Gfeller H, Haupt T, Ellis G. 2015. Predicting skin sensitizer potency based on *in vitro* data from KeratinoSens and kinetic peptide binding: global versus domain-based assessment. Toxicol Sci. 143(2):319-332.

Olive P, Banáth J. 2006. The comet assay: a method to measure DNA damage in individual cells. Nat Protoc. 1:23–29.

Roberts DW, Schultz TW, Wolf EM, Aptula AO. 2010. Experimental reactivity parameters for toxicity modeling: Application to aquatic toxicity of Sn2 electrophiles to Tetrahymena pyriformis. Chem Res Toxicol. 23:228-234.

Schultz TW, Carlson RE, Cronin MTD, Hermens JLM, Johnson R, O'Brien PJ, Roberts DW, Siraki A, Wallace KD, Veith GD. 2006. A conceptual framework for predicting toxicity of reactive chemicals: Models for soft electrophilicity. SAR QSAR Environ Res. 17:413-428.

Schultz TW, Rogers K, Aptula AO. 2009. Read-across to rank skin sensitization potential: Subcategories for the Michael acceptor domain. Contact Dermatitis. 60:21-31.

Schultz TW, Yarbrough JW, Johnson EL. 2005. Structure-activity relationships for glutathione reactivity of carbonyl-containing chemicals. SAR QSAR Environ Res. 16:313-322.

Schultz TW, Yarbrough JW, Koss SK. 2006. Identification of reactive toxicants: Structure-activity relationships for amides. Cell Biol Toxicol. 22:339-349.

Schultz TW, Yarbrough JW, Hunter RS, Aptula AO. 2007. Verification of the structural alerts for Michael acceptors. Chem Res Toxicol. 20:1359-1363.

Serafimova R, Todorov M, Pavlov T, Kotov S, Jacob E, Aptula A, Mekenyan O. 2007. Identification of the structural requirements for mutagencitiy, by incorporating molecular flexibility and metabolic activation of chemicals. II. General Ames mutagenicity model. Chem Res Toxicol. 20:662-676.

[TIMES] TIssue MEtabolism Simulator [prediction module]. 2018. Ver. 2.27.19. Bourgas (BG): University "Prof. Dr. Assen Zlatarov", Laboratory of Mathematical Chemistry.

Urbisch D, Mehling A, Guth K, Ramirez T, Honarvar N, Kolle S, Landsiedel R, Jaworska J, Kern P, Gerberick F, et al. 2015. Assessing skin sensitization hazard in mice and men using non-animal test methods. Regulat Toxicol Pharmacol. 71:337-351.

Yarbrough JW, Schultz TW. 2007. Abiotic sulfhydryl reactivity: A predictor of aquatic toxicity for carbonyl-containing α,β-unsaturated chemicals. Chem Res Toxicol. 20:558-562.

14. APPENDIX V: Rules for Classification and Confidence Scoring of Mode of Toxic Action (MoA)

Mode of toxic action (MoA) seeks the consensus between *in silico* profiling and tissue residue-based approaches to determine narcotic or "non-specific" interactions (Hazard Class 1) from those that have specific interactions and are thus more potent (Hazard Class 3). Greater confidence is given when both approaches agree. Modes of action are first based on known observed MoA (from Baron et al. 2015); however, these are only available for a few substances in ERC2. Therefore, MoAs was determined using *in silico* results when no empirical MoA was available. The rules and approaches for *in silico*, tissue residue and final consensus MoA are discussed below.

1. *In silico*

In silico profiling of mode of action involved a consensus approach using six MoA QSARs

- iSafeRat MechoA Profiler (Bauer et al. 2018)
- ASTER (Russom et al. 1997; ASTER 1999)
- TEST Acute Fish MoA Assignment (Martin et al. 2013; Martin et al. 2015)
- OASIS Acute Fish MoA Assignment (Dimitrov et al. 2003)
- Verharr (Verharr et al. 1992; Verharr et al. 2000; Enoch et al. 2008)
- OECD Toolbox Profiler for Uncouplers of Oxidative Phosphorylation (S. Enoch and Laboratory of Mathematical Chemistry)

Table 1: Hazard classification and confidence rules for *in silico* profiling of MoA

Rule No.	Rule Description	Hazard Class	Confidence Score
1	>=5 models agree	1 or 3	5
2	4 models agree	1 or 3	4
3	3 models agree	1 or 3	3
4	2 models agree	1 or 3	2
5	Single model	1 or 3	1
6	1-3 models agree for both classes (equivocal)	3	1

The MoA for alpha, beta-unsaturated alcohols, esters, diesters, phenols and anilines (if not an uncoupler in ASTER or TEST), quinolines, chlorodiester, and amines was set to "narcosis" in all models except iSafeRat due to metabolite considerations. Evidence suggests that potency of these classes does not exceed baseline narcosis when examined using tissue residue approaches (Armitage and Bonnell 2017). Phenols and anilines determined by OASIS were parsed out for uncouplers using ASTER, TEST or OECD

approaches. When MoA was determined to be equivocal (same number of models suggest specific or non-specific MoA), then a specific MoA was selected as a precautionary approach. Although there is no impact to hazard classification or confidence score from the selection of type of specific MoA, the following order of QSAR approaches for selection of specific MoA type was used: iSafeRat; ASTER; TEST; OASIS; Verharr.

