MEMORANDUM

To: Eric Hooker, Matthew Cho, and Kristina Hatlelid, CPSC

From: Samantha Snow, Joei Robertson, and Elizabeth Martin, ICF
Mark Bradley and Lynne Haber, University of Cincinnati

Date: October 18, 2024

Re: Subtask 3A: Identify approaches to perform class-based risk assessment
(CPSC Contract Number 61320622A0005, CPSC Order Number 61320623F2030)

---

1. Overview of Memorandum

Under CO-9, Subtask 3A seeks to identify approaches for conducting class-based risk assessment. To find these methods, a targeted literature search was conducted followed by title/abstract and full text screening to identify relevant references.

Class-based risk assessment approaches are described in the below memo. The steps of risk assessment were identified, and methods used to extrapolate within a chemical class are detailed by the steps that integrate into risk assessment. In addition to summarizing the approaches found, this memo notes their feasibility for use in Subtask 3B, which will involve the development of quantitative hazard indexes or risk values for chemicals in the polyhalogenated organophosphate (PHOP) organohalogen flame retardants (OFRs) subclass.

The polyhalogenated bisphenol aliphatics and functionalized (PHBAF) subclass will not be discussed in this report as the class-based risk assessment approaches depend on findings from CO-1 (Call Order 61320622F2011, “Guidance Document for Conducting Class-based Hazard Assessment of Organohalogen Flame Retardants” under CPSC Blank Purchase Agreement (BPA 61320622A0005)) and CO-3 (Call Order 61320622F2012, “Class-based Exposure Assessment Approaches for Organohalogen Flame Retardants (OFRs): Process Guide and Case Study Application for up to Two Classes” under CPSC Blank Purchase Agreement (BPA 61320622A0005)), which were only conducted on chemicals within the PHOP subclass.

2. Literature Search for Describing Approaches to Class-based Risk Assessment

ICF and UC conducted a search for relevant methods papers for this subtask based on three data streams:

1) Literature identified via searches of scientific literature databases (i.e., PubMed and Web of Science) with no date restrictions;
2) Literature identified via a search of the grey literature, or publications produced outside of traditional scientific publishing avenues, including reports from government agencies and authoritative bodies not generally available in databases of published literature; and
3) Publications and reports identified by subject matter experts.

Under Subtask 2A, we conducted a literature search to identify approaches for conducting class-based risk assessment. Following deduplication and prioritization (Section 3.1), the papers identified were screened in the title/abstract step (Section 3.2). The search strings and strategies for this subtask are presented in Appendix A. Refer to the Subtask 2A report for the complete literature search methods.

2.1. Literature Search Strategies

The following sections describe the literature search strategies used for databases and additional sources. The literature search strategies included searches within core literature databases (i.e., PubMed and Web of Science) as well as relevant domestic and international non-periodical grey literature, such as technical reports, monographs, and symposium proceedings prepared by select committees or bodies (e.g., National Academy of Sciences).

2.2. Seed Studies

Subject matter experts identified 33 references to use as positive seed studies to help verify the accuracy and efficiency of our literature search (Attachment 1; Figure 1). These relevant references were identified from reference compilations identified by the report authors as part of other projects or from general monitoring of the literature. The goal was to identify key references related to the use of approaches for applying class-based risk assessment, with particular focus on data-rich and data-poor chemicals. These studies were used as seed studies for the database searches described in Section 2.3 and in the ensemble supervised clustering described in Section 3.1. Of the 33 seed studies, 7 studies were captured in our database search (Section 2.3) and 1 study was captured in our grey literature search (Section 2.4). Although the low capture rate for the seed studies in the initial literature search may raise concerns about the adequacy of the initial literature searches, the initial searches were heavily supplemented by expert review of the grey literature and tree searching, as described below. In addition, missing individual studies are of less concern for this task than for an assessment, since the goal was to capture the general approaches that are in the literature.

2.3. Database Searches

The literature searches for peer-reviewed publications were conducted on December 4, 2023 in Pubmed and Web of Science with no date limit. Removing duplicates across the searches resulted in a total of 4,672 unique references (Attachment 1; Figure 1). These references were moved to a topic extraction step (Section 3.1), before going forward into title/abstract screening using Litstream®.

2.4. Identification of Grey Literature

There were 11 agency websites identified by ICF and UC, and confirmed by CPSC, for review of relevant grey literature (0). Grey literature was pulled from these relevant websites using an ICF-developed Google webscraping tool. This tool crawled the Google results page for a search that had been limited to a specific domain and pulled links and metadata associated with the search results into an Excel file. Each search was limited to pull in only the top 10 hits per source (for 11 sources) per search string (for 3 topic specific search strings), meaning a maximum of 330 grey literature search results would be pulled by the webscraper.

A total of 259 references/websites were returned from the Google webscraping (Attachment 1; Figure 1). These items underwent a first pass screening by a subject matter expert to assign a priority rating (i.e., not relevant, maybe relevant, include) and to identify the topic or key take-

away of the references identified as maybe relevant or include. If the initial webscraping tool identified a general website, that website was reviewed to identify any useful/relevant references, and those references were flagged for downloading if they were not included in the list from the webscraping tool. A second review was conducted of the literature marked as maybe relevant or include, in order to identify the most recent relevant references. Key topics were those that addressed methods for read across, class-based assessments, cumulative risk/combined exposure, and use of NAMs in risk assessment. Please note that not all literature was reviewed in detail.

2.5. Expert Identified Literature

Key agency guidance documents that the authors were aware of but had been missed by the webscraping tool were downloaded and categorized as expert identified literature. In addition, subject matter experts may have identified additional peer-reviewed or grey literature sources during their synthesis writing (e.g., as a reference cited in a paper that was captured in our literature search) that resulted in them pulling in expert identified literature. A total of 50 expert identified literature were included in this report (Figure 1; Attachment 1).

3. Prioritization and Literature Screening

3.1. Prioritization of Literature for Screening

The peer-reviewed literature identified via database searches (n = 4,672) underwent a prioritization step prior to title/abstract screening. To prioritize these 4,672 studies, three unique approaches were applied: (1) keyword analysis, (2) unsupervised clustering, and (3) supervised clustering. Keyword analysis was selected as the best approach to optimize inclusion of likely relevant literature and resulted in 1,489 papers to be screened at the title/abstract level. See Subtask 2A report for complete details on screening prioritization.

3.2. Title/Abstract Screening

A total of 1,489 studies were identified using keyword analysis and uploaded into ICF's Litstream® software. A screening form was developed using inclusion/exclusion criteria to guide decisions about the relevance of the literature. Screeners reviewed the title and abstract to determine whether the reference should be included, excluded, or marked unclear. If a reference was marked as unclear, it underwent a QC step by a senior screener to make the final decision. For a reference to be included for this subtask, it needed to describe an approach to perform class-based risk assessment.

Each screener completed a pilot of 30 references, which were then QC'd and feedback provided if there were any discrepancies between the primary screener and senior screener. After successfully completing the pilot, screeners continued working until all 1,489 references were reviewed. Each reference was only reviewed by one screener following the pilot. At the completion of title/abstract screening, there were 593 references tagged to approaches to perform class-based risk assessment.

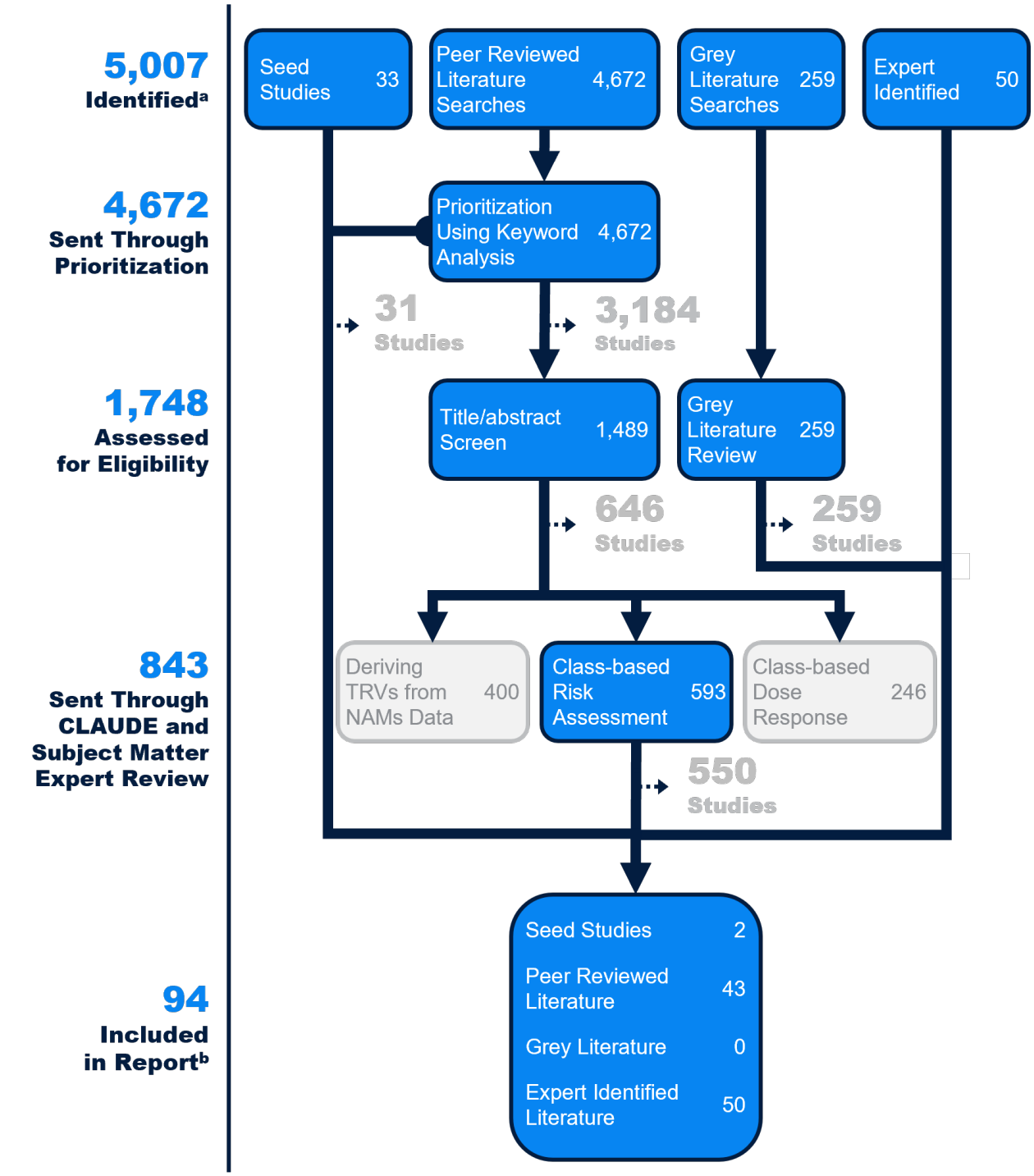
3.3. Full Text Screening

From title/abstract screening, 593 peer-reviewed studies were marked as 'include' for identifying an approach to perform class-based risk assessment. Of those, 569 had PDFs that were retrievable. To prioritize full text PDFs, a generative artificial intelligence (GenAI) workflow

utilizing the large language model Claude 3.0 was deployed. Screening literature is not an unintended or prohibited use of Claude 3.0. For this workflow PDFs were converted to text by Azure Document Intelligence. 25 PDFs were too large for conversion and were removed from subsequent analysis leaving 545 summarized studies. The 25 large PDFs were reviewed by subject matter experts to prevent loss of information. Next, a set of 9 prompts were tested against 10 expert curated PDFs (Appendix Table B-1). These prompts were then graded for accuracy (correct, mostly correct, incorrect, or hallucinating[1]) and were determined to be sufficient for use with this subject matter topic. These 9 prompts were used to prioritize papers into bins based on report directives. Additionally, parameters used for this prioritization included "Acting as a toxicologist" and Claude 3.0 (Appendix Table B-2).

[1]Hallucination = When a large language model makes up an answer that is completely unrelated to the source material.

Figure 1. Literature Search Results



[a] The total number of studies that were identified as a seed study, through the peer reviewed literature searches, through the grey literature searches, and as an expert identified study is higher than the total number of studies included in the report because several studies (n=7) were identified from multiple sources.
[b] The total number of studies listed in the final blue box is higher than the total number of studies included in the report because one study was identified from multiple sources. Specifically, one seed study was also captured in our peer reviewed literature search (Maffini et al., 2023).

4. Overview of Traditional Risk Assessment for Single Chemicals and Extrapolation to Class-based Risk Assessment

4.1. Traditional Risk Assessment

Risk assessment is the process of characterizing the impacts to human health following exposure to a given chemical. Generally, the Environmental Protection Agency (EPA) breaks this down into four basic steps: hazard identification, dose-response assessment, exposure assessment, and risk characterization (U.S. EPA, 2024). Hazard identification/hazard characterization describes the types of health effects caused by a given pollutant including weight of evidence considerations such as mode of action (MOA), human relevance of effects observed in experimental animals, and the dose/route of exposure associated with different health effects. Dose-response assessment describes the relationship between the amount of exposure and the observed health effect (e.g., incidence and severity). Exposure assessment describes the magnitude, frequency and duration of exposure, and the size and characteristics of the population exposed. All of these evidence streams feed into risk characterization. Risk characterization includes both a quantitative description of the risk associated with an exposure, and a qualitative description of the associated uncertainties.

