United States
Consumer Product Safety Commission

July 22, 2025

CPSC Staff Statement on: Guidance Document for Conducting Class-Based Risk Assessment

The U.S. Consumer Product Safety Commission (CPSC or Commission) contracted with ICF (Contract BPA No. 61320622A0005, Order No. 61320623F2030) to identify approaches to class-based risk assessment and to complete a guide that can be used to conduct class-based risk assessments. This statement was prepared by the CPSC staff. ICF produced the accompanying report: "Guidance Document for Conducting Class-Based Risk Assessment" (the Risk Assessment Guide), also known as Subtask 4B in the contract. The statement and report have not been reviewed or approved by, and may not represent the views of, the Commission.

The purpose of the Risk Assessment Guide is to outline technical guidance for developing quantitative risk value estimates for chemicals belonging to a chemical class or subclass. The Risk Assessment Guide describes a method that can be used to provide estimates both for cancer and noncancer outcomes. Before deploying this approach, it is necessary to perform class-based dose-response assessment and class-based exposure assessment as described, respectively, in the Subtask 4A and in reports from an earlier call order.[1]

The Risk Assessment Guide explains probabilistic risk assessment methods that can be incorporated with more traditional risk assessment approaches to produce an assessment of risk within a population. The approaches described in the Risk Assessment Guide are flexible based on data availability, and new data can be integrated into assessments to improve risk estimates.

The document also provides lessons learned from performing the class-based risk assessment for the polyhalogenated organophosphate (PHOP) organohalogen flame retardants (OFRs) subclass in Subtask 3B of this contract.

[1] The Risk Assessment Guide cites supporting documents from other subtasks of Contract BPA No. 61320622A0005, Order No. 61320623F2030 and from Contract BPA No. 61320622A0005, Order No. 61320622F2012 (Guidance Document for Conducting Class-based Exposure Assessments for Organohalogen Flame Retardants) all of which can be found at https://www.cpsc.gov/Business--Manufacturing/Organohalogen-Flame-Retardant-Chemicals-Assessment.

Guidance Document for Conducting Class-based Risk Assessment

Call Order No. 61320623F2030
CPSC BPA No. 61320622A0005
February 5, 2025

Submitted To:
U.S. Consumer Product Safety Commission
5 Research Place
Rockville, MD 20850

Submitted By:
ICF
1902 Reston Metro Plaza
Reston, VA 20190 USA

Table of Contents

Tables

Figures

1. Introduction

Under CO-9, Subtask 4B seeks to outline guidance to apply class-based risk assessment analyses to develop quantitative risk value estimates for chemicals belonging to a single chemical class. Here we describe the process by which probabilistic risk assessment methods can be incorporated with more traditional risk assessment metrics, such as a hazard quotient (HQ), to produce a relatively easy to interpret assessment of risk within a population. Additionally, we note lessons learned from performing class-based risk assessment for the polyhalogenated organophosphate (PHOP) organohalogen flame retardants (OFRs) subclass.

2. Risk Assessment

Human health risk assessment is the process of characterizing the impacts to human health following exposures to given chemical(s) and source(s). 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/characterization describes the types of health effects caused by a given chemical 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 at a given level of exposure). Exposure assessment describes the magnitude, frequency and duration of exposure, and the size and characteristics of the exposed population. All of these evidence streams feed into risk characterization, which includes both a quantitative description of the risk(s) associated with a exposure(s), and a quantitative or qualitative description of the associated uncertainties.

Risk assessment for single chemicals is conducted using different methods for threshold endpoints (i.e., noncancer effects and cancer effects thought to be induced by a non-mutagenic MOA) and non-threshold effects (i.e., cancer effects thought to be induced by a mutagenic MOA or for which the MOA is insufficiently characterized).

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 (UF) 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 EPA, also include factors for database uncertainty (UFD) and extrapolation from subchronic to chronic duration (UFS). Different organizations may use different sets of UF. 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 UF approach used, risk characterization is then conducted by comparing the resulting TRV to a measured or estimated level of real-world exposure, resulting in a HQ.

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

Generally, HQ < 1 indicates adverse effects are unlikely, while HQ > 1 indicates potential for adverse effects, though HQ does not specifically quantify the magnitude of this potential.

Another way of characterizing risk is the margin of exposure (MOE) approach where a POD is divided directly by a measured or estimated real-world exposure. The MOE is then compared to a minimum margin of safety. This minimum margin of safety considers factors similar to the UF 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 lifetime average daily dose (LADD) to estimate lifetime excess risk (LER):