The following specific MoAs, in no particular order, were identified in ERC2 according to the QSAR approaches above (note there are overlapping MoAs due to slightly different model output descriptions):

Carbonyl reactivity (aldehyde eq. #3); Narcosis; Reactivity; Uncoupler of oxidative phosphorylation; Carbonyl reactivity (aldehyde eq. #3); Receptor docking; Enzyme disruption; Alkylation/arylation-based reactivity; Hydrazine-based reactivity; OP-mediated AChE inhibition; Pyridinium chemicals; Carbonyl reactivity (aldehyde eq. # 2); Quaternary ammonium chemicals; Sulfhydryl based reactivity; Isocyanate based reactivity; Neurotoxicant: Strychnine; Carbonyl reactivity (aldehyde eq. # 1); Neurotoxicant: Caffeine; Reactive unspecified; Neurotoxicant: Cyclodiene-type; Reactive dinitroaromatic group; Acylation based reactivity; Neurotoxicity; Acrylate toxicity; AChE inhibition; Beta-Halogenated alcohol-based inhibition; Uncoupler; Nitroso-based reactivity; Carbonyl-based reactivity; Reactive diketones; Reactive difunctional acrylate; N-Halogenated acetophenone inhibition; Neurotoxicant: Pyrethroid; Malonitrate-based reactivity; Oxime-base reactivity; Carbamate-mediated AChE inhibition; Carbonyl reactivity (aldehyde eq. #1); Alpha, Beta-unsaturated alcohols; Acetamidophenol reactivity; Phenols and Anilines; Nitrile-based reactivity; Carbonyl reactivity (aldehyde eq. #2)

2. Tissue residue approaches

The calculation of tissue residues (TR) was performed for neutral organic substances only due to limitations of TR approaches when applied to ionizing chemicals. Three TR approaches were examined; however, two of the three approaches (CBR50 and CMC50) used three different calculation routines (described below) based on different partition coefficients. In total, seven approaches were investigated to determine if consensus exists among the different methods. All of the TR approaches have a basis in determining median lethality (LD50) from a critical tissue concentration either on a whole body basis or at the cell membrane. Approaches using critical body residues for median lethality (CBR50) and median lethal activity (LA50) are whole body based while the critical membrane concertation approach (CMC50) is parametrized at the cell membrane and thus closer to or at the site of toxic action. The TR approaches require the input of median lethal toxicity data. Data for fish species only was used in ERC2 because the TR methods have been proven largely using fish toxicity data and MoA QSARs have largely used fish toxicity data as well. The TR approaches used in ERC2 are described in detail in Armitage (2018) and Armitage et al. (2018) and are summarized below.

Average empirical fish toxicity data were preferentially used for TR calculation (n=~1600). Confidence scoring for empirical data was set higher than TR using predicted fish toxicity. Rules for the selection of acute median lethal fish toxicity (LC50) are summarized in Table 4 below. Empirical fish toxicity data were extracted from the OECD QSAR Toolbox v4.3/4.4. If an empirical value was not available, a hybrid best model/model consensus approach was used based on a seven acute fish toxicity QSAR model

comparison conducted by ECCC using all substances in ERC2. The analysis confirmed that the best performing models (based on residual mean square error and correlation analysis) were ECCC's AIEPS and US EPA's TEST QSAR models which performed equally well (see Appendix I for model description). These results confirm a previous analysis by ECCC and colleagues (Moore et al. 2008), which showed that using a small halogenated organic data set, the best performing model was a probabilistic neural network (PNN) model. The PNN model has the same basis as ECCC's AIEPS model.

CBR50

The Critical Body Residue (CBR) approach has been evident in the scientific literature for several years (McCarty et al. 1992; McCarty and Mackay 1993; Barron et al. 1997; Barron et al. 2002; Meador et al. 2011; Barron et al. 2015). The CBR approach is useful for identifying baseline toxicants expected to elicit toxicity at a relatively narrow range of concentrations in the tissue of organisms (e.g., 1–10 mmol/kg for hydrophobic neutral organics) even though the external lethal concentration (e.g., LC50 in aquatic toxicity test) or dose may vary over orders of magnitude. Chemicals exerting toxicity through specific modes of action (MoA) have been observed to have CBRs well below this threshold, thereby allowing its use to discriminate between baseline toxicants (narcosis) and non-baseline toxicants (Armitage 2018).

McCarty et al. (1992) proposed the following equation to estimate the CBR corresponding to baseline toxicity for neutral organic chemicals with log K_{ow} values between -1.5 and 6.0:

$$\mathrm{CBR50}=2.5\ \mathrm{mmol/kg} + 50/K_{\mathrm{ow}}$$

Across this range of log K_{ow}, the CBR for baseline toxicity ranges from 1580 mmol/kg to 2.5 mmol/kg respectively. Note that once the log K_{ow} of the chemical is greater than 1, the estimated CBR falls into the commonly applied "rule of thumb' range for CBRs (i.e., 1–10 mmol/kg). The advantage of applying the equation presented above is that estimated CBRs for hydrophilic chemicals are "corrected" for the fact that lipids are no longer the dominant sorption phase within the organism.

Empirical or QSAR-based LC50s for neutral organic chemicals can also be converted to CBRs (mmol/kg) using the following expression:

$$\mathrm{CBR50} = \mathrm{LC50\ x\ K_{BW}}$$

where LC50 (median lethal concentration) is in units of mmol/L and K_{BW}, the equilibrium biota-water partition coefficient, is in units of L/kg. Biota-water partitioning for neutral organic chemicals was estimated using two methods:

$$\mathrm{K_{BW}} = 0.05\mathrm{K_{ow}} + 0.8$$

$$\mathrm{K_{BW}} = 0.0375\mathrm{K_{SW}} + 0.0125\mathrm{K_{MW}} + 0.8$$

where 0.05 = lipid content of the organism; 0.8 is the water content of the organism; 0.0375 is the assumed storage lipid content of the organism; K_{SW} is the storage lipid-water partition coefficient; 0.0125 is the assumed membrane lipid content of the organism; and K_{MW} is the membrane-water partition coefficient. K_{SW} values were estimated using pp-LFERs (polyparameter linear free energy

relationship) reported in the UFZ LSER database available from http://www.ufz.de/lserd. K_{MW}s in this expression are also based on pp-LFERs. The inclusion of the water content of the organisms in the equation above is critical when calculating the CBRs of hydrophilic chemicals (log K_{ow} < 2) because it cannot be assumed that lipids are the main storage phase within the organism. The CBR50 range of 1-10 mmol/kg for baseline toxicity implicitly assumes that the majority fraction of chemical load is sorbed to lipid phases (i.e., membranes). For this reason, the baseline toxicity CBR range of 1-10 mmol/kg (median of ~3 mmol/kg) is valid only for chemicals with log K_{ow} greater than one. Therefore, domain of application rules were developed for the TR rules in the tables below.