Risk assessment for single chemicals is conducted using different methods for threshold endpoints (including non-cancer effects and cancer effects thought to be induced by a non-mutagenic MOA), and non-threshold effects (cancer effects thought to be induced by a mutagenic MOA).

For chemicals causing threshold effects, a toxicity reference value (TRV) such as a reference dose (RfD) or Acceptable Daily Intake (ADI) is calculated by dividing a point of departure (POD) by a composite of uncertainty factors representing interspecies differences (UFA), intraspecies differences (UFH), and use of a lowest-observed-adverse-effect level (LOAEL) instead of a no-observed-adverse-effect level (NOAEL) or benchmark dose lower confidence limit (BMDL) (UFL). Some organizations, such as the US EPA, also include factors for database uncertainty (UFD), and extrapolation from subchronic to chronic duration (UFS). Different organizations may use different sets of uncertainty factors. While CPSC generally uses UFA, UFH, and UFL for traditional single-chemical assessments, we are also considering UFS for use in a class-based risk assessment.

$$TRV = \frac{POD}{UFA \times UFH \times UFL \times UFS}$$

Regardless of the specific uncertainty factor approach used, the risk characterization is then conducted by comparing the resulting TRV to a level of real-world exposure, resulting in a hazard quotient (HQ)

$$HQ = \frac{Exposure}{TRV}$$

Generally, HQ < 1 indicates adverse effects are unlikely, while HQ > 1 indicates an increased potential for adverse effects, though HQ does not specifically quantify the degree to which this potential increases.

Another way of characterizing risk is the margin of exposure (MOE) approach, where a POD is divided directly by a real-world exposure. The MOE is then compared to a minimum margin of

safety. This minimum margin of safety considers factors similar to the uncertainty factors used to derive TRVs. Depending on the regulatory organization, the MOE approach can be used for both noncancer assessments and evaluation of chemicals causing cancer via a mutagenic MOA.

$$MOE = \frac{POD}{Exposure}$$

For chemicals causing non-threshold effects, a cancer slope factor (CSF) can be calculated as:

$$CSF = \frac{Response}{POD}$$

For example, when tumor data are modeled, the POD is typically the lower 95% confidence limit on the benchmark dose corresponding to a 10% response (BMDL10); and the corresponding response is 0.1 (for the 10% response at the BMD10). In the risk characterization step, the CSF (a measure of cancer potency, in units of risk per unit dose) is multiplied by the dose to estimate risk:

$$Lifetime\ Excess\ Risk = Dose\ \times CSF$$

Note that, unlike the HQ or MOE, the result here is an estimate of risk.

For chemical mixtures (assuming the chemicals are similar in terms of toxic effects, MOA, kinetics, etc.), the HQ for each individual chemical can be summed to create a hazard index (HI). While mixture assessments and class-based assessments are not necessarily the same, mixture assessment methods may still inform class-based assessments.

$$HI = HQ_{ChemA} + HQ_{ChemB} + HQ_{ChemC}$$

Similarly, the cancer risk for a mixture can be determined by adding the risk for individual chemicals.

$$Total\ Lifetime\ Excess\ Risk = Lifetime\ Excess\ Risk_{ChemA} + Lifetime\ Excess\ Risk_{ChemB} + Lifetime\ Excess\ Risk_{ChemC}$$

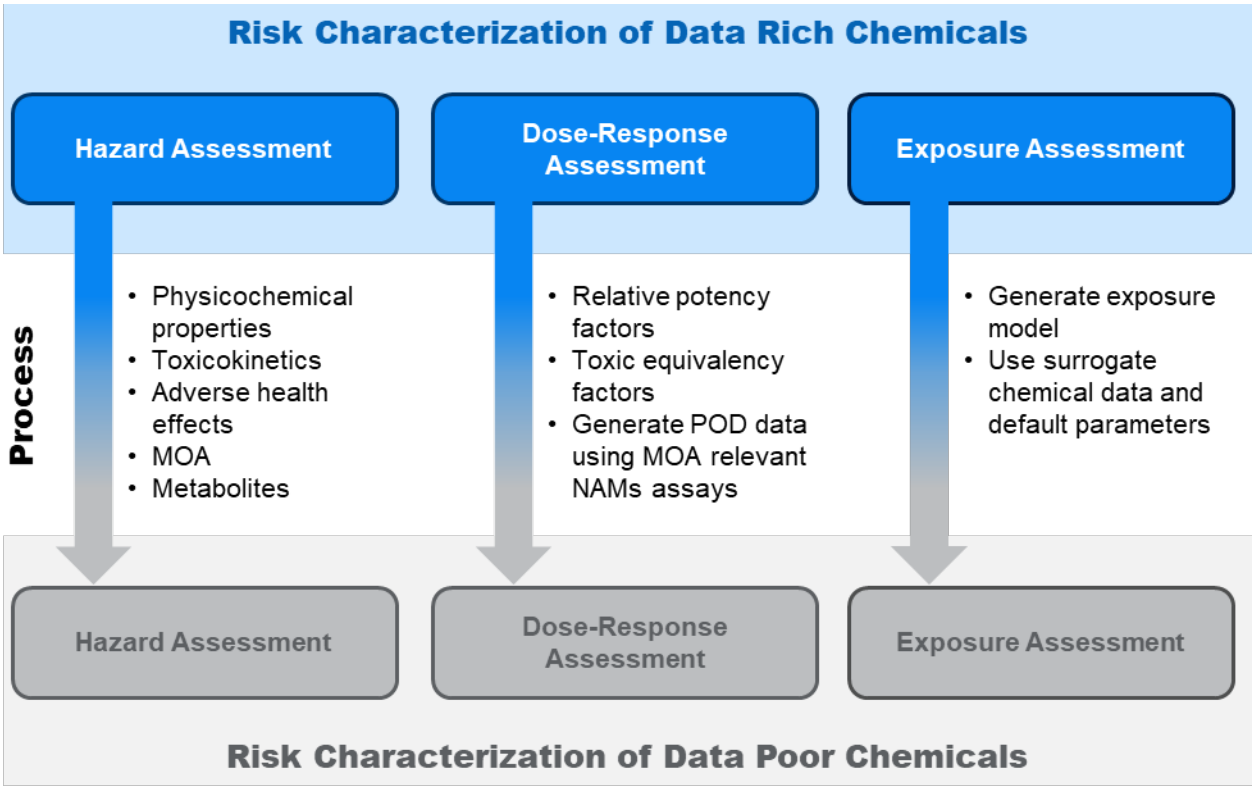
4.2. Class-based Risk Assessment

Although there is a relatively extensive literature on grouping chemicals into classes for cumulative risk assessment, the literature on using class-based approaches for filling data gaps, particularly for dose-response and exposure assessment, is more limited. One well conducted example of a class-based cumulative risk assessment is the one prepared by EPA Office of Pesticide Program (OPP) for the Pyrethrins/Pyrethroid group. The Pyrethrins/Pyrethroid group is a group of 21 chemicals that qualify as a Common Mechanism Group (CMG) (U.S. EPA, 2011). Considering the existing and proposed registered uses of the chemicals, EPA identified three exposure pathways of interest including food, drinking water, and residential/non-occupational settings. A screening level assessment was conducted for the oral, dermal, and inhalation routes of exposure using multiple high-end inputs, resulting in conservative exposure estimates. Specific chemicals assessed for each pathway varied, as chemicals were excluded if evidence did not exist for exposure potential, exposure data was not available, or exposure was limited to minor exposure pathways, routes, or uses. Hazard criteria were also used to exclude chemicals, including chemicals unlikely to have an impact on the estimated cumulative risk such as those with a relative potency factor (RPF) significantly lower than that of the index chemical.

Considering these exposure and hazard criteria, only 15 of the 21 chemicals in the CMG were ultimately included in the quantitative cumulative risk assessment.

The goal of this report is to describe methods of class-based risk assessment, meaning the process of extrapolating risk information about data rich chemicals to data poor chemicals within a class (Figure 2). Here we describe methods used to extrapolate exposure (Section 4), hazard (Section 5), dose-response (Section 6) and risk characterization (Section 8), with the ultimate goal of applying these methods to the PHOP chemical class of OFRs. The general principles that govern these processes are found both in the OPP cumulative risk guidance (U.S. EPA, 2002) and CPSC Chronic Hazard Advisory Panel on Phthalates (CHAP) report (CPSC, 2014). Importantly, applications of these methods have noted that lack of any data about a chemical's MOA, environmental prevalence, and potential for toxic interaction among members of a chemical class add uncertainty to estimates of risk (Ahlers et al., 2019; Hines et al., 2018; Safe, 1997).

Figure 2. Overall Description of Class - based Risk Assessment Processes



MOA = mode of action; POD = point of departure; NAMs = new approaches/alternative methods.

5. Class - based Exposure Assessment Methods

Exposure assessment is "the process of estimating or measuring the magnitude, frequency, and duration of exposure to an agent and the size and characteristics of the population exposed" (U.S. EPA, 2019). Exposure assessments have generally been conducted for single-pathway and single-chemical evaluations, with fewer assessments conducted for aggregate exposure (i.e., exposure to a single chemical from multiple sources and pathways) or cumulative exposure (i.e., exposure to multiple chemicals from multiple sources and pathways that affect a

single biological target). Conducting single chemical exposure assessments for numerous chemicals at the same time can be (i) time consuming and (ii) challenging if data are not available. This section investigates class-based exposure assessment approaches, where multiple chemicals are assessed at one time in a group or class.

Based on the literature search workflow described in Section 2, few studies were identified that proposed and/or implemented a class-based exposure assessment framework. Of the relatively few exposure assessment studies identified in the literature search, most were single chemical assessments – the low number of exposure assessment studies is an artifact of the literature search strategy which used a risk assessment keyword search string in addition to an exposure keyword search string (specifically combined, aggregate, or cumulative exposure) and therefore likely excluded exposure-only studies. The literature search workflow and an additional targeted internet search identified a limited number of class-based exposure assessments, as well as studies that were amenable to class-based use. These studies are described below, along with a summary of the class-based exposure assessment work performed under a different call order (Contract BPA No. 61320622A0005, Call Order No. 61320622F2012 [CO-3]).

5.1. Existing Class - based Exposure Assessments or Frameworks

Class-based approaches to exposure assessments have been used in agency risk assessments and in regulatory settings. Examples include the EPA OPP cumulative risk framework and assessments for pesticides with a common mechanism of toxicity (U.S. EPA, 2002) and CPSC, Health Canada, and EPA Office of Pollution Prevention and Toxics (OPPT) proposed approach for cumulative risk assessment of phthalates (CPSC, 2014; Health Canada, 2020; U.S. EPA, 2023). While these risk assessments focus on a group of chemicals, the approach to estimate exposure is similar to individual chemical assessments, where exposure is estimated using one or more of the following approaches:

- **Indirect or scenario-based estimation:** estimates exposure for a specified exposure scenario based on chemical concentration (in the exposure media), exposure factors, and exposure models. Surrogate datasets are frequently used from other chemicals or from similar use patterns and products to fill in data gaps.
- **Exposure reconstruction or biomonitoring (reverse dosimetry):** uses human biomonitoring data to estimate the total dose for all sources and pathways to which a person is exposed.
- **Direct or point-of-contact measurement:** measures a person's direct exposure to a chemical over an identified time period (e.g., monitoring data in food).

Typically, aggregate exposure for a chemical is then calculated as the sum of the doses from different sources (e.g., total exposure for di-isodecyl phthalate (DIDP) was equal to the sum of exposures from diet, toys, furniture, gloves, and dust). For cumulative exposure to phthalates, CPSC and Health Canada both use the reverse dosimetry approach while EPA OPPT's proposed approach uses a relative potency factor approach. Recent work for CPSC under a different call order (Contract BPA No. 61320622A0005, Call Order No. 61320622F2012 [CO-3]) has explored the possibility of addressing data-poor chemicals where empirical data are not available. We provide below a summary of two example assessments.

5.1.1. EPA OPP: Cumulative Risk Assessment (CRA) for Pyrethrins/Pyrethroid

As discussed above in Section 4.1, US EPA OPP conducts cumulative risk assessments for chemicals with a common mechanism of toxicity. For the Pyrethins/Pyrethriods CRA (as an example), the exposure assessment used both chemical specific monitoring data and fate and transport modeling methodologies. Specifically, food exposures were determined in a probabilistic manner from pesticide residue monitoring data collected by the United States Department of Agriculture (USDA) Pesticide Data Program (PDP) combined with USDA dietary consumption data. PDP samples with non-detectable residues were treated in this assessment as "zero" values. For drinking water exposures, monitoring data with sufficiently low detection limits were not available, therefore EPA modeled drinking water concentrations using multiple models (e.g., FIRST, SCI_GROW, and PRZM/EXAMS), which incorporated application rates from pesticide product labels. In cases where modeled estimates were greater than the limit of solubility for the chemical, the chemical's limit of solubility was used. Residential exposures focused on the main use scenarios (e.g., application and post-application activities associated with turf, indoor environment, garden, and pet pesticide use) and were quantified in a deterministic manner using chemical specific application rates from product labels or chemical-specific residue data from monitoring/residue transfer studies, combined with various default exposure factor assumptions. In residential cases where monitoring data were not available for an individual chemical or specific scenario, surrogate data (i.e., most conservative or most representative) were used with proper adjustments. Since residential exposure from different routes occur simultaneously, all routes were combined within a scenario. Additionally, residential scenarios that could potentially co-occur were combined, for example a child contacting a treated pet, treated turf, and a treated indoor surface on the same day. Finally, the dietary and residential pathways were combined into a screening level pyrethroid cumulative assessment, using the 99.9th percentile MOE for the dietary pathway.