$$LER = LADD \times CSF$$

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

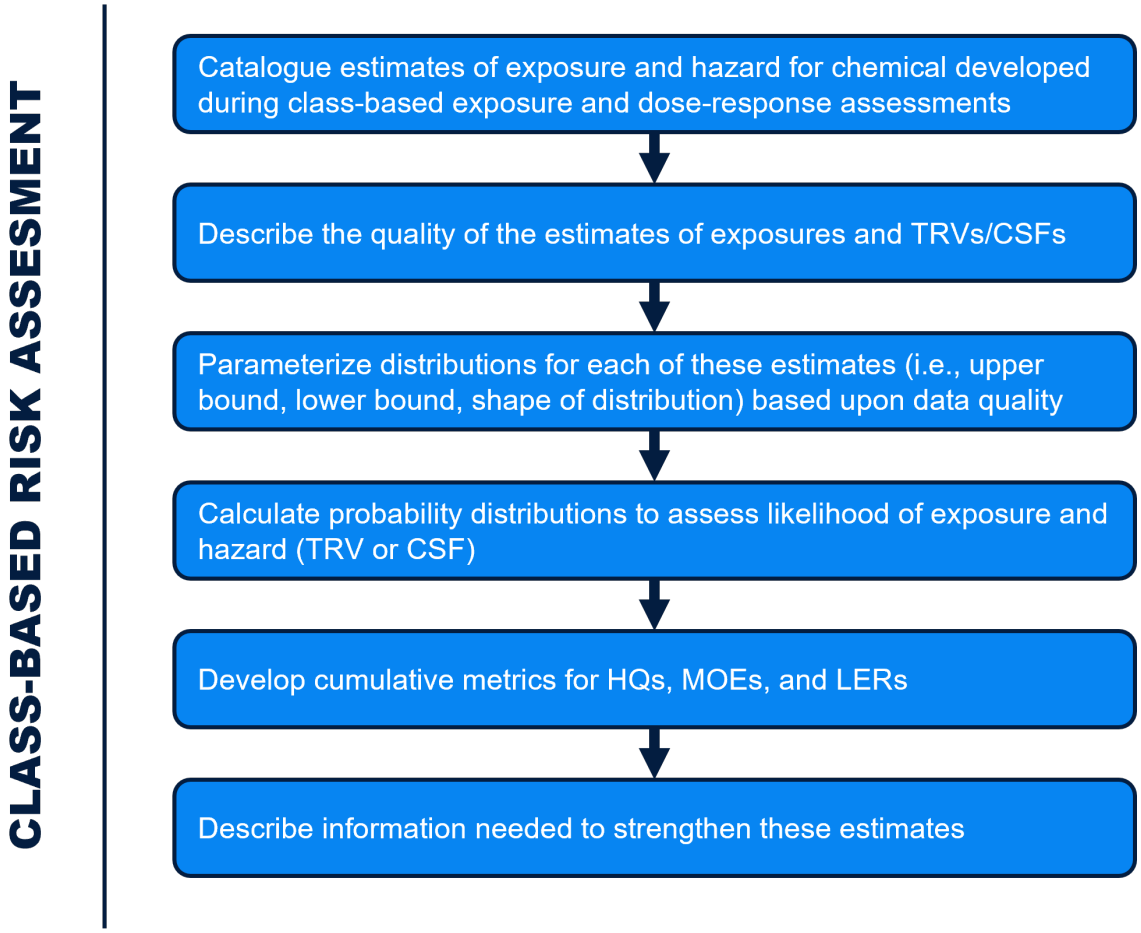
3. Class-based Risk Assessment

Although there is a relatively extensive literature on grouping chemicals into classes for hazard identification and cumulative risk assessment, the literature on using class-based approaches for filling data gaps, particularly for dose-response and exposure assessment, is limited. Furthermore, there are even fewer available methods for deriving risk values, such as TRVs and CSFs, for chemicals with limited data.

In describing how to conduct a class-based risk assessment, we assume that assessors have conducted class-based exposure assessment and class-based dose-response assessment. Specific guidance for these two processes is described in the CO-3 Report (Call Order No. 61320622F2012 under CPSC Blank Purchase Agreement (BPA 61320622A0005)) and Subtask 4A report.[1] Here, we focus on a tiered approach for parameterizing probability distributions for both exposure and dose-response estimates derived using those methods. This framework is loosely based on the approach proposed by Alexander-White et al. (2022), and the EPA Office of Pesticides Program (OPP) (U.S. EPA, 2002) framework for addressing uncertainties in data. To this end, we propose performing a probabilistic risk assessment that draws on ideas presented in Silva et al. (2020), the basic steps of which are outlined in Figure 1.

[1] CO-3 Report, "Exposure Assessment of PHOP Flame Retardants Using Human Biomonitoring Data" is available at https://www.cpsc.gov/Business--Manufacturing/Organohalogen-Flame-Retardant-Chemicals-Assessment.

Figure 1. Basic Steps of Class-based Risk Assessment



TRV = toxicity reference value; CSF = cancer slope factor; HQ = hazard quotient; MOE = margin of exposure; LER = lifetime excess risk.

It is important to note that there are other probabilistic frameworks. The most notable of these is the Approximate PROBabilistic Analysis tool (APROBA), which is based on the International Programme on Chemical Safety (IPCS) guidance on evaluating and expressing uncertainty in hazard characterization (WHO, 2018), but is not solely focused on human risk assessment. The IPCS guidance focuses on calculating the target human dose associated with an effect size of magnitude M and population incidence I (HD_M^I) (Chiu et al., 2018; Chiu & Slob, 2015). While these methods work well to derive TRVs, they are somewhat data intensive and require specific choices from the risk manager about acceptable levels to be observed in the population (Bokkers et al., 2017).

For this reason, and for ease of interpretation, we chose to base this guide on a method that incorporates the idea of uncertainty in the estimate by replacing single point estimates with distributions (van der Voet & Slob, 2007). This allows for the calculation of more traditional risk metrics, such as the HQ and MOE, while incorporating the idea that

there may be a lack of certainty in the estimates, particularly for data-poor class members.

In this report, we describe the approach proposed based upon the strength of the estimates derived from available data. We then note factors that increase and decrease