CMC50

As previously mentioned the CMC50 is more focused at the site of toxic action for narcosis (i.e., the membranes) (Escher et al., 2002). Similar to the CBR50 approach, the CMC50 is associated with baseline toxicants eliciting at a similar membrane concentration, proposed to centre around 100 mmol/kg lipid (Endo et al 2011; Goss and Endo, 2016; Goss et al 2018).

CMC50 (mmol/kg lipid) can be calculated for neutral organic chemicals *and ionizing chemicals* from LC50s and the membrane-water partition coefficient using the following expression:

$$\mathrm{CMC50} = \mathrm{K_{MW}\ x\ LC50}$$

where LC50 is in units of mmol/L and K_{MW} is in units of L/kg lipid and can be obtained using various methods such as pp-LFERs (e.g., via UFZ LSER website), extrapolated from K_{ow}, measured empirically (Droge et al. 2017) or calculated using other quantum chemistry *in silico* tools such as COSMOmic (Klamt et al. 2008). The CMC approach is regarded as being more broadly applicable (i.e., also to chemicals with log K_{ow} < 2) because the calculation explicitly deals with the site of action for baseline toxicity (i.e., membranes). However, there are uncertainties with regard to its application to superhydrophilics (i.e., not well studied). Therefore, similar K_{ow} domain of application rules for CB50 were applied to CMC50.

LA50

The merits and potential limitations of the chemical activity concept, here called lethal activity (LA50), in the context of ecotoxicity is also well described in the scientific literature (e.g., Mayer and Reichenberg 2006; Mackay et al. 2009; Mackay et al. 2011; Thomas et al. 2015; Goss and Endo 2016; Thomas et al. 2016). The underlying concept of LA50 is that baseline toxicants can be expected to elicit toxicity at similar chemical activities (0.01–0.1), even if the LC50 varies over orders of magnitude. LA50s (unitless) can be calculated from LC50s for neutral organic chemicals using the following expression:

$$\mathrm{LA50} = \mathrm{LC50}/\mathrm{S_W}$$

where S_W is the water solubility (mg/L) of the chemical of a liquid or the sub-cooled liquid water solubility if substance is a solid (mg/L). Sub-cooled liquid water solubility of solids can be calculated

using the fugacity ratio (F) (Mackay 2001; Mackay et al. 2011). The LA50 approach assumes that chemical equilibrium is achieved between all phases (i.e., external, internal) and therefore the LA50 estimated from the LC50 is equal to the whole body chemical activity including cell membranes. If non-equilibrium conditions apply (e.g., rapid metabolism), the external LA50 may not be representative of the internal LA50. The lethal activity concept is not applicable to fully miscible substances (i.e., LA50 becomes unbounded) and is far less well-developed for ionizing chemicals as there is no consensus in the literature on how to apply it as well as issues with measurement of S_W.

Table 2: Hazard Classification and Confidence Rules for Empirical and Predicted Tissue Residue Approaches

RULES FOR TISSUE RESIDUE APPROACHES USING PREDICTED TOXICITY DATA			
Rule No.	**Rule Description**	**Hazard Classification**	**Confidence Score**
1	7 approach consensus	1 or 3	5
2	6 approach consensus	1 or 3	4
3	If fewer than 6 approach consensus, then MoA based on critical membrane concentration approach	1 or 3	3
RULES FOR TISSUE RESIDUE APPROACHES USING EMPIRICAL TOXICITY DATA			
Rule No.	**Description**	**Hazard Classification**	**Confidence Score**
4	7 approach consensus	1 or 3	10
5	6 approach consensus	1 or 3	9
6	If fewer than 6 approach consensus, then MoA based on critical membrane concentration approach	1 or 3	8

Note: Out of domain for empirical and predicted tissue residue approaches = log K_{ow} < 2, Log K_{ow} > 6, water solubility < LC50

Table 3: Hazard Classification and Confidence Rules for Final MoA Classification

RULES FOR FINAL MOA CLASSIFICATION				
Rule No.	**Rule Description**	**Hazard Classification**	**Confidence Score**	**Approach Used to Select MoA**
1	Empirical MoA available	1 or 3	15	MoATox
2	QSAR and CBR MoA agree	1 or 3	1 to 15	QSAR
3	QSAR and CBR MoA do not agree; MOA=CBR empirical	1 or 3	8 to 10	CBR
4	QSAR and CBR MoA do not agree; MOA=QSAR High Confidence	1 or 3	4 to 5	QSAR

5	QSAR and CBR MoA do not agree; MOA=CBR Predicted High Confidence	1 or 3	4 to 5	CBR
6	CBR=NA; MoA=QSAR lower confidence	1 or 3	1 to 3	QSAR
7	if CBR MoA out of domain; MOA=QSAR	1 or 3	1 to 5	QSAR

Table 4: Rules for Selecting Acute Median Lethal Fish Toxicity (LC50)

Rule No.	Rule Description
1	Log K_{ow} must be <9; if >9, then no fish toxicity data accepted (out of domain of acute toxicity testing and QSARs)
2	Empirical fish toxicity data selected preferentially to QSAR even if greater than estimated water solubility (WS), but cannot exceed rule of >9
3	Model predictions used for organics salts of metals
4	If LC50 is greater than predicted or measured WS, then WS used as LC50
5	Predicted water solubility calculated using the mean of the AIEPS and TEST QSAR models (best performing models in five WS model comparison performed by ECCC) unless no WS prediction possible in either or both models, then single model WS used, or N/A
6	If MoA was determined as narcosis (non-specific), then fish LC50 was generated using AIEPS, unless no prediction was possible, then TEST fish LC50 used
7	If MoA was determined to be specific, then then TEST MoA-based fish LC50 used, unless no prediction was possible then, AIEPS fish LC50 used
8	If no prediction possible in AIEPS and TEST QSAR models, then fish LC50 value set to average of all other models used in ECCC seven fish LC50 QSAR model comparison analysis (ECOSAR, OASIS, ASTER, ACD Labs, TOPKAT) if still np = no prediction possible
9	If no empirical data of QSAR prediction possible, result set to N/A

References

[ASTER] Assessment Tools for the Evaluation of Risk. 1999. Duluth (MN): US Environmental Protection Agency, Mid-Continent Ecology Division. [restricted access].