While the approach could be refined by using probabilistic estimates for all routes and linking with a physiologically-based pharmacokinetic (PBPK) model, which would allow calculation of distributions across the US population and investigate uncertainty, EPA decided that the screening approach was sufficient. It utilized a simple point-estimate aggregate methodology that adds a high-end dietary (i.e., food and water) exposure to a high-end residential exposure, assuming co-occurrence of dietary and non-dietary exposure.

5.1.2. CPSC, Health Canada, EPA: Cumulative Risk Assessment for Phthalates

In cumulative risk assessment approaches for phthalates, U.S. CPSC (CPSC, 2014), Health Canada (Health Canada, 2020), and EPA OPPT (proposed) (U.S. EPA, 2023) estimated intake using a scenario-based approach employing indirect exposure estimates (similar to methodologies used by EPA OPP) and also estimated intake using a reverse dosimetry approach with human biomonitoring data, specifically metabolites in urine. Exposure methodologies generally followed EPA's Guidelines for Human Exposure Assessment (U.S. EPA, 2019). According to EPA OPPT, Health Canada and U.S. CPSC found that both approaches resulted in daily intake values that were generally similar in magnitude. Uncertainties and limitations of the two approaches are discussed in Table 6-5 of U.S. EPA (2023).

5.1.3. CPSC: Polyhalogenated Organophosphate Flame Retardants

Under CO-3, exposure to PHOP flame retardants was assessed using four approaches:

i) Approach 1: mechanistic models to estimate indoor environmental concentrations and/or doses due to consumer product use;

ii) Approach 2: empirical measurements including migration rates into saliva, handwipe data, personal air monitoring data, and product emissions/parameters to estimate contact (direct) and mediated (indirect) consumer exposure;

iii) Approach 3: indoor dust monitoring data to estimate total dose for all sources related to indoor dust; and

iv) Approach 4: reverse dosimetry with human biomonitoring and chemical-specific toxicokinetic data to estimate dose for all sources and pathways to which a person is exposed.

Under Approach 1, consumer exposure was estimated individually for all PHOP chemicals with known or potential use in consumer products using a mechanistic model (Consumer Exposure Model, CEM) although several limitations and uncertainties were noted. However, based on data availability, only a limited number of chemicals could be assessed using the other three approaches. The correlation between modeled CEM doses and each of six physicochemical parameters were evaluated, with the goal of identifying the key predictor of absorbed dose, which is defined as the dose entering body tissues. For the dermal and total (i.e., sum across dermal, inhalation, and absorption) exposure pathways, molecular weight was the key predictor, where the absorbed dose increased as the molecular weight of the chemical decreased. For the inhalation pathway, the absorbed dose increased when vapor pressure increased. For the ingestion pathway, the absorbed dose increased when the octanol-air partition coefficient increased. These key predictors were then used to estimate relative dose for the PHOPs with no empirical data (i.e., data-poor) for Approaches 2 and 3. For indoor dust data, semi-quantitative doses were estimated for 21 chemicals with product use data where the reported dose was relative to a data-rich chemical. Both empirical measurements (specifically handwipes, personal air, and product emissions) and reverse dosimetry approaches had data for only three or fewer chemicals and the doses from empirical data were generally calculated using a limited number of data sets, leading to inconclusive extrapolation results.

The analyses performed under CO-3 demonstrated how different data types and sources could be combined together and used to extrapolate doses from data-rich to data-poor chemicals in a class. In addition, the different approaches of calculating exposure also allowed for comparison and corroboration of whether an approach is over- or underpredicting exposure. Additional work is still needed to address the limitations identified and incorporate more exposure data.

5.2. Potential Exposure - based Chemical Groupings

The example assessments and frameworks noted in the previous section were for chemical groupings based on hazard similarities (i.e., chemicals that behave in a similar biological manner with shared endpoints of interest). While most chemical groupings are often based on hazard similarities, mixture studies that examine the co-occurrence of chemicals in the context of exposure may provide a framework for grouping chemicals for class-based exposure assessments.

As part of an iterative approach for assessing exposures to environmental mixtures, Rice et al. (2008) proposed several ways for grouping chemicals:

(i) chemicals can be grouped by the primary release mechanism that controls the initial contamination and migration behavior in the environment where examples of release processes include volatilization to air and leaching to groundwater;

(ii) chemicals can be grouped based on physicochemical properties that are relevant to fate in soil and air where examples of relevant parameters include volatility, water solubility, and partition coefficients; and

(iii) chemicals can be grouped based on both the media they partition into and the timing of when the chemical is expected to occur in the exposure media (i.e., a 2x2 matrix with four possible groupings).

Grouping by physicochemical properties was also suggested by Cattaneo et al. (2023), in addition to grouping based on regulatory criteria, exposure scenarios, functional groups, chemical structures, and toxicological properties. Bopp et al. (2018) suggested groupings similar to Cattaneo et al. (2023) (i.e., chemical class, biological effect) but also proposed grouping by product/use, co-emission (considering origin from one source), and receptor. Groupings will depend on the context and purpose of the assessment, with Bopp et al. (2018) providing examples of groupings based on structure (e.g., phthalates), technical function (e.g., solvents), and biological effect (e.g., pesticides toxic to the thyroid and central nervous system). While Diamond et al. (2018) did not discuss chemical groupings, the representative chemicals for their mixture assessment of down-the-drain chemicals in wastewater treatment plant effluent were selected by first identifying activities that would results in disposal of chemicals down the drain, suggesting product/use and fate as the drivers.

5.3. Data Sources Amenable for Class-based Exposure Assessment

As previously mentioned in Section 5.1, three approaches are commonly used to estimate exposure. Exposure estimates (as a concentration, dose, or loading) can be obtained from monitoring studies, quantified through mechanistic models, or derived from empirical evidence. All estimation methods may be suitable to class-based exposure assessment. This section focuses on studies which describe aspects of these exposure methods or quantify exposure for multiple chemicals. For class-based exposure assessment, this ideally includes defined groups/classes of chemicals or a wide range of chemical classes with grouping of chemicals performed after the fact. In addition, chemical mixture studies were also considered as they may provide a framework for grouping chemicals. The amenable study types and studies identified are discussed below.

5.3.1. Monitoring Studies

Monitoring studies typically provide concentrations of chemicals in environmental media (e.g., indoor dust) or biomonitoring matrices which can then be used to estimate dose or intake. High-throughput monitoring studies generate exposure estimates for tens-to-hundreds-to-thousands of chemicals at once. Studies may be (i) targeted, where samples are analyzed for a pre-defined list of chemicals or (ii) non-targeted, where the aim is to identify all chemicals present.

Targeted studies provide quantitative estimates, but the number of chemicals measured is usually limited by the availability of analytical methods. For example, European surveys for polar compounds in surface and wastewaters were typically limited to ~50 target chemicals, but the development of wide-scope chemical target screening methods in recent years has allowed the quantification of concentrations of hundreds of chemicals (Finckh et al., 2022). In Finckh et al. (2022), 366 out of 499 target emerging chemicals, primarily pharmaceuticals and pesticides, were quantified in wastewater treatment plant effluent samples with concentrations ranging from <1 ng/L to >100 µg/L. Often, a targeted study may choose to focus on only one or two chemical groups (see Section 5.1.3 for potential chemical groupings) such as in Lopez et al. (2017), where the authors focused on quantifying only pesticides and reported concentration of 40

pesticides in ambient air. In addition to pesticides, other chemical groupings in monitoring studies include per- and polyfluoroalkyl substances (Fiedler et al., 2022; Genualdi et al., 2022; Ulrich et al., 2024), disinfection byproducts (specifically trihalomethanes and haloacetic acids, each of which is regulated and monitored as a group) (Almonacid Garrido et al., 2024; Rodriguez et al., 2003; Zhao et al., 2020), and polychlorinated biphenyls (Du et al., 2020; Saktrakulkla et al., 2020) among many others. Targeted monitoring data from individual studies and other sources such as government programs are sometimes compiled into databases such as EPA Multimedia Monitoring Database (Isaacs et al., 2022), Center for Disease Control (CDC), National Health and Nutrition Examination Survey (NHANES), or the United States Geological Service (USGS) Water Quality Portal, which provides a single source to obtain data on a large number of chemicals. However, depending on the database, the metadata may not be as detailed as found in the original source, providing some uncertainties in the reported values, and may require statistical analysis prior to implementation in exposure models. Additionally, even collectively, most sources often do not cover a wide breadth of geographic locations or different populations.

Non-targeted studies can cover a wider chemical space (i.e., thousands of chemicals) but their results are primarily qualitative, focusing on chemical identification. For example, Rostkowski et al. (2019) identified or tentatively[2] identified over 2,300 compounds in a composite house dust sample, Zhao et al. (2013) assigned approximately 3,000 molecular formulas to high resolution mass spectra of cloud water samples, and Cui et al. (2023) reported a total of 30, 78, 103, 20, and 265 frequently identified annotated features[3] (and between 700 and 5,100 total features) in food, dust, soil, water, and urine samples, respectively, from households with children. In a review on how non-targeted analysis data can be used in the chemical risk assessment process, McCord et al. (2022) noted that while there is currently no universally accepted method for quantifying the results from non-targeted screening, several approaches exist that allows for semi-quantitative estimates to be used for prioritization. The three approaches described in McCord et al. (2022) are: (i) relative quantification, (ii) quantification by surrogate standard, and (iii) response modeling from chemical structure.

As noted by McCord et al. (2022), existing risk paradigms ultimately require quantitative estimates of chemical abundance to support regulatory decisions by identifying concentrations from targeted studies, but non-targeted studies can fill data gaps by informing on the presence of chemicals and putative sources, routes, and pathways.

5.3.2. Modeling Studies

Exposure models, whether mechanistic or empirical, can provide predicted concentrations, doses, or other key parameters needed to calculate dose. Mechanistic models are based on well-established mechanistic processes, while empirical models are typically data-fitting regressions using empirical measurements (e.g., product testing emissions data, migration data).

High-throughput exposure models can generate estimates for a large number of chemicals with relatively minimal input requirements, for example, with just physicochemical properties and in the case of consumer exposures, product composition and product usage characteristics.

[2] Tentative identification defined as confidence level 2 and confidence level 3, where confidence is based on the level system in Schymanski et al. (2014).
[3] A feature is described by three coordinates: accurate mass, retention time, and signal intensity in Minkus et al. (2022).

Depending on the model, one or more exposure pathways[4] may be modeled, and results are often used for screening and/or prioritization purposes. Examples of high-throughput mechanistic exposure models and corresponding case studies using these models include Stochastic Human Exposure and Dose Simulation – high-throughput (SHEDS-HT), which models exposures from consumer product categories or food groups (Isaacs et al., 2014), and ExpoCast ("Exposure Forecasting"), a companion project to ToxCast, which uses two existing multimedia mass balance models, USEtox and RAIDAR, to estimate the amount of chemical in air, water, and soil due to industrial releases to the environment (Wambaugh et al., 2013). Many additional models are also available in the peer-reviewed literature although these models do not necessarily have a model acronym/name. For example, Ernstoff et al. (2017) developed a high-throughput migration model that estimates the fraction of chemical migrating from food packaging to food and this model was later modified by Aurisano et al. (2022) for estimating chemical migration from solid products to saliva during mouthing.

Not all mechanistic exposure models are high-throughput models. Some models may require more complex input parameters and/or the available tool/platform does not lend itself to running multiple exposure scenarios at once. Dutch National Institute for Public Health and the Environment's (RIVM) three exposure modeling tools (with case studies provided) – Probabilistic Aggregate Consumer Exposure Model (PACEM), which estimates aggregate consumer exposure (Delmaar et al., 2023)], ConsExpo, which estimates consumer exposure (Delmaar & Meesters, 2020), and DustEx, which estimates dust-mediated exposures to sources in homes (Ball et al., 2016) – are all available as web-based tools but there is currently no option that allows for the user to perform batch runs (e.g., uploading an input file with inputs for hundreds of runs). EPA's CEM (which estimates consumer exposure) is available as an Access database file but requires the user to set up and run each individual exposure scenario. However, for class-based exposure assessment, these models can still be used, especially for smaller chemical classes where running the models for each chemical/scenario may not be too burdensome.

Empirical models can also be used to estimate exposure or key parameters needed to calculate exposure. For example, in Aurisano et al. (2022), in addition to the mechanistic model for estimating chemical migration to saliva, the authors also developed a regression-based model using data available in the literature from 60 chemicals. Liang et al. (2019) measured the emissions of organophosphate ester flame retardants from foam materials at different temperatures and determined the material-phase diffusion coefficient (D_m) and the material/air partition coefficient (K_{ma}) at each temperature. The authors then developed correlations between temperature and emission parameters (i.e., steady-state emission rate, D_m, K_{ma}).

While existing mechanistic and empirical models are available that can be applied to class-based exposure assessment, most currently are not efficient at running large numbers of chemicals at the same time and may require expert level professional input to parameterize the models.

[4] Exposure pathways considered under different call orders (Contract BPA No. 61320622A0005, Call Order Nos. 61320622F2014 [CO-2] and 61320622F2012 [CO-3]) included: (i) ingestion of indoor dust; (ii) gas-phase air transfer to skin; (iii) inhalation of particle dust; (iv) inhalation of gas; (v) dermal contact; and (vi) mouthing. Additional pathways not considered in these call orders include the dietary pathway.

6. Class-based Hazard Assessment Methods

There is no single method for grouping chemicals for hazard assessment. Grouping of chemicals relies on methods that have been developed for read across approaches and/or mixtures modeling. Read-across methods are methods that are used to group chemicals together based upon structural similarities to extrapolate information about the toxicity of a chemical without significant experimental testing. Alternatively, if chemicals have associated toxicological data, such as in the case of pesticides, they can be grouped by MOA or health effect concerns. Mixtures modeling methods are methods that are used to estimate toxicity for a group of chemicals that occurs together in the environment. Once a group is identified, the extrapolation process between chemicals relies on the identification of anchor or surrogate chemicals (i.e., chemicals within a class for which a significant amount of toxicological data has been collected).

In the previous Subtask 2A report, we described guidelines and examples of grouping of chemicals for risk assessment. This included a review of guides by regulatory and adjacent organizations, a brief overview of literature examples that group chemicals for dose response, and a description of an approach developed by ICF specific to organohalogen flame retardants for CPSC as a part of the overarching contract (Contract BPA No. 61320622A0005, Call Order No. 61320622F2011 [CO-1]).To briefly summarize, Organization for Economic Co-operation and Development (OECD), European Chemicals Agency (ECHA), and European Food Safety Authority (EFSA) have published guides on conducting read across, each reflecting their specific purposes. The Research Institute for Fragrance Materials (RIFM) also uses read across but primarily publishes single chemical assessments using the analog method. These resources provide methods to group chemicals using the category approach to read across, which links chemicals by shared characteristics or property trends. Each organization focuses on methods that are unique to their regulatory processes and goals. These references are directly meant for grouping of chemicals for hazard assessment, although they can be applied to grouping for dose-response.

Additionally, examples of chemical grouping by shared MOA were described. However, this method is limited to data-rich classes of chemicals, such as well studied pesticides and polycyclic aromatic hydrocarbons (PAHs). MOA and mechanisms of toxicity are robust links for chemical grouping, but this method is severely limited in application since the MOA or mechanisms are only reported for a few chemicals. Further details on the OECD, ECHA, EFSA, and RIFM guides, or examples of grouping chemicals by MOA specifically for dose response can be found in the Subtask 2A report and in Maffini et al.'s detailed description of each agency's framework (Maffini et al., 2023).

6.1. Grouping Chemicals by Mode of Action

During hazard characterization, one of the key steps is identifying the mechanism of effect. This idea encompasses absorption, deposition, metabolism, and excretion (ADME) profiles and MOA (i.e., a high-level description of the key steps – termed key events – in the process by which a chemical may act to cause toxicity). During grouping exercises, chemicals are identified as being a part of a class based on shared toxicity characteristics. Often the focus has been on which characteristics should be chosen to delineate a class. Chemicals can be grouped based on their ability to cause various assay endpoints during high-throughput screening or by computational predictions, which estimate toxicity potential. For example, chemicals can be grouped together by those that cause cytotoxicity (Malich et al., 1998), hepatotoxicity (Li et al., 2021; Rowan-Carroll et al., 2021), genotoxicity (Bus et al., 2015; Czaja et al., 2020; Dourson et al., 2013; Enoch et al., 2022a, 2023; Enoch et al., 2022b; Paini et al., 2021; Verdon et al.,

2022), or neurotoxicity (Crofton et al., 2022). Some efforts have also focused on well-defined MOAs to define chemical subclasses, such as in the case of cobalt-containing substances (Danzeisen et al., 2022), or structure-dependent MOAs (Liu et al., 2019). Other important screening assays that can help prioritize and group chemicals include metabolic assays that can inform whether a genotoxic metabolite will form (Enoch et al., 2023) or bioaccumulate (Salvito et al., 2020).

Importantly, read across methods to group chemicals and extrapolate health hazards to data poor chemicals rely on the likelihood of these shared chemical structures to imply shared MOAs (Lai et al., 2022; Salvito et al., 2020). Extrapolation that relies on chemical structure rather than on assay readout is of limited utility for MOA grouping given that diverse chemical structures can have the same molecular target, such as the diverse structure of estrogen receptor agonists (Fang et al., 2001). Additionally, similarly structured chemicals can have diverse molecular targets, such as the diversity of carcinogenic potential among trihalomethanes (Woo et al., 2002). Both scenarios highlight the difficulty of grouping chemicals by MOAs without assay data, and the advantages of understanding how MOAs can map to adverse outcomes.

6.2. Grouping Chemicals by Health Effect

While grouping chemicals by their toxicological MOA offers higher confidence due to its specificity, grouping by adverse effects allows for the use of more diverse data types and does not necessarily require the same MOA. Without shared MOA, quantitative cumulative risk assessment of a group of chemicals becomes more challenging and the results share more uncertainty.

Several agencies have identified classes of chemicals grouped by structure and associated health effects including PAHs (U.S. EPA, 2008), phthalates (CPSC, 2010; U.S. EPA, 2012), and organophosphate pesticides (Brock et al., 2003; EFSA Scientific Committee et al., 2019; U.S. EPA, 2006); many of these are well studied and shared mechanisms have been identified. For chemicals that do not fit into agency classes (most chemicals), grouping can be accomplished by shared target organs and/or systems, or even broadly by their ability to bioaccumulate in mammalians. Although not as strong of an association, a shared ecotoxicity or nonmammalian bioaccumulation target can also be used to support grouping for human hazard assessment (Danforth et al., 2020).

Extrapolating adverse effects to data poor chemicals relies on shared characteristics with the chemicals that have adverse outcome data. Characteristics such as chemical substructures (Enoch et al., 2022a; Enoch et al., 2022b; Schor et al., 2022; Smith et al., 2005), physicochemical properties (predicted or experimental) (Grimm et al., 2016; Redman et al., 2014), toxicokinetic properties (predicted or experimental) (Al-Malahmeh et al., 2017; van den Berg et al., 2012), and/or metabolites (Enoch et al., 2023; Olesti et al., 2021) all provide support for including chemicals without data into a group classified by health effect. For example, there are thousands of known per- and polyfluoroalkyl substances (PFAS) chemicals of ongoing interest to researchers and regulatory agencies for class-based hazard assessment. PFAS vary widely in structure, size, and even properties, but groups of similarly structured PFAS have been shown to have shared health effects, such as hepatoxicity potency of perfluoro carboxylic and sulfonic acids of a certain carbon chain length range (Bil et al., 2021; Reardon et al., 2021). As fully explored in CO-1 (Contract BPA No. 61320622A0005, Call Order No. 61320622F2011), OFRs are similar to PFAS in this regard, and structural grouping is useful for determining health effects by subclass (NASEM, 2019).

7. Class-based Dose-response Modeling

7.1. General Framework for Class-based Dose-response Methods

Once chemical groupings have been established, as is described in Section 6, a variety of approaches for conducting class-based dose-response can be utilized (Bennekou, 2019; Crump et al., 2010; Lutz et al., 2014). These are discussed in detail in the Subtask 2A report. As for hazard characterization, these approaches draw largely from read across and from cumulative risk methods and consider how to integrate new approaches/alternative methods (NAMs) data with more traditional data. The Subtask 2A report highlights relative potency factors (RPFs) and toxic equivalency factors (TEFs) as key methods from cumulative risk that may be useful in a class-based dose-response assessment with both methods assuming dose additivity. TEFs require greater similarity among chemicals (similarity specific to target organs, routes, and effects) than do RPFs (similarity to either target organs, routes, or effects), and so RPFs are likely to be useful in a greater number of class-based dose-response assessments. Broadly, the key steps for conducting a class-based dose-response assessment were outlined as:

i) identification of toxicity data for read across, including NAMs data in the case of OFR dose response,

ii) conducting dose-response evaluation and calculation of relative potency,

iii) identification of appropriate uncertainty factors, and

iv) calculation of TRVs.

7.2. Use of Mode of Action in Read Across

While the method above describes a relative potency framework, read across methods can be used in the absence of sufficient high-quality quantitative data. Risk assessors have the greatest confidence in the results of read across when there is a clear understanding of the MOA for the class of chemicals that includes the analog and the target, as well as information on how toxicity varies with structure. For example, if the MOA is known and toxicity can be related to a structural characteristic, a regression or a structure activity relationship (SAR) based on that characteristic could be used quantitatively in a read across approach. In other cases, when the MOA is well understood and an endpoint related to the MOA is used to group the chemicals, a relative potency approach can be taken in which chemicals are ranked relative to each other and toxicity data is extrapolated within the group. In addition, it is important to ensure that MOAs unique to individual chemicals within a structurally-related grouping are considered. For example, n-hexane is neurotoxic, but n-heptane and n-pentane are not. The difference is due to the metabolism of n-heptane to the gamma diketone 2,5-hexanedione (Moser et al., 2019).

Conversely, when the MOA is not known or well understood, but there is a regular pattern of variation of the strength of an effect (e.g., toxicity varying with the length of an alkyl chain), toxicity can be predicted based on the pattern, or based on the most toxic within a class (ECHA, 2017). Importantly, this approach has uncertainties not associated with MOA, but it could be used to address data gaps in a health-protective manner. An additional approach used when there is no MOA data, is a weight of evidence approach that may be used to show that a surrogate is similar to other related chemicals. Such decisions need to take into account multiple lines of evidence, similar to those described for grouping chemicals into classes

(Section 6), The absence of MOA information also means that there is higher uncertainty in the estimate.

Finally, the estimate of dose-response from an anchor or surrogate chemical can be applied to data poor members of the class. These methods are addressed by a number of publications and guidance documents (Ball et al., 2016; ECHA, 2017; Patlewicz et al., 2018; Schultz et al., 2019). This approach assumes that there are shared biological endpoints of interest and requires less chemical-specific information than a relative potency approach, but is associated with greater uncertainty, similar to cases where the MOA is not known. Thus, this approach is better suited to qualitative prediction (e.g., is the chemical a carcinogen) than quantitative predictions, and the specific data requirements depend on the problem formulation. For example, the information requirements are higher in a regulatory context than those for informing internal risk management decisions.

7.3. Use of New Approaches Methods to Address Data Gaps

In light of the data needs for conducting class-based dose-response analyses, it is expected that NAMs, particularly transcriptomics, will play an increasing role in conducting such assessments, and can be used to calculate RPFs (see the Subtask 2A report for additional details). High-throughput methods make it possible to test multiple doses of multiple chemicals under the same experimental conditions, something that is rarely done with *in vivo* testing. Consideration of adverse outcome pathways (AOPs) can aid in integrating non-traditional animal model data with rodent data (Hines et al., 2018), as well as choosing the most useful *in vitro* assays for conducting a risk assessment (Clewell et al., 2020).

Further, NAMs can be used to gather molecular MOA data in a high-throughput manner, and benchmark doses (BMDs) calculated from such data can be used to derive PODs (De Abrew et al., 2019). In the Subtask 2A report, we highlighted methods such as mechanistic mapping of chemicals (De Abrew et al., 2019), integration of pathway-based transcriptomic data for quantitative risk assessment, AOP mapping (Pitzer et al., 2023), and generation of new data utilizing *in vitro* systems (Rowan-Carroll et al., 2021). In general, these efforts to gather MOA data to calculate PODs can be coupled with *in vitro* to *in vivo* extrapolation (IVIVE) and high-throughput exposure estimation to calculate MOE, and, less commonly, to calculate TRVs. This method of leveraging molecular data is the most promising effort to perform class-based dose-response assessment (ICCVAM, 2024).

In general, uncertainty factors for a class-based dose-response analysis will depend on the data used. Discussion of uncertainty factors for NAMs (likely to be necessary for many class-based dose-response analyses) have modified traditional uncertainty factors and/or added additional uncertainty factors (Dourson et al., 2022; Health Canada, 2020). For example, if the POD comes from a human cell line, the interspecies uncertainty factor would not be needed, but it may be appropriate to add additional factor(s) accounting for limitations of *in vitro* assays or extrapolation from *in vitro* to *in vivo* effects.

8. Class-based Risk Characterization

Information from the exposure assessment, hazard identification assessment, and dose-response modeling is synthesized in the risk characterization step to describe the risk under the specified exposure conditions, and to clearly communicate key findings and their strengths and limitations. This integrative step typically involves expert judgement and an accounting of variability and uncertainty to reach an overall conclusion about the risk of a substance to human health. While a quantitative risk estimate is a part of that conclusion, a qualitative description of the considerations made in determining the quantitative risk estimate is critical to provide the information necessary to interpret the risk results in a way that is transparent, clear, consistent, and reasonable (U.S. EPA, 2000).