Armitage J, Arnot JA, Bonnell M. 2018. Comparing mode of action (MOA) classification using body residues, membrane concentrations and chemical activity for chemical prioritization [Abstract]. Society of Environmental Toxicology and Chemistry (SETAC) North America 39th Annual Meeting; 2018 Nuvember 4-8; Sacramento, California.

Armitage J. 2018. Generation of tissue residue, critical membrane concentration and toxicity ratios for organic substances on the Canadian DSL. Gatineau (QC): Environment and Climate Change Canada.

Baron MG, Lilavois CR, Martin TM. 2015. MOATox: A comprehensive mode of action and acute toxicity database for predictive model development. Aquat Toxicol. 161:102-107.

Barron MG, Anderson MJ, Lipton J, Dixon DG. 1997. Evaluation of critical body residue QSARs for predicting organic chemical toxicity to aquatic organisms. SAR QSAR Environ Res. 6:47-62.

Barron MG, Hansen JA, Lipton J. 2002. Association Between contaminant tissue residues and effects in aquatic organisms. Rev Environ Contam Toxicol. 173:1-37.

Barron MG, Lilavois CR, Martin TM. 2015. MOAtox: A comprehensive mode of action and acute aquatic toxicity database for predictive model development. Aquat Toxicol. 161:102-107.

Bauer F, Thomas PC, Fouchard SY, Neunlist SJM. 2018. A new classification algorithm based on mechanisms of action. Computat Toxicol. 5:8-15.

Dimitrov SD, Mekenyan OG, Sinks GD, Schultz TW. 2003. Global modeling of narcotic chemicals: ciliate and fish toxicity. J Mol Structr. 622(1-2):62-70.

Droge ST, Hermens JL, Gutsell S, Rabone J, Hodges G. 2017. Predicting the phospholipophilicity of monoprotic positively charged amines. Environ Sci Process Impacts. 19:307-323.

Endo S, Escher BI, Goss KU. 2011. Capacities of membrane lipids to accumulate neutral organic chemicals. Environ Sci Tech. 45:5912-5921.

Enoch S J, Hewitt M, Cronin MTD, Azam S, Madden JC. 2008. Classification of chemicals according to mechanism of aquatic toxicity: an evaluation of the implementation of the Verhaar scheme in Toxtree. Chemosphere. 73(3):243−248.

Escher BI, Eggen RIL, Schreiber U, Schreiber Z, Vye E, Wisner B, Schwarzenbach RP. 2002. Baseline toxicity (narcosis) of organic chemicals determined by *in vitro* membrane potential measurements in energy-transducing membranes. Environ Sci Tech. 36:1971-1979.

Goss KU, Bittermann K, Henneberger L, Linden L. 2018. Equilibrium biopartitioning of organic anions - a Case study for humans and fish. Chemosphere. 199:174-181.

Goss KU, Endo S. 2016. Comment on application of the activity framework for assessing aquatic ecotoxicology data for organic chemicals. Environ Sci Tech. 50:4139-4140.

Klamt A, Huniar U, Spycher S, Keldenich J. 2008. COSMOmic: A mechanistic approach to the calculation of membrane−water partition coefficients and internal distributions within membranes and micelles. J Phys Chem. 112(38):12148-12157.

Mackay D, Arnot JA, Petkova EP, Wallace KB, Call DJ, Brooke LT, Veith GD. 2009. The physicochemical basis of QSARs for baseline toxicity. SAR QSAR Environ Res. 20:393-414.

Mackay D, Arnot JA, Wania F, Bailey RE. 2011. Chemical activity as an integrating concept in environmental assessment and management of contaminants. Integr Environ Assess Manage. 7:248-255.

Maeder V, Escher BI, Scheringer M, Hungerbühler K. 2004. Toxic ratio as an indicator of the intrinsic toxicity in the assessment of persistent, bioaccumulative, and toxic chemicals. Environ Sci Technol. 38(13):3659-3666.

Martin TM, Grulke CM, Young DM, Russom CL, Wang NY, Jackson CR, Barron MG. 2013. Prediction of aquatic toxicity mode of action using linear discriminant and random forest models. J Chem Inf Model. 53(9):2229−2239.

Martin TM, Young DM, Lilavois CR, Barron MG. 2015. Comparison of global and mode of action-based models for aquatic toxicity. SAR QSAR Environ Res. 26(3):245−262.

Mayer P, Reichenberg F. 2006. Can highly hydrophobic organic substances cause aquatic baseline toxicity and can they contribute to mixture toxicity? Environ Toxicol Chem. 25:2639-2644.

McCarty LS, Arnot JA, Mackay D. 2013. Evaluation of critical body residue for acute narcosis in aquatic organisms. Environ Toxicol Chem. 32(10):2301-2314.

McCarty LS, Mackay D. 1993. Enhancing ecotoxicological modeling and assessment: critical body residues and modes of toxic action. Environ Sci Technol. 27(9):1719-1728.

Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE, Drummond RA. 1997. Predicting modes of toxic action from chemical structure: acute toxicity in the fathead minnow (Pimephales promelas). Environ Toxicol Chem. 16(5):948−967.

Thomas P, Dawick J, Lampi M, Lemaire P, Presow S, van Egmond R, Arnot JA, Mackay D, Mayer P, Burgos MG. 2015. Application of the activity framework for assessing aquatic ecotoxicology data for organic chemicals. Environ Sci Technol. 49:12289-12296.

Thomas P, Mackay D, Mayer P, Arnot J, Burgos MG. 2016. Response to comment on “Application of the activity framework for assessing aquatic ecotoxicology data for organic chemicals”. Environ Sci Tech. 50:4141-4142.