This section focuses on risk characterization as its own step. But it is worth noting that the same key principles described above can and should be applied to individual characterization of the exposure assessment, hazard identification assessment, and dose-response assessment steps. In particular, Alexander-White et al. (2022) proposed a framework for the use of read across in risk assessment, including final steps analogous to risk characterization (Text Box 1). Specifically, they lay out questions to consider when assessing overall confidence when SAR approaches are utilized (Text Box 2). Similar guiding questions would be useful in a class-based risk assessment.

Text Box 1. 10 Steps for Risk Assessments using Read Across

- Tier 0
  - Step 1: Identify exposure/use scenarios
  - Step 2: Identify molecular structure of target chemical
  - Step 3: Collate supporting data on target chemical(s) and define data gap
  - Step 4:
    - a) Identify analogue(s),
    - b) collate existing data, and
    - c) determine similarity hypothesis (at this stage still mostly structure-based but can be refined to biological similarity – mode/mechanism of action (MOA) where data is existing)
- Tier 1 refinement
  - Step 5: Supporting similar bioavailability/ADME (absorption, distribution, metabolism, and excretion) of target chemical and analogues
  - Step 6: Supporting similar MOA hypothesis
- Tier 2 refinement
  - Step 7:
    - a) Targeted testing to strengthen MOA hypotheses and/or
    - b) Biokinetic refinements of target chemical and analogues
  - Step 8: Performing a risk assessment to derive a point of departure
  - Step 9: Perform a margin of safety/margin of internal exposure evaluation using risk assessment
  - Step 10: Assessing confidence in the risk assessment

Text Box 2. Prompts for Assessing Confidence in Risk Assessments using Read Across

- Describe overall level of confidence (low, medium or high) that the read across is appropriately conservative as part of an exposure driven risk assessment. Throughout the whole process, uncertainties and the level of confidence in the data should be captured as transparently as possible and integrated to provide an overall level of confidence in the assessment.
- Examples of questions to address:
  - What type of category formation was attempted and was it suitable for the context of the read across?
  - How well made was the premise or hypothesis of the read across argument?
  - What rationale was used to select the NAMs used and how did they support the decision making?
  - How was mechanism of action considered supported and assessed?
  - How was similarity in toxicodynamics/effects defined and assessed to support the MOA?
  - How was similarity in toxicokinetics/potency defined and assessed?
  - What were the uncertainties in the toxicological data for read across data and how did they allow for an assessment of robustness of these data?
  - How were NAMs applied and did they assist in the reduction of uncertainty?
  - What are the key strengths of the case study?
  - What are the key limitations of the case study?

Prompts were described in Step 10: Assessing confidence in the risk assessment in Table 2 of Alexander-White et al. (2022).

8.1. Cumulative Risk Assessment Methods

Much of the past work on class-based risk assessment has been done in the context of cumulative risk from mixtures (Section 4). While the specifics of cumulative risk are not directly applicable to the current work on OFRs, as these guidelines focus on concurrent exposure to similar chemicals rather than individual exposure to similar chemicals, some considerations are informative. There are a number of models and guidance documents for cumulative risk assessment, as summarized in the Subtask 2A report. Two of these approaches are particularly worthy of a more detailed consideration here. One is the assessment of phthalates (CPSC, 2014), since that is a peer-reviewed CPSC product. The second general approach is that of EPA's OPP (U.S. EPA, 2002), since that organization has conducted cumulative risk assessments for several different groups of pesticides. A key aspect of OPP's cumulative risk assessment approach is the identification of the cumulative assessment group, which includes chemicals that share a common toxic mechanism and will contribute more than minimally to cumulative hazard and exposure. For the OFRs, the risk characterization could be done on an individual chemical basis, comparing the modeled or estimated exposure for each chemical with the corresponding TRV for the chemical (CPSC, 2014).

The CPSC phthalate assessment used the HQ/HI approach, with a few noteworthy adaptations. First, because male developmental and reproductive toxicity via an antiandrogenic MOA was determined to be the critical effect, the HQ/HI calculation was based on such endpoints, even if the relevant TRV for a given individual phthalate was not based on antiandrogenicity. Instead, CPSC derived potency estimates for antiandrogenicity (PEAA), when needed, to ensure that the dose-response for each chemical was based on comparable endpoints. These PEAA values were described as having a purpose solely in cumulative risk assessment, and not to be

confused with a TRV that is used in a regulatory context. The PEAA for a given chemical was determined by dividing a POD specific to the anti-androgen MOA by an uncertainty factor specific to that POD. Three different sources for PEAAs were used, as a measure of uncertainty. The first source was from a published cumulative risk assessment for mixtures of phthalates. The second used values from "recently published and highly reliable relative potency comparisons across phthalates from the same study," and the third source was from values from the *de novo* literature review conducted for the phthalates assessment. The daily intake for a given phthalate was then divided by the PEAA for that phthalate, resulting in the HQ. The assessment for the PHOPs is similarly based on a combination of published assessments and new data, but the large number of chemicals with no TRVs or no reliable relative potency data makes it impractical to follow a similar approach for the PHOPs.

CPSC (2014) noted that, as an alternative to the HQ, the point of departure index [PODI; see also Reffstrup et al. (2010)] was considered as the basis for the cumulative risk assessment. In the PODI approach, daily intake is compared with the POD, rather than with the PEAA, which is based on both the POD and case-specific uncertainty factors. The PODI approach uses one overall uncertainty factor instead of having the potential for different values of the total uncertainty factor for different phthalates. Because the quality of the toxicological data (i.e., certainty based upon experimental design) varied across the phthalates, the PEAA approach including substance-specific uncertainty factors was preferred. Similar reasoning would apply to the OFRs, in light of the varying quality of data for them.

According to the OPP cumulative risk assessment approach, in order to calculate a cumulative MOE, the exposure is first expressed in concentration equivalents of an index chemical. That is, rather than adjusting the potency measurement, the exposure measurement is adjusted to reflect potency. Presumably this approach is followed so that differences in uncertainty factor are separated from empirical potency differences. A cumulative MOE is calculated for the class for each route of exposure, expressed in equivalents of the index chemical.

$$MOE_{Route} = \frac{POD_{Index}}{\sum Exposure_{Route}}$$

8.2. Uncertainties in Class-based Risk Assessment

For risk characterization, OPP (U.S. EPA, 2002) emphasizes describing confidence in the data and models used for toxicity and exposure estimations, including uncertainty as well as any potential biases (direction and magnitude). This is also the step where sensitivity analyses may be conducted, and decisions are made about uncertainty/safety factors. Some key issues noted related to the choice of uncertainty/safety factors include "the identification of the nature of the common toxic effect and the scope of the toxicity and exposure databases relative to the cumulative assessment."

A potential challenge for class-based risk characterization is in handling data gaps as many chemicals that have exposure data do not have dose-response data. Use of computational predictions to apply methods such as RPFs is one method, but as noted in Section 7, has its own set of limitations. This is likely to be relevant to class-based assessments of OFRs, given the generally sparse data availability.

8.3. Polyhalogenated Organophosphate Flame Retardants

For the PHOPs, as noted in Section 5.1.3 for work performed under CO-3, it was possible to calculate consumer exposures for all PHOPs that had product use data. However, for the few data rich chemicals where dose estimates based on NHANES were possible, the aggregate

consumer exposure estimates based on the mechanistic models were higher, suggesting that further refinements to the modeling input parameters are needed before the estimates from mechanistic models can be used in bounding risk characterization as is more fully discussed in CO-3 (Contract BPA No. 61320622A0005, Call Order No. 61320622F2012) Estimated doses based on empirical data were calculated for only a few of the PHOPs. Semi-quantitative estimates of dose from indoor dust were predicted for 21 chemicals even in the absence of dust monitoring data, where the reported dose was relative to a data rich chemical (Contract BPA No. 61320622A0005, Call Order No. 61320622F2012 [CO-3]). Although there is substantial uncertainty in the dose estimates using the semi-quantitative estimates, these estimates do allow for a rough risk characterization. Comparing these dose estimates with the TRV for the most toxic class member might allow one to set aside the chemical as not of concern (if there is a substantial MOE) or identify the chemical for further testing and analysis.

9. Discussion

In this report, we describe methods used for class-based risk assessment, building on the prior work in class-based exposure assessment in CO-3, class-based hazard characterization in CO-1, and class-based dose-response assessment in the Subtask 2A report of the current call order. The application of methods for individual chemicals is noted, with particular attention to how these methods are modified for class-based assessments. The work to date on the application of class-based methods to the available data for PHOPs is discussed, with particular attention to the way in which the approach has evolved with progress through different steps of the class-based assessment. Generally, our findings suggest that we can extrapolate information at each step of the risk assessment process (hazard identification to group chemicals, exposure assessment to gather measured or estimated values, and dose-response assessment to estimate TRVs) that can then be used to calculate HQs, HIs, and lifetime excess risk. These values will rely upon the extrapolation methods that are used at each phase of the risk assessment. Based on our review of the literature, it does not seem that there is a good way to holistically estimate any of these values without relying on estimation in prior steps. Furthermore, there is no agency level guidance on class-based risk assessment at this time.

We found that several different organizations have conducted class-based assessments, particularly in the context of cumulative risk. Structural similarity and shared mechanism or MOA were used in grouping the chemicals. Most of these assessments appear to have been done using a “bottom up” approach, starting with individual chemicals, which were then formed into groups based on the available data. In contrast, the OFR task has many elements of a “top down” process, having evolved by the following process. The National Academies of Science, Engineering, and Medicine organized an initial list of chemicals into subclasses (NASEM, 2019), which have been refined in subsequent work under CO-1. The key implication of starting with the larger list of chemicals for the OFR analysis is that many of the chemicals are data poor or completely lacking in data. This means that a key aspect of the current work is addressing data gaps, while most of the existing literature on class-based assessments (aside from the literature on read across for hazard and dose-response) appears to focus on using MOA to identify classes, consideration of the potential for toxic interactions, and addressing data gaps only with the classic uncertainty factors.

The exposure assessment step is perhaps the most well described with regard to addressing chemical-specific data gaps in a risk assessment context. Section 5 of the current report summarized a variety of approaches to class-based exposure assessment. Many of the class-based exposure assessments were done in the context of cumulative risk assessment and were based on empirical exposure data for chemicals acting via a common MOA (Sections 5.1).

However, exposure modeling has been done in both high-throughput (Isaacs et al., 2014) (Wambaugh et al., 2013) and a variety of lower throughput models, including ones that estimate consumer exposure (Section 5.3.2) (Delmaar & Meesters, 2020; Delmaar et al., 2023; Wang et al., 2023). Other exposure studies have grouped or proposed grouping chemicals by a variety of properties or considerations, including exposure scenarios, chemical function, and toxicological properties (Bopp et al., 2018; Cattaneo et al., 2023), but do not appear to have used this grouping to fill data gaps.

In CO-3 (class-based exposure), empirical exposure measurements for indoor dust, handwipe, personal air, and product emissions data were available for only a few of the PHOPs (Contract BPA No. 61320622A0005, Call Order No. 61320622F2012). It was possible to estimate consumer exposure using mechanistic models for all of the PHOPs that had product use data. However, for the few data rich chemicals, the aggregate consumer exposure estimates based on the mechanistic models were higher than the estimates for the same chemicals based on reverse dosimetry from NHANES, suggesting an overinclusion of sources and the need for further refinement of the modeling input parameters. Using indoor dust data, semi-quantitative estimates of dose relative to data rich PHOPs were predicted for 21 PHOPs with product use data; these dose estimates represented the dose from all sources related to indoor dust. Thus, at least some estimate of dose (from consumer products and indoor dust sources) is available for all of the PHOPs with use data.

In CO-1 (read-across), data on the PHOPs were evaluated based on five profiles: metabolite, PBTK, physicochemical properties, mechanistic effects, and adverse effects. Five OFRs were determined to be poor fits to the PHOP subclass, because they were outliers in the five physicochemical and biological profiles (Contract BPA No. 61320622A0005, Call Order No. 61320622F2011). The remaining PHOPs were grouped as to the confidence that the chemical belongs to the PHOP subclass. Four anchor chemicals with relatively well-defined profiles were identified and rated as very high confidence; two additional chemicals were included in the very high confidence category. The remainder of the chemicals in the subclass were divided about evenly into categories of high and medium confidence that they are part of the PHOPs subclass.

Section 7 of the current report describes methods for class-based dose-response assessment, including relative potency and read across approaches. These methods were described in additional detail in the Subtask 2A report. Class-based dose-response assessment has been used extensively in pesticide evaluations of cumulative risk of common mechanism groups and using read across under REACH.

Of the four elements of the risk assessment paradigm, the dose-response assessment of the PHOPs faces the largest data gaps. The available dose-response information on PHOPs has been summarized and analyzed in prior reports. The Subtask 1A deliverable summarized existing TRVs developed in authoritative assessments, and the Subtask 1B report reviewed the toxicological literature to develop TRVs for chemicals with relevant *in vivo* data but no authoritative TRV. The Subtask 1C report reviewed the transcriptomic data for the chemicals considered in Subtask 1A and Subtask 1B, with the goal of using the transcriptomic data to fill data gaps. However, the transcriptomic data were too limited to address data gaps, both because the data were from only a small group of chemicals, and because most of the transcriptomic data were not from the (potential) target tissues. The CO-1 Task 4B case study report laid out a series of recommendations for targeted testing of the PHOPs identified as belonging to the subclass with medium confidence or higher (i.e., all of the chemicals remaining in the class after the chemicals with poor fits are excluded). After review of additional toxicity data on this class as part of the current Call Order, this testing strategy continues to be recommended.

As such, for the current Call Order, it is likely that several different methods will be needed for a class-based risk assessment, based on the approaches described in Section 8, depending on the available data for the individual PHOP and confidence that the chemical is a member of the subclass. For example, based on the available data, it is possible that an RPF approach could be used for the PHOPs with some *in vivo* data, while read across from an analog may be the only feasible approach for the PHOPs with limited data. For the PHOPs lacking any toxicity data, it is possible that only a bounding analysis (i.e., evaluating the potential results using plausible inputs from the extremes) can be done, comparing the modeled exposure with a toxicity benchmark from an analog. The initial ideas presented here will need to be further investigated in the context of the data for the specific chemicals in future work. Regardless of the approach chosen, the overall assessment for each chemical will need to take into account the overall weight of evidence. This analysis will need to consider the implications of data gaps and ensure that the target chemicals are not more toxic than the analog used for read across, based on the limited available data and any plausible trends or associations of available parameters with toxicity. If such limited data indicate that using the analog data is not sufficiently protective, but are insufficient of themselves for developing a TRV, then another dose-response surrogate, such as the threshold of toxicological concern (TTC), would need to be employed. In general, however, the TTC approach is expected to be a highly conservative measure of toxicity, and so is usually used to show that a chemical exposure is not of concern, or that more data is needed on the chemical, rather than saying that a particular exposure is of concern.

Appendix A. Literature Search String and Strategy

A.1. Peer Database Search

The peer database search was executed on 12/04/2023.

A.1.1. Describing Approaches to Perform Class-based Risk Assessment

A.1.1.1. PubMed

Search String:

("Risk Assessment/methods"[Mesh] OR "Risk Assessment/standards"[Mesh] OR Method*[ti] OR Approach*[ti]) AND ("chemical class" OR “Chemical classes” OR "Chemical Grouping" OR “Chemical Groupings” OR "Chemical Group" OR “Chemical Groups” OR “Surrogate chemical” OR “Surrogate Chemicals” OR “Read Across” OR “Chemical Mixture” OR “Chemical Mixtures” OR “Combined exposure” OR “combined exposures” OR “aggregate exposure” OR “Aggregate exposures” OR “Cumulative Exposure” OR “Cumulative Exposures” OR ("Class-based"[tiab] AND Chemical*))

Limits: None

A.1.1.2. Web of Science

Search String:

TS=("Risk Assessment" AND (Standard* OR Method*)) AND TS=("chemical class" OR “Chemical classes” OR "Chemical Grouping" OR “Chemical Groupings” OR "Chemical Group" OR “Chemical Groups” OR “Surrogate chemical” OR “Surrogate Chemicals” OR “Read Across” OR “Chemical Mixture” OR “Chemical Mixtures” OR “Combined exposure” OR “combined exposures” OR “aggregate exposure” OR “Aggregate exposures” OR “Cumulative Exposure” OR “Cumulative Exposures” OR ("Class-based" AND Chemical*))

Limits: None

A.2. Grey Literature Google Webscraping

All grey literature searches were executed on 12/04/2023. Table A-1 lists the grey literature sources that were searched using the Google webscraping tool.

Table A-1. Grey Literature Sources

Source	Type	URL Domain
US Environmental Protection Agency (EPA)	Federal, Regulatory	epa.gov
National Institute of Environmental Health Sciences (NIEHS)	Federal, Research	niehs.nih.gov
Centers for Disease Control and Prevention (CDC)	Federal, Regulatory	cdc.gov
National Technical Reports Library (NTRL)	Federal, Repository	ntrl.ntis.gov
National Academies of Sciences, Engineering, and Medicine (NASEM)	NGO, Research	nationalacademies.org
National Institute for Public Health and the Environment (RIVM)	International, Regulatory	rivm.nl

Source	Type	URL Domain
European Chemicals Agency (ECHA)	International, Regulatory	echa.europa.eu
European Food Safety Authority (EFSA)	International, Regulatory	efsa.europa.eu
Organisation for Economic Co-operation and Development (OECD)	International, Research	oecd.org
Health Canada (HC)	International, Regulatory	canada.ca
Japanese Ministry of Health	International, Regulatory	mhlw.go.jp

A.2.1. Describing Approaches to Perform Class-based Risk Assessment

The below search string was used. Due to strict character limits in Google searching, the string was edited from the peer database string

("Risk Assessment" AND (Standard* OR Method*)) AND ("chemical class" OR “Chemical classes” OR "Chemical Grouping" OR "Chemical Group" OR “Surrogate chemical” OR “Read Across” OR “Chemical Mixture” OR “Combined exposure” OR “combined exposures” OR “aggregate exposure” OR “Cumulative Exposure” OR ("Class-based" AND Chemical))

Appendix B. Generative Artificial Intelligence (GenAI) for Full Text Screening

B.1. Prompts Used for Full Text Screening

A set of 9 prompts were developed to categorize 10 expert curated PDFs according to topic. Subject matter experts graded these prompts for accuracy against the seed studies and confirmed they were sufficient for this subject matter topic. These 9 prompts were used to review the 545 peer reviewed papers identified as containing potential approaches to performing class-based risk assessment.

Table B-1 provides an overview of the 9 prompts that were used for testing. Table B-2 lists the various parameters used for the GenAI.

Table B-1. Final Prompt Questions

Variable Name	Prompt
Exposure	Does this paper estimate exposure, dose, or intake for a chemical with no to limited data, or provide a method for doing so?
Physicochemical	Does this paper use physicochemical properties (e.g., molecular weight, Kow, Koa, Henry's law constant, vapor pressure, solubility) to estimate exposure, dose, or intake; or provide a method for doing so?
Material Type	Does this paper extrapolate using material type to estimate exposure, dose, or intake?
Hazard Property	Does the paper discuss an approach for grouping chemicals based on hazard property or adverse effect or toxic effect or apical effect or target organ or target organ system?
Toxicokinetics	Does the paper discuss an approach for grouping chemicals based on toxicokinetics or how the body absorbs, distributes, metabolizes, and eliminates specific chemicals?
Toicodynamics	Does the paper discuss an approach for grouping chemicals based toxicodynamics or mode of action or mechanism of action or adverse outcome pathway or key characteristic of specific chemicals?
Weight of Evidence	Does the paper describe the weight of evidence, considering the amount, type (i.e., human or animal), and quality of studies?
Risk Estimate	Does the paper discuss a framework for assessing a risk estimate or margin of exposure or risk characterization ratio or hazard quotient or hazard index or margin of safety?
Limitations	Does the paper discuss a framework for assessing the limitations or assumptions or uncertainties of a risk estimate or margin of exposure or risk characterization ratio or hazard quotient or hazard index or margin of safety?

Table B-2. Final Parameters Used for GenAI Full Text Screening

Parameter	Final Details
GenAI	Claude version 3.0
Preamble	Please act as a toxicologist. Wrap each response in an XML tag based on the prefix 'response' concatenated with the question number.
Post-amble	If yes, justify your response with quotes. If you don't know, say "I don't know."

Appendix C. References

Ahlers, J., Nendza, M., & Schwartz, D. (2019). Environmental hazard and risk assessment of thiochemicals. Application of integrated testing and intelligent assessment strategies (ITS) to fulfil the REACH requirements for aquatic toxicity. *Chemosphere*, *214*, 480-490. https://doi.org/10.1016/j.chemosphere.2018.09.082

Al-Malahmeh, A. J., Al-Ajlouni, A., Wesseling, S., Soffers, A. E., Al-Subeihi, A., Kiwamoto, R., Vervoort, J., & Rietjens, I. M. (2017). Physiologically based kinetic modeling of the bioactivation of myristicin. *Arch Toxicol*, *91*(2), 713-734. https://doi.org/10.1007/s00204-016-1752-5

Alexander-White, C., Bury, D., Cronin, M., Dent, M., Hack, E., Hewitt, N. J., Kenna, G., Naciff, J., Ouedraogo, G., Schepky, A., Mahony, C., & Europe, C. (2022). A 10-step framework for use of read-across (RAX) in next generation risk assessment (NGRA) for cosmetics safety assessment. *Regul Toxicol Pharmacol*, *129*, 105094. https://doi.org/10.1016/j.yrtph.2021.105094

Almonacid Garrido, M. C., Villanueva-Suárez, M. J., Álamo Gómez, A., Garcia-Alonso, A., & Tenorio Sanz, M. D. (2024). Trihalomethane Levels in Municipal Drinking Water of Madrid (Spain) and Health Risk Assessment for Oral Exposure. *Exposure and Health*, *16*(4), 911-924. https://doi.org/10.1007/s12403-023-00599-0

Aurisano, N., Fantke, P., Huang, L., & Jolliet, O. (2022). Estimating mouthing exposure to chemicals in children's products. *J Expo Sci Environ Epidemiol*, *32*(1), 94-102. https://doi.org/10.1038/s41370-021-00354-0

Ball, N., Cronin, M. T. D., Shen, J., Blackburn, K., Booth, E. D., Bouhifd, M., Donley, E., Egnash, L., Hastings, C., Juberg, D. R., Kleensang, A., Kleinstreuer, N., Kroese, E. D., Lee, A. C., Luechtefeld, T., Maertens, A., Marty, S., Naciff, J. M., Palmer, J.,…Hartung, T. (2016). Toward Good Read-Across Practice (GRAP) guidance. *ALTEX - Alternatives to animal experimentation*, *33*(2), 149-166. https://doi.org/10.14573/altex.1601251

Bennekou, S. H. (2019). Moving towards a holistic approach for human health risk assessment - Is the current approach fit for purpose? *EFSA J*, *17*(Suppl 1), e170711. https://doi.org/10.2903/j.efsa.2019.e170711

Bil, W., Zeilmaker, M., Fragki, S., Lijzen, J., Verbruggen, E., & Bokkers, B. (2021). Risk Assessment of Per- and Polyfluoroalkyl Substance Mixtures: A Relative Potency Factor Approach. *Environ Toxicol Chem*, *40*(3), 859-870. https://doi.org/10.1002/etc.4835

Bopp, S. K., Barouki, R., Brack, W., Dalla Costa, S., Dorne, J. C. M., Drakvik, P. E., Faust, M., Karjalainen, T. K., Kephalopoulos, S., van Klaveren, J., Kolossa-Gehring, M., Kortenkamp, A., Lebret, E., Lettieri, T., Norager, S., Ruegg, J., Tarazona, J. V., Trier, X., van de Water, B.,…Bergman, A. (2018). Current EU research activities on combined exposure to multiple chemicals. *Environ Int*, *120*, 544-562. https://doi.org/10.1016/j.envint.2018.07.037

Brock, W. J., Rodricks, J. V., Rulis, A., Dellarco, V. L., Gray, G. M., & Lane, R. W. (2003). Food Safety: Risk Assessment Methodology and Decision-Making Criteria. *International Journal of Toxicology*, *22*(6), 435-451. https://doi.org/10.1177/109158180302200605

Bus, J. S., Banton, M. I., Faber, W. D., Kirman, C. R., McGregor, D. B., & Pourreau, D. B. (2015). Human health screening level risk assessments of tertiary-butyl acetate (TBAC): calculated acute and chronic reference concentration (RfC) and Hazard Quotient (HQ) values based on toxicity and exposure scenario evaluations. *Crit Rev Toxicol*, *45*(2), 142-171. https://doi.org/10.3109/10408444.2014.980884

Cattaneo, I., Kalian, A. D., Di Nicola, M. R., Dujardin, B., Levorato, S., Mohimont, L., Nathanail, A. V., Carnessechi, E., Astuto, M. C., Tarazona, J. V., Kass, G. E. N., Liem, A. K. D.,

Robinson, T., Manini, P., Hogstrand, C., Price, P. S., & Dorne, J. (2023). Risk Assessment of Combined Exposure to Multiple Chemicals at the European Food Safety Authority: Principles, Guidance Documents, Applications and Future Challenges. *Toxins (Basel)*, *15*(1). https://doi.org/10.3390/toxins15010040

Clewell, R. A., Leonard, J. A., Nicolas, C. I., Campbell, J. L., Yoon, M., Efremenko, A. Y., McMullen, P. D., Andersen, M. E., Clewell, H. J., 3rd, Phillips, K. A., & Tan, Y. M. (2020). Application of a combined aggregate exposure pathway and adverse outcome pathway (AEP-AOP) approach to inform a cumulative risk assessment: A case study with phthalates. *Toxicol In Vitro*, *66*, 104855. https://doi.org/10.1016/j.tiv.2020.104855

U.S. Consumer Product Safety Commission. (CPSC). (2010). *Phthalates*. https://www.cpsc.gov/Regulations-Laws--Standards/Voluntary-Standards/Topics/Phthalates