Verhaar HJM, van Leeuwen CJ, Hermens JLM. 1992. Classifying environmental pollutants: Part 1. Structure-activity relationships for prediction of aquatic toxicity. Chemosphere. 25:471-491.

Verhaar HJM, Solbe J, Speksnijder J, van Leeuwen CJ and Hermens JLM. 2000. Classifying environmental pollutants: Part 3. External validation of the classification system. Chemosphere. 40:875-883.

15. APPENDIX VI: Rules for Classification and Confidence Scoring of Food Web Toxicity

The RAIDAR model hazard assessment factor (HAF) can be equated to a risk quotient based on internal toxicity thresholds and using a default emission rate to the aquatic environment. The HAF classification rules in Table 1 represent the relative hazard among ERC2 substances based on their intrinsic ability to bioaccumulate in the model food web to levels associated with deleterious acute effects using a default rate of emission (1 kg/hr). Substances with HAF values closer to or exceeding one suggest that there is a lower margin of exposure using a default rate of emission and therefore were classified as a higher concern. The RAIDAR risk assessment factor or RAF (Arnot et al. 2006; Arnot et al. 2008), which replaces the unit rate of emission (C_U of 1 kg/hr) with the actual rate of emission (E_A of 1 kg/hr) to estimate food web toxicity (i.e., the HAF multiplied by the ratio of the actual vs. default rate of emission) was not used in ERC2. This was due to the lack of reliable estimates of tonnage for almost all ERC2 substances. Confidence scoring of the HAF values accounted for fugacity and extreme property warnings developed for version 3.0 of the model. These warnings suggest that physical-chemical property data used to parameterize the model may not be of good quality (extreme low and high values) and may result is mass-balance concentrations exceeding maximum solubility (fugacity warnings) in various media. A lower confidence score was given when these warnings were triggered.

Table 1: Hazard Classification Rules for RAIDAR HAF

Rule No.	Rule Description	Approach	Hazard Class
1	HAF is >=0.01	Single Model Mass-Balance	3
2	HAF is <0.01 to 0.001	Single Model Mass-Balance	2
3	HAF is <0.001	Single Model Mass-Balance	1

Table 2: Confidence scoring used for RAIDAR HAF values

Rule No.	Rule Description	Confidence Score
1	No fugacity or extreme property warnings	10
2	Fugacity or extreme property warnings	5

References

Arnot JA, Mackay D. 2008. Policies for chemical hazard and risk priority setting: Can persistence, bioaccumulation, toxicity and quantity information be combined? Environ Sci Technol. 42(13):4648-4654.

Arnot JA, Mackay D, Webster E, Southwood J. 2006. Screening level risk assessment model for chemical fate and effects in the environment. Environ Sci Technol. 40:2316-2323.

16. APPENDIX VII: Rules for Classification and Confidence Scoring of Cumulative Toxicity

The MoA for determining IECs and TR_{50} estimates (i.e., CMC50) was based on the final outcomes of the consensus MoA in the MoA descriptor (Appendix V). The tissue residue values used for internal effect concentrations (IEC, mmol/kg ww) were extracted from IEC distribution data in Escher et al. (2011), which represent a variety of IEC data sources (e.g., McCarty et al. 1993; Baron et al. 2002; Traas et al. 2004; Hendriks et al. 2005) compiled by the authors as well as their own data contributions. IEC values (mol/kg ww) from the distribution shown in Figure 5 of Escher et al. (2011) converted to mmol/kg. Table 1 summaries the IEC values for the majority of the data used for IEC representing median lethality in fish, however, data for daphnids and algae are also included. Table 1 IECs were selected from minimum values in Escher et al. (2011), except for narcosis. A median value for narcosis (mmol/kg) is based on the default IEC for narcosis used in the RAIDAR model and represents the median of the range of IECs given in McCarty et al. (1993). The reason for this was to easily distinguish the IEC for specific MoAs vs. non-specific MoAs, which can overlap at extremes (Escher et al. 2011; Kienzler et al. 2019). The IEC for receptor docking MoA in Table 1 includes a measurement uncertainty factor of 100 because only a single value is available in Escher et al. (2011) for estrogenic activity (1 mmol/kg). Further, if ERC2 rules determined that a substance's final receptor-mediated hazard classification was class 2 or class 3, but was considered to be a narcotic by MoA consensus, then MoA for the purposes of an IEC was set to receptor docking.

IECs for quaternary ammonium chemicals (surfactant QACs) were not available in Escher et al. (2011). For the purposes of IEC derivation, QACs were treated as "super polar narcotics" based on the activity of the permanently charged nitrogen molecule and its high binding affinity for the largely negatively charged cell membrane (evidenced by high K_{MW} values). For QACs, IEC potency can be derived using the toxicity ratio approach (Maeder et al. 2004; Escher et al. 2011) described in Appendix V. The CMC50 for narcotics (100 mmol/kg lipid) was compared to the CMC50 calculated for linear alkyl QACs ranging from C12-C18 according to the equation describing CMC50 in Appendix V. An average K_{MW} was selected from Droge (2017a) for seven linear alkyl QACs (including benzyl alkyl) estimated using the COSMOmic model referenced in Appendix V. K_{MW} derived using COSMOmic are strongly correlated with empirical values for cationic surfactants (Timmer and Droge 2017; Droge 2017b). The median of acute EC50 and LC50 data collected by ECCC for several linear alkyl QACs ranging in chain length from C12-C18 was used for the external toxicity value (~0.8 mg/L). The effects data include measured and nominal concentrations (including water accommodated fraction tests) and were not rigorously screened for data quality issues given the difficulty of interpreting toxicity outcomes for these "difficult-to-test" substances. A deeper data quality analysis would be conducted at the risk assessment stage.

The toxicity ratio analysis for QACs revealed that, when expressed as an IEC (mmol/kg ww), their potency is approximately three orders of magnitude greater than that of non-polar and polar narcotics. The IEC for QAC is likely to best represent dermal and gill tissue "sensitization" effects rather than

internal effects as little distribution within organisms such as fish is expected (McLaughlin et al. 2019). Unknown MoA and not available values for MoA according to Appendix V resulted in an NA IEC.

Table 1: Selected Internal Effect Concentrations (mmol/kg ww) based on data from Escher et al. (2011)

Rule No.	Mode of Toxic Action	IEC (mmol/kg ww)
1	All reactive MoAs	0.01
2	Uncoupling and blocking of oxidative phosphorylation	0.01
3	C12-C18 alkyl quaternary ammonium chemicals	0.003
4	Narcosis (nonpolar+polar)	3
5	Receptor docking and ERC2 receptor-mediated	0.01
6	Enzyme disruption	0.01
7	Neurotoxicity (chloride transport inhibition+cation transport inhibition+acetylcholine esterase inhibition)	0.00001

References

Barron MG, Hansen J, Lipton J. 2002. Association between contaminant tissue residues and effects in aquatic animals. Rev Environ Contam Toxicol. 173:1–37.

Droge, STJ (Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam (NL). 2017a. Derivation and evaluation of partitioning properties for the bioaccumulation assessment of CMP priority ionogenic substances. Gatineau (QC): Environment and Climate Change Canada.

Timmer N, Droge ST. 2017. Sorption of cationic surfactants to artificial cell membranes: Comparing phospholipid bilayers with monolayer coatings and molecular simulations. Environ Sci Technol. 51(5):2890-2898.

Droge STJ, Hermens JLM, Gutsell S, Rabone J, Hodges G. 2017. Predicting the phospholipophilicity of monoprotic positively charged amines. Environ Sci Processes Impacts. 19:307-323.

Escher BI, Ashauer R, Dyer S, Hermens JLM, Lee J-H, Leslie HA, Mayer P, Meador JP, Warne MSJ. 2011. Crucial role of mechanisms and modes of toxic action for understanding tissue residue toxicity and internal effect concentrations of organic chemicals. Integr Environ Assess Manag. 7(1):28-49.

Hendriks AJ, Traas TP, Huijbregts MAJ. 2005. Critical body residues linked to octanol–water partitioning, organism composition, and LC50 QSARs: Meta-analysis and model. Environ Sci Technol. 39:3226–3236.

McCarty LS, Mackay D. 1993. Enhancing ecotoxicological modeling and assessment: Body residues and modes of toxic action. Environ Sci Technol. 27:1719–1728.

McLachlan M, Kierkgaard A, Strandell M, Yuan B, Chen C, Armitage J, Arnot J, Droge S. 2019. Bioconcentration of cationic surfactants in rainbow trout [presentation]. Society of Environmental Toxicology and Chemistry (SETAC) North America 40th Annual Meeting; 2019 November 3-7; Toronto, Ontario.

Maeder V, Escher BI, Scheringer M, Hungerbühler K. 2004. Toxic ratio as an indicator of the intrinsic toxicity in the assessment of persistent, bioaccumulative, and toxic chemicals. Environ Sci Technol. 38(13):3659-3666.

Traas TP, van Wezel AP, Hermens JLM, Zorn M, van Hattum AGM, Van Leeuwen CJ. 2004. Environmental quality criteria for organic chemicals predicted from internal effect concentrations and a food web model. Environ Toxicol Chem. 23:2518–2527.

Barron MG, Lilavois CR, Martin TM. 2015. MOAtox: A comprehensive mode of action and acute aquatic toxicity database for predictive model development. Aquat Toxicol. 161:102-107.

Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE, Drummond RA. 1997. Predicting modes of toxic action from chemical structure: acute toxicity in the fathead minnow (Pimephales promelas). Environ Toxicol Chem. 16(5):948−967.

17. APPENDIX VIII: Rules for Classification and Confidence Scoring of Response Time

Table 1 summarizes the exposure classification and confidence scoring rules for response time. Thresholds used for exposure classification are based on the P_{ov} values for benchmark substances used in the OECD POPs Tool v2.2[27] (Wegmann et al. 2008) as defined by Klasmeier et al. (2006) for the POPs Tool.

Table 1: Exposure classification and confidence scoring rules for response time

Classification Rules		
Rule No.	**Rule Description**	**Exposure Class**
1	97% Response Time is >=10 years (i.e., PCB-like) (>=3650 days)	3
2	97% Response Time is 1 to <10 years (e.g., a-HCH) (365-3649 days)	2
3	97% Response Time is <1 year (e.g., cresol, biphenyl) (<365 days)	1
Confidence Rules		
Rule No.	**Rule Description**	**Confidence Score**
1	P_{ov} has no fugacity or property warnings	10
2	P_{ov} has fugacity and/or property warnings	5

References

Klasmeier J, Matthies M, MacLeod M, Fenner K, Scheringer M, Stroebe M, Le Gall AC, McKone TE, van de Meent D, Wania F. 2006. Application of multimedia models for screening assessment of long-range transport potential and overall persistence. Environ Sci Tech. 40:53-60.

Wegmann F, Cavin L, MacLeod M, Scheringer M, Hungerbühler K. 2008. The OECD software tool for screening chemicals for persistence and long-range transport potential. Environ Model Software. 24(2):228-237.

[27] Available from the OECD at http://www.oecd.org/chemicalsafety/risk-assessment/oecdpovandlrtpscreeningtool.htm

18. APPENDIX IX: Rules for Classification and Confidence Scoring of Chemical Mobility

Table 1 summarizes the exposure classification and confidence scoring rules for chemical mobility. Thresholds used for exposure classification are based on the characteristic travel distance (CTD) values for benchmark substances from Beyer et al. (2000) and Zarfl et al. (2011).

Table 1: Classification and confidence rules for chemical mobility

Classification Rules		
Rule No.	**Rule Description**	**Exposure Class**
	CTD is >2000 km (continental to global scale)	3
	CTD is 700-2000 km (regional to continental scale)	2
	CTD is <700 km (local to regional scale)	1
	CTD is <5 km (local scale)	1
Confidence Rules		
Rule No.	**Rule Description**	**Confidence Score**
1	P_{ov} has no fugacity or property warnings	10
2	P_{ov} has fugacity and/or property warnings	5

References

Beyer A, Mackay D, Matthies M, Wania F, Webster E. 2000. Assessing long-range transport potential of persistent organic pollutants. Environ Sci Tech. 34:699-703.