U.S. Consumer Product Safety Commission. (CPSC). (2014). *Chronic hazard advisory panel on phthalates and phthalate alternatives*. https://web.archive.org/web/20170202160318/https://www.cpsc.gov/s3fs-public/CHAP-REPORT-With-Appendices.pdf

Crofton, K. M., Bassan, A., Behl, M., Chushak, Y. G., Fritsche, E., Gearhart, J. M., Marty, M. S., Mumtaz, M., Pavan, M., Ruiz, P., Sachana, M., Selvam, R., Shafer, T. J., Stavitskaya, L., Szabo, D. T., Szabo, S. T., Tice, R. R., Wilson, D., Woolley, D., & Myatt, G. J. (2022). Current status and future directions for a neurotoxicity hazard assessment framework that integrates in silico approaches. *Comput Toxicol*, *22*. https://doi.org/10.1016/j.comtox.2022.100223

Crump, K. S., Chen, C., Chiu, W. A., Louis, T. A., Portier, C. J., Subramaniam, R. P., & White, P. D. (2010). What role for biologically based dose-response models in estimating low-dose risk? *Environ Health Perspect*, *118*(5), 585-588. https://doi.org/10.1289/ehp.0901249

Cui, D., Cox, J., Mejias, E., Ng, B., Gardinali, P., Bagner, D. M., & Quinete, N. (2023). Evaluating non-targeted analysis methods for chemical characterization of organic contaminants in different matrices to estimate children's exposure. *J Expo Sci Environ Epidemiol*, *33*(4), 589-601. https://doi.org/10.1038/s41370-023-00547-9

Czaja, K., Strucinski, P., Korcz, W., Minorczyk, M., Hernik, A., & Wiadrowska, B. (2020). Alternative toxicological methods for establishing residue definitions applied for dietary risk assessment of pesticides in the European Union. *Food Chem Toxicol*, *137*, 111120. https://doi.org/10.1016/j.fct.2020.111120

Danforth, C., Chiu, W. A., Rusyn, I., Schultz, K., Bolden, A., Kwiatkowski, C., & Craft, E. (2020). An integrative method for identification and prioritization of constituents of concern in produced water from onshore oil and gas extraction. *Environ Int*, *134*, 105280. https://doi.org/10.1016/j.envint.2019.105280

Danzeisen, R., Weight, D., Blakeney, M., & Boyle, D. (2022). A tiered approach to investigate the inhalation toxicity of cobalt substances. Introduction: Cobalt's essential role in nature and technology. *Regul Toxicol Pharmacol*, *130*, 105125. https://doi.org/10.1016/j.yrtph.2022.105125

De Abrew, K. N., Shan, Y. K., Wang, X., Krailler, J. M., Kainkaryam, R. M., Lester, C. C., Settivari, R. S., LeBaron, M. J., Naciff, J. M., & Daston, G. P. (2019). Use of connectivity mapping to support read across: A deeper dive using data from 186 chemicals, 19 cell lines and 2 case studies. *Toxicology*, *423*, 84-94. https://doi.org/10.1016/j.tox.2019.05.008

Delmaar, C., & Meesters, J. (2020). Modeling consumer exposure to spray products: an evaluation of the ConsExpo Web and ConsExpo nano models with experimental data. *J Expo Sci Environ Epidemiol*, *30*(5), 878-887. https://doi.org/10.1038/s41370-020-0239-x

Delmaar, C. J. E., Schreurs, R., Bakker, M. I., Minnema, J., & Bokkers, B. G. H. (2023). PACEMweb: a tool for aggregate consumer exposure assessment. *J Expo Sci Environ Epidemiol*, *33*(6), 971-979. https://doi.org/10.1038/s41370-022-00509-7

Diamond, J., Altenburger, R., Coors, A., Dyer, S. D., Focazio, M., Kidd, K., Koelmans, A. A., Leung, K. M. Y., Servos, M. R., Snape, J., Tolls, J., & Zhang, X. (2018). Use of prospective and retrospective risk assessment methods that simplify chemical mixtures associated with treated domestic wastewater discharges. *Environ Toxicol Chem*, *37*(3), 690-702. https://doi.org/10.1002/etc.4013

Dourson, M., Becker, R. A., Haber, L. T., Pottenger, L. H., Bredfeldt, T., & Fenner-Crisp, P. A. (2013). Advancing human health risk assessment: integrating recent advisory committee recommendations. *Crit Rev Toxicol*, *43*(6), 467-492. https://doi.org/10.3109/10408444.2013.807223

Dourson, M., Ewart, L., Fitzpatrick, S. C., Barros, S. B. M., Mahadevan, B., & Hayes, A. W. (2022). The Future of Uncertainty Factors With In Vitro Studies Using Human Cells. *Toxicol Sci*, *186*(1), 12-17. https://doi.org/10.1093/toxsci/kfab134

Du, S., Rodenburg, L., Patterson, N., Chu, C., Riker, C. D., Yu, C. H., & Fan, Z. (2020). Concentration of polychlorinated biphenyls in serum from New Jersey biomonitoring study: 2016–2018. *Chemosphere*, *261*, 127730. https://doi.org/https://doi.org/10.1016/j.chemosphere.2020.127730

European Chemicals Agency. (ECHA). (2017). *Read-Across Assessment Framework (RAAF)* (ECHA-17-R-01-EN). https://echa.europa.eu/documents/10162/17221/raaf_en.pdf/614e5d61-891d-4154-8a47-87efebd1851a

EFSA Scientific Committee, More, S. J., Bampidis, V., Benford, D., Bragard, C., Halldorsson, T. I., Hernández-Jerez, A. F., Hougaard Bennekou, S., Koutsoumanis, K. P., Machera, K., Naegeli, H., Nielsen, S. S., Schlatter, J. R., Schrenk, D., Silano, V., Turck, D., Younes, M., Gundert-Remy, U., Kass, G. E. N.,…Wallace, H. M. (2019). Guidance on the use of the Threshold of Toxicological Concern approach in food safety assessment. *EFSA Journal*, *17*(6), e05708. https://doi.org/https://doi.org/10.2903/j.efsa.2019.5708

Enoch, S. J., Hasarova, Z., Cronin, M. T. D., Bridgwood, K., Rao, S., Kluxen, F. M., & Frericks, M. (2022a). Sub-structure-based category formation for the prioritisation of genotoxicity hazard assessment for pesticide residues: Sulphonyl ureas. *Regul Toxicol Pharmacol*, *129*, 105115. https://doi.org/10.1016/j.yrtph.2022.105115

Enoch, S. J., Hasarova, Z., Cronin, M. T. D., Bridgwood, K., Rao, S., Kluxen, F. M., & Frericks, M. (2023). Metabolism-based category formation for the prioritisation of genotoxicity hazard assessment for plant protection product residues (part 3): Strobilurins. *Regul Toxicol Pharmacol*, *144*, 105484. https://doi.org/10.1016/j.yrtph.2023.105484

Enoch, S. J., Hasarova, Z., Cronin, M. T. D., & Frericks, M. (2022b). Sub-structure-based category formation for the prioritisation of genotoxicity hazard assessment for pesticide residues (part 2): Triazoles. *Regul Toxicol Pharmacol*, *134*, 105237. https://doi.org/10.1016/j.yrtph.2022.105237

Ernstoff, A. S., Fantke, P., Huang, L., & Jolliet, O. (2017). High-throughput migration modelling for estimating exposure to chemicals in food packaging in screening and prioritization tools. *Food Chem Toxicol*, *109*(Pt 1), 428-438. https://doi.org/10.1016/j.fct.2017.09.024

Fang, H., Tong, W., Shi, L. M., Blair, R., Perkins, R., Branham, W., Hass, B. S., Xie, Q., Dial, S. L., Moland, C. L., & Sheehan, D. M. (2001). Structure-activity relationships for a large diverse set of natural, synthetic, and environmental estrogens. *Chem Res Toxicol*, *14*(3), 280-294. https://doi.org/10.1021/tx000208y

Fiedler, H., Sadia, M., Baabish, A., & Sobhanei, S. (2022). Perfluoroalkane substances in national samples from global monitoring plan projects (2017–2019). *Chemosphere*, *307*, 136038. https://doi.org/https://doi.org/10.1016/j.chemosphere.2022.136038

Finckh, S., Beckers, L. M., Busch, W., Carmona, E., Dulio, V., Kramer, L., Krauss, M., Posthuma, L., Schulze, T., Slootweg, J., Von der Ohe, P. C., & Brack, W. (2022). A risk based assessment approach for chemical mixtures from wastewater treatment plant effluents. *Environ Int*, *164*, 107234. https://doi.org/10.1016/j.envint.2022.107234

Genualdi, S., Beekman, J., Carlos, K., Fisher, C. M., Young, W., DeJager, L., & Begley, T. (2022). Analysis of per- and poly-fluoroalkyl substances (PFAS) in processed foods from FDA's Total Diet Study. *Analytical and Bioanalytical Chemistry*, *414*(3), 1189-1199. https://doi.org/10.1007/s00216-021-03610-2

Grimm, F. A., Iwata, Y., Sirenko, O., Chappell, G. A., Wright, F. A., Reif, D. M., Braisted, J., Gerhold, D. L., Yeakley, J. M., Shepard, P., Seligmann, B., Roy, T., Boogaard, P. J., Ketelslegers, H. B., Rohde, A. M., & Rusyn, I. (2016). A chemical-biological similarity-based grouping of complex substances as a prototype approach for evaluating chemical alternatives. *Green Chem*, *18*(16), 4407-4419. https://doi.org/10.1039/c6gc01147k

Health Canada. (Health Canada). (2020). *Screening assessment - Phthalate substance grouping* (En14-393/2019E-PDF). https://www.canada.ca/en/environment-climate-change/services/evaluating-existing-substances/screening-assessment-phthalate-substance-grouping.html

Hines, D. E., Edwards, S. W., Conolly, R. B., & Jarabek, A. M. (2018). A Case Study Application of the Aggregate Exposure Pathway (AEP) and Adverse Outcome Pathway (AOP) Frameworks to Facilitate the Integration of Human Health and Ecological End Points for Cumulative Risk Assessment (CRA). *Environ Sci Technol*, *52*(2), 839-849. https://doi.org/10.1021/acs.est.7b04940

Interagency Coordinating Committee on the Validation of Alternative Methods. (ICCVAM). (2024). *Validation, Qualification, and Regulatory Acceptance of New Approach Methodologies*. https://ntp.niehs.nih.gov/whatwestudy/niceatm/resources-for-test-method-developers/submissions

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

Isaacs, K. K., Wall, J. T., Williams, A. R., Hobbie, K. A., Sobus, J. R., Ulrich, E., Lyons, D., Dionisio, K. L., Williams, A. J., Grulke, C., Foster, C. A., McCoy, J., & Bevington, C. (2022). A harmonized chemical monitoring database for support of exposure assessments. *Sci Data*, *9*(1), 314. https://doi.org/10.1038/s41597-022-01365-8

Lai, A., Clark, A. M., Escher, B. I., Fernandez, M., McEwen, L. R., Tian, Z., Wang, Z., & Schymanski, E. L. (2022). The Next Frontier of Environmental Unknowns: Substances of Unknown or Variable Composition, Complex Reaction Products, or Biological Materials (UVCBs). *Environ Sci Technol*, *56*(12), 7448-7466. https://doi.org/10.1021/acs.est.2c00321

Li, S., Yu, Y., Bian, X., Yao, L., Li, M., Lou, Y. R., Yuan, J., Lin, H. S., Liu, L., Han, B., & Xiang, X. (2021). Prediction of oral hepatotoxic dose of natural products derived from traditional Chinese medicines based on SVM classifier and PBPK modeling. *Arch Toxicol*, *95*(5), 1683-1701. https://doi.org/10.1007/s00204-021-03023-1

Liang, Y., Liu, X., & Allen, M. R. (2019). The influence of temperature on the emissions of organophosphate ester flame retardants from polyisocyanurate foam: Measurement and modeling. *Chemosphere*, *233*, 347-354. https://doi.org/10.1016/j.chemosphere.2019.05.289

Liu, X., Zhang, H., Pan, W., Xue, Q., Fu, J., Liu, G., Zheng, M., & Zhang, A. (2019). A novel computational solution to the health risk assessment of air pollution via joint toxicity prediction: A case study on selected PAH binary mixtures in particulate matters. *Ecotoxicol Environ Saf*, *170*, 427-435. https://doi.org/10.1016/j.ecoenv.2018.12.010

Lopez, A., Yusa, V., Munoz, A., Vera, T., Borras, E., Rodenas, M., & Coscolla, C. (2017). Risk assessment of airborne pesticides in a Mediterranean region of Spain. *Sci Total Environ*, *574*, 724-734. https://doi.org/10.1016/j.scitotenv.2016.08.149

Lutz, W. K., Lutz, R. W., Gaylor, D. W., & Conolly, R. B. (2014). Dose–Response Relationship and Extrapolation in Toxicology. Mechanistic and Statistical Considerations. In F.-X. Reichl & M. Schwenk (Eds.), *Regulatory Toxicology* (pp. 547-568). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35374-1_72