Zarfl C, Scheringer M, Matthies M. 2011. Screening criteria for long-range transport potential of organic substances in water. Environ Sci Technol. 45(23):10075-10081.

19. APPENDIX X: Rules for Exposure Classification and Confidence Scoring of Emission Pattern

Two mode of entry (MOE) scenarios were considered for the RAIDAR simulations, which affect the calculation of P_{ov}. The first is "predominantly release to water" and the second is "100% to soil". Given an emission to water, chemicals with neutral form log air-water partition coefficients (log $K_{AW,N}$) <-6 were not considered to be multi-media chemicals given that significant escape from water to air and then deposition to soil is possible during industrial use and release. Table 1 describes the "mostly water" mass fraction division of the emission quantity according to log K_{AW}, noting that chemicals identified as acids or bases were assumed to be 100% released to water regardless of K_{AW}.

Table 1. Mode of entry (MOE) assumptions for chemical release in the "predominantly to water" simulation scenario

Log $K_{AW,N}$	Air	Water	Soil
<-6 or "acid" or "base"	0%	100%	0%
>= -6 and <-5	20%	75%	5%
>-5 and <-4	30%	65%	5%
>-4 and <-3	40%	55%	5%
>-3 and <-2	50%	45%	5%
>-2 and <-1	65%	30%	5%
>-1 and <0	75%	20%	5%
>0	85%	10%	5%

Classification and confidence rules for emission pattern are based on chemical quantity and overall persistence based on either an emission to water or soil. These are presented in Table 2. All quantity data represent the combined total of import and manufactured tonnages of a substance in pure form (i.e., not the quantity contained in products or articles). When chemical quantity data (kt/yr) was available (n=425) from any CEPA section 71 survey (2012-2017), it was selected preferentially over DSL 1986 data. When multiple years of CEPA section 71 data were available, the most recent reporting year value was used and if a CEPA section 71 value was not reported, it was assumed no longer in commerce. When CEPA section 71 data were not available (>99% of ERC2 substances), geometric mean values from 1986 DSL tonnage data were calculated for each tonnage bin used in the 1986 survey as summarized in Table 3 (from ARC 2018). A 10-fold scaling factor was applied to the 1986 mean values based on an ECCC comparative analysis of recent DSLIU survey data and the corresponding 1986 mean chemical quantity

for ~1750 substances. This analysis revealed that quantity data from 1986 generally under predict more recent CEPA section 71 survey data by a factor of seven on average. If a factor of 10 is used only 5% of the ERC2 substances are under-predicted using CEPA section 71 data. However, this analysis remains biased by the fact that chemical quantity data from 1986 was reported as ranges (bins) whereas more recent quantity data are single values. Therefore, confidence assignment to all extrapolated quantity data remains low. All inorganics, organometallics, organic-metal salts, and polymers were removed from this analysis to avoid bias for potential high volume non-organic chemicals. Survey and extrapolated chemical quantity values were then applied as the "actual emission" rate (E_a) in RAIDAR fate and exposure modeling.

Table 2: Rules for Selection of Chemical Quantity (Mass)

P_{ov} Classification Rules		
Rule No.	**Description**	**Exposure Class**
0	If WWTS sludge adsorption is >85%, then soil P_{ov} used, else P_{ov} water used	NA
1	P_{ov} >=60 days	3
2	P_{ov} = 21-60 days	2
3	P_{ov} <=21 days	1
P_{ov} Confidence Rules		
Rule No.	**Description**	**Confidence Value**
1	P_{ov} has no fugacity or property warnings	5
2	P_{ov} has fugacity and/or property warnings	3
Quantity Classification Rules		
Rule No.	**Description**	**Confidence Value**
1	Chemical quantity is >100 000 kg/yr (>0.1 kt/yr)	3
2	Chemical quantity is 10 000-100 000 kg/yr (0.01-0.1 kt/yr)	2
3	Chemical quantity is < 10 000 kg/yr (<0.01 kt/yr)	1
4	No quantity data available	NA
Quantity Confidence Rules		
Rule No.	**Rule Description**	**Confidence Score**
1	CEPA section 71 survey data available	5
2	CEPA section 71 survey data unavailable	3
Emission Pattern Classification Rules		
Rule No.	**Rule Description**	**Exposure Class**
1	Class 3 P_{ov} + Class 3 Quantity	3
2	Class 2 P_{ov} + Class 3 Quantity	2

3	Class 3 P_{ov} + Class 2 Quantity	2
4	Does not meet rules 1-3	1
Emission Pattern Confidence Rules		
Rule No.	**Rule Description**	**Confidence Score**
3	P_{ov} + Quantity Confidences	6-10
2	P_{ov} + Quantity Confidences	6-10
1	P_{ov} + Quantity Confidences	6-10

Table 3: Selected geometric mean values for chemical quantity ranges (bins) reported for DSL organic substances in 1986

DSL Reported Values (kt)	**Selected Emission Rate (E_A; kt/y)**
$0 – 10^{-4}$	3.16×10^{-5}
10^{-4} or $>10^{-4}$	3.16×10^{-3}
$10^{-4} – 10^{-3}$	3.16×10^{-3}
$0 – 10^{-3}$	3.16×10^{-4}
$10^{-3} – 1$	3.16×10^{-1}
$10^{-3} – 10^{-2}$	3.16×10^{-2}
10^{-2}	0.1
$>10^{-2}$	30
>1	100
Not Available	3.16×10^{-3}

References

[ARC] ARC Arnot Research & Consulting. 2018. Generation of physical-chemical property data and the application of models for estimating fate and transport and exposure and risk potential for organic substances on the Canadian DSL. Gatineau (QC): Environment and Climate Change Canada.

20. APPENDIX XI: Rules for Exposure Classification and Confidence Scoring of Food Web Exposure

Classification thresholds for food web exposure using the RAIDAR exposure assessment factor (EAF) are given in Table 1. The thresholds were determined by examining the range of EAF values to determine the relative distribution of each exposure class according to the thresholds selected. Values closer to 1 represent increased ability of the environment to deliver a substance to a vulnerable food web species in the RAIDAR Model (ARC 2018).

Table 1: Rules for classification and confidence scoring of food web exposure

Exposure Classification Rules		
Rule No.	**Rule Description**	**Exposure Class**
1	EAF is >=0.001	3
2	EAF is 0.0009-0.00009	2
3	EAF is <0.00009	1
Confidence Rules		
Rule No.	**Rule Description**	**Confidence Class**
1	No property or fugacity warnings	10
2	Property or fugacity warnings present	5

References

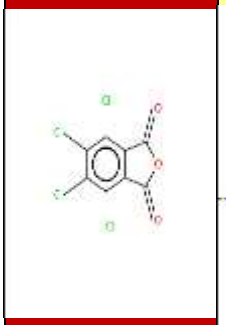
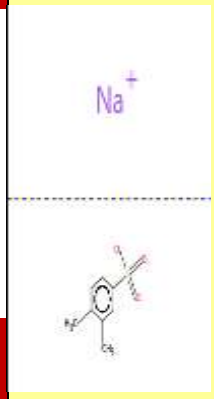
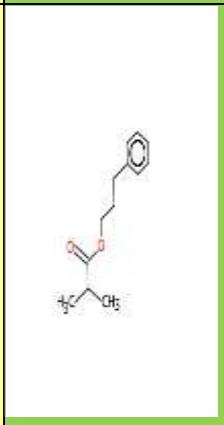
[ARC] ARC Arnot Research & Consulting. 2018. Generation of physical-chemical property data and the application of models for estimating fate and transport and exposure and risk potential for organic substances on the Canadian DSL. Gatineau (QC): Environment and Climate Change Canada.

21. APPENDIX XII: Excel Table of Main ERC2 Results for ~12 200 Organic Substances

Provide link [here] for public release

22. Appendix XIII: ERC2 Profile and Classification Output for Three Example Substances

Table 1: Example ERC2 profile and classification results for high, moderate and low risk outcomes

ERC2 Component	Descriptor	Target Endpoint or Property	High Priority Substance	Moderate Priority Substance	Low Priority Substance
Chemical Identity	Chemical name and CAS RN	Chemical name and CAS RN	1,3-isobenzofurandione, 4,5,6,7-tetrachloro-, [CAS RN 117-08-8]	xylenesulfonic acid, sodium salt, CAS RN [1300-72-7]	3-phenylpropyl ester, CAS RN [103-58-2]
	2D structure	2D structure			
Use Pattern	Known or estimated use	Use	Colorant (predicted), Flame retardant/chemical intermediate (known)	Ubiquitous uses (e.g., emulsion stabilizer, solvent in household products; printing paste additive)	Fragrance/ flavorant
	Estimated tonnage	kt/yr	100	100	0.0003
ADME	Bioavail-ability	Permeability	permeable	permeable	permeable
	Distribution	Plasma Binding	not plasma distributed	not plasma distributed	not plasma distributed
Hazard Profile	Receptor Mediated	Estrogen (ER), androgen (AR), thyroid (THY), aryl hydrocarbon	in vivo effects unconfirmed to be resulting from receptor mediated interactions	very weak or no interactions	very weak or no interactions

		receptor (AhR)			
	Chemical Reactivity and Geno-toxicity	Proteins and nucleic acids	in vivo effects plausibly linked to DNA and protein interactions	possible interactions with proteins	no interactions
	Mode of Action	Various	Acyrlation-based reactivity	narcosis (polar)	narcosis (esters)
	Foodweb Toxicity	RAIDAR Hazard Assessment Factor (HAF) (unitless)	High potential foodweb toxicant	low potential foodweb toxicant	low potential foodweb toxicant
	Cumm. Toxicity	According to Mode of Action	not calculated	not calculated	not calculated
Exposure Profile	Response Time	Time to 97% removal from all environmental media (yrs)	14.5	0.6	0.2
	Mobility	travel distance (km)	441 (local to regional scale)	370 (local to regional scale)	500 (local to regional)
	Emission Pattern	exposure potential/mode of entry	High tonnage substance, long residence time, emitted to water and soil via biosolids	High tonnage substance, short residence time, emitted to water only	Low tonnage substance, short residence time, emitted to water only
	Foodweb Exposure	RAIDAR Exposure Assessment Factor (EAF) (unitless)	Significant accumulation potential in foodwebs	Not significantly accumulated in foodwebs	Not significantly accumulated in foodwebs

	Margin of Exposure	Ratio of critical emission (kg/hr)/actual emission (kg/hr)	< 1000 (439)	< 1000 (15)	> 1000 (~93000)
Hazard Outcomes	Classification	Numerical Score and Category	3 (high)	2 (moderate)	1 (low)
	Confidence	Numerical Score and Category	68 (high)	2 (very low)	29 (moderate)
	Severity	Numerical Score and Category	4 (very high)	1 (low)	0 (very low)
Exposure Outcomes	Classification	Numerical Score and Category	3 (high)	2 (moderate)	1 (low)
	Confidence	Numerical Score and Category	28 (moderate)	16 (low)	38 (high)
	Severity	Numerical Score and Category	3 (high)	1 (low)	0 (very low)
Risk Outcomes	Classification	Numerical Score and Category	6 (high)	4 (moderate)	1 (low)
	Confidence	Numerical Score and Category	96 (moderate)	18 (very low)	67 (moderate)
	Severity	Numerical Score and Hazard+Exposure Category	7 (very high-high)	2 (low-low)	0 (very low-very low)
Assessment Flags	Flag to consider if nominated for further assessment	Margin of Exposure <1000, low permeability, plasma distributed, >99% fraction ionized, organic salt of metal, UVCB	Margin of Exposure	Margin of Exposure, >=99% fraction ionized	no flags

Possible Assessment Activity	**Assessment, data gathering, no further action**	**N/A**	**Assessment**	**Data gathering**	**No further action**