Maffini, M. V., Rayasam, S. D. G., Axelrad, D. A., Birnbaum, L. S., Cooper, C., Franjevic, S., MacRoy, P. M., Nachman, K. E., Patisaul, H. B., Rodgers, K. M., Rossi, M. S., Schettler, T., Solomon, G. M., & Woodruff, T. J. (2023). Advancing the science on chemical classes. *Environ Health*, *21*(Suppl 1), 120. https://doi.org/10.1186/s12940-022-00919-y

Malich, G., Markovic, B., & Winder, C. (1998). Human cell line toxicity of binary and ternary chemical mixtures in comparison to individual toxic effects of their components. *Arch Environ Contam Toxicol*, *35*(3), 370-376. https://doi.org/10.1007/s002449900390

McCord, J. P., Groff, L. C., 2nd, & Sobus, J. R. (2022). Quantitative non-targeted analysis: Bridging the gap between contaminant discovery and risk characterization. *Environ Int*, *158*, 107011. https://doi.org/10.1016/j.envint.2021.107011

Minkus, S., Bieber, S., & Letzel, T. (2022). Spotlight on mass spectrometric non-target screening analysis: Advanced data processing methods recently communicated for extracting, prioritizing and quantifying features. *Anal Sci Adv*, *3*(3-4), 103-112. https://doi.org/10.1002/ansa.202200001

Moser, V. C., Aschner, M., Richardson, J. R., Bowman, A. B., & Richardson, R. J. (2019). Toxic Responses of the Nervous System. In C. D. Klaassen (Ed.), *Casarett & Doull's Toxicology: The Basic Science of Poisons, 9th edition*. McGraw-Hill Education. accesspharmacy.mhmedical.com/content.aspx?aid=1158499879

National Academies of Sciences, Engineering, and Medicine. (NASEM). (2019). *A class approach to hazard assessment of organohalogen flame retardants*. http://dx.doi.org/10.17226/25412

Olesti, E., Gonzalez-Ruiz, V., Wilks, M. F., Boccard, J., & Rudaz, S. (2021). Approaches in metabolomics for regulatory toxicology applications. *Analyst*, *146*(6), 1820-1834. https://doi.org/10.1039/d0an02212h

Paini, A., Worth, A., Kulkarni, S., Ebbrell, D., & Madden, J. (2021). Assessment of the predictive capacity of a physiologically based kinetic model using a read-across approach. *Comput Toxicol*, *18*, 100159. https://doi.org/10.1016/j.comtox.2021.100159

Patlewicz, G., Cronin, M. T. D., Helman, G., Lambert, J. C., Lizarraga, L. E., & Shah, I. (2018). Navigating through the minefield of read-across frameworks: A commentary perspective. *Computational Toxicology*, *6*, 39-54. https://doi.org/https://doi.org/10.1016/j.comtox.2018.04.002

Pitzer, E. M., Shafer, T. J., & Herr, D. W. (2023). Identification of neurotoxicology (NT)/developmental neurotoxicology (DNT) adverse outcome pathways and key event linkages with in vitro DNT screening assays. *Neurotoxicology*, *99*, 184-194. https://doi.org/10.1016/j.neuro.2023.10.007

Reardon, A. J. F., Rowan-Carroll, A., Ferguson, S. S., Leingartner, K., Gagne, R., Kuo, B., Williams, A., Lorusso, L., Bourdon-Lacombe, J. A., Carrier, R., Moffat, I., Yauk, C. L., & Atlas, E. (2021). Potency Ranking of Per- and Polyfluoroalkyl Substances Using High-Throughput Transcriptomic Analysis of Human Liver Spheroids. *Toxicol Sci*, *184*(1), 154-169. https://doi.org/10.1093/toxsci/kfab102

Redman, A. D., Parkerton, T. F., Comber, M. H., Paumen, M. L., Eadsforth, C. V., Dmytrasz, B., King, D., Warren, C. S., den Haan, K., & Djemel, N. (2014). PETRORISK: a risk assessment framework for petroleum substances. *Integr Environ Assess Manag*, *10*(3), 437-448. https://doi.org/10.1002/ieam.1536

Reffstrup, T. K., Larsen, J. C., & Meyer, O. (2010). Risk assessment of mixtures of pesticides. Current approaches and future strategies. *Regul Toxicol Pharmacol*, *56*(2), 174-192. https://doi.org/10.1016/j.yrtph.2009.09.013

Rice, G., MacDonell, M., Hertzberg, R. C., Teuschler, L., Picel, K., Butler, J., Chang, Y. S., & Hartmann, H. (2008). An approach for assessing human exposures to chemical mixtures in the environment. *Toxicol Appl Pharmacol*, *233*(1), 126-136. https://doi.org/10.1016/j.taap.2008.05.004

Rodriguez, M. J., Vinette, Y., Sérodes, J.-B., & Bouchard, C. (2003). Trihalomethanes in Drinking Water of Greater Québec Region (Canada): Occurrence, Variations and Modelling. *Environmental Monitoring and Assessment*, *89*(1), 69-93. https://doi.org/10.1023/A:1025811921502

Rostkowski, P., Haglund, P., Aalizadeh, R., Alygizakis, N., Thomaidis, N., Arandes, J. B., Nizzetto, P. B., Booij, P., Budzinski, H., Brunswick, P., Covaci, A., Gallampois, C., Grosse, S., Hindle, R., Ipolyi, I., Jobst, K., Kaserzon, S. L., Leonards, P., Lestremau, F.,…Yang, C. (2019). The strength in numbers: comprehensive characterization of house dust using complementary mass spectrometric techniques. *Anal Bioanal Chem*, *411*(10), 1957-1977. https://doi.org/10.1007/s00216-019-01615-6

Rowan-Carroll, A., Reardon, A., Leingartner, K., Gagne, R., Williams, A., Meier, M. J., Kuo, B., Bourdon-Lacombe, J., Moffat, I., Carrier, R., Nong, A., Lorusso, L., Ferguson, S. S., Atlas, E., & Yauk, C. (2021). High-Throughput Transcriptomic Analysis of Human Primary Hepatocyte Spheroids Exposed to Per- and Polyfluoroalkyl Substances as a Platform for Relative Potency Characterization. *Toxicol Sci*, *181*(2), 199-214. https://doi.org/10.1093/toxsci/kfab039

Safe, S. (1997). Limitations of the toxic equivalency factor approach for risk assessment of TCDD and related compounds. *Teratog Carcinog Mutagen*, *17*(4-5), 285-304. https://www.ncbi.nlm.nih.gov/pubmed/9508738

Saktrakulkla, P., Lan, T., Hua, J., Marek, R. F., Thorne, P. S., & Hornbuckle, K. C. (2020). Polychlorinated Biphenyls in Food. *Environmental Science & Technology*, *54*(18), 11443-11452. https://doi.org/10.1021/acs.est.0c03632

Salvito, D., Fernandez, M., Jenner, K., Lyon, D. Y., de Knecht, J., Mayer, P., MacLeod, M., Eisenreich, K., Leonards, P., Cesnaitis, R., Leon-Paumen, M., Embry, M., & Deglin, S. E. (2020). Improving the Environmental Risk Assessment of Substances of Unknown or Variable Composition, Complex Reaction Products, or Biological Materials. *Environ Toxicol Chem*, *39*(11), 2097-2108. https://doi.org/10.1002/etc.4846

Schor, J., Scheibe, P., Bernt, M., Busch, W., Lai, C., & Hackermuller, J. (2022). AI for predicting chemical-effect associations at the chemical universe level-deepFPlearn. *Brief Bioinform*, *23*(5). https://doi.org/10.1093/bib/bbac257

Schultz, T. W., Richarz, A.-N., & Cronin, M. T. D. (2019). Assessing uncertainty in read-across: Questions to evaluate toxicity predictions based on knowledge gained from case studies. *Computational Toxicology*, *9*, 1-11. https://doi.org/https://doi.org/10.1016/j.comtox.2018.10.003

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., & Hollender, J. (2014). Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol*, *48*(4), 2097-2098. https://doi.org/10.1021/es5002105

Smith, R. L., Cohen, S. M., Doull, J., Feron, V. J., Goodman, J. I., Marnett, L. J., Portoghese, P. S., Waddell, W. J., Wagner, B. M., Hall, R. L., Higley, N. A., Lucas-Gavin, C., & Adams, T. B. (2005). A procedure for the safety evaluation of natural flavor complexes used as ingredients in food: essential oils. *Food and Chemical Toxicology*, *43*(3), 345-363. https://doi.org/https://doi.org/10.1016/j.fct.2004.11.007

U.S. Environmental Protection Agency. (U.S. EPA). (2000). *Science policy council handbook: Risk characterization handbook* [EPA Report](EPA/100/B-00/002). https://www.epa.gov/risk/risk-characterization-handbook
U.S. Environmental Protection Agency. (U.S. EPA). (2002). *Guidance on cumulative risk assessment of pesticide chemicals that have a common mechanism of toxicity* [EPA Report].
U.S. Environmental Protection Agency. (U.S. EPA). (2006). *Organophosphorus cumulative risk assessment. 2006 Update. Continuation*. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P1010KM6.txt
https://archive.epa.gov/epawaste/hazard/wastemin/web/pdf/pahs.pdf
U.S. Environmental Protection Agency. (U.S. EPA). (2011). *Pyrethrins/pyrethroid cumulative risk assessment*. https://www.regulations.gov/document/EPA-HQ-OPP-2011-0746-0019
U.S. Environmental Protection Agency. (U.S. EPA). (2012). *Assessing and managing chemicals under TSCA: Phthalates*. https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/phthalates
U.S. Environmental Protection Agency. (U.S. EPA). (2019). *Guidelines for human exposure assessment* [EPA Report](EPA/100/B-19/001). https://www.epa.gov/sites/production/files/2020-01/documents/guidelines_for_human_exposure_assessment_final2019.pdf
U.S. Environmental Protection Agency. (U.S. EPA). (2023). *Draft Proposed Approach for Cumulative Risk Assessment of High-Priority Phthalates and a Manufacturer-Requested Phthalate under the Toxic Substances Control Act* (EPA-740-P-23-002). https://www.regulations.gov/document/EPA-HQ-OPPT-2022-0918-0009
U.S. Environmental Protection Agency. (U.S. EPA). (2024). *Conducting a human health risk assessment*. https://www.epa.gov/risk/conducting-human-health-risk-assessment
Ulrich, H., Macherius, A., Kunkel, U., Sengl, M., & Letzel, T. (2024). Novel PFAS-specific monitoring approach for highly impacted surface waters. *Chemosphere*, *349*, 140893. https://doi.org/https://doi.org/10.1016/j.chemosphere.2023.140893
van den Berg, S. J., Punt, A., Soffers, A. E., Vervoort, J., Ngeleja, S., Spenkelink, B., & Rietjens, I. M. (2012). Physiologically based kinetic models for the alkenylbenzene elemicin in rat and human and possible implications for risk assessment. *Chem Res Toxicol*, *25*(11), 2352-2367. https://doi.org/10.1021/tx300239z
Verdon, R., Stone, V., Murphy, F., Christopher, E., Johnston, H., Doak, S., Vogel, U., Haase, A., & Kermanizadeh, A. (2022). The application of existing genotoxicity methodologies for grouping of nanomaterials: towards an integrated approach to testing and assessment. *Part Fibre Toxicol*, *19*(1), 32. https://doi.org/10.1186/s12989-022-00476-9
Wambaugh, J. F., Setzer, R. W., Reif, D. M., Gangwal, S., Mitchell-Blackwood, J., Arnot, J. A., Joliet, O., Frame, A., Rabinowitz, J., Knudsen, T. B., Judson, R. S., Egeghy, P., Vallero, D., & Cohen Hubal, E. A. (2013). High-throughput models for exposure-based chemical prioritization in the ExpoCast project. *Environ Sci Technol*, *47*(15), 8479-8488. https://doi.org/10.1021/es400482g
Wang, C., Eichler, C. M. A., Bi, C., Delmaar, C. J. E., Xu, Y., & Little, J. C. (2023). A rapid micro chamber method to measure SVOC emission and transport model parameters. *Environ Sci Process Impacts*, *25*(4), 818-831. https://doi.org/10.1039/d2em00507g
Woo, Y.-T., Lai, D., McLain, J. L., Manibusan, M. K., & Dellarco, V. (2002). Use of mechanism-based structure-activity relationships analysis in carcinogenic potential ranking for drinking water disinfection by-products. *Environmental Health Perspectives*, *110*(suppl 1), 75-87. https://doi.org/doi:10.1289/ehp.02110s175
Zhao, H., Yang, L., Li, Y., Xue, W., Li, K., Xie, Y., Meng, S., & Cao, G. (2020). Environmental occurrence and risk assessment of haloacetic acids in swimming pool water and drinking

water [10.1039/D0RA02389B]. *RSC Advances*, *10*(47), 28267-28276. https://doi.org/10.1039/D0RA02389B

Zhao, Y., Hallar, A. G., & Mazzoleni, L. R. (2013). Atmospheric organic matter in clouds: exact masses and molecular formula identification using ultrahigh-resolution FT-ICR mass spectrometry. *Atmospheric Chemistry and Physics*, *13*(24), 12343-12362. https://doi.org/10.5194/acp-13-12343-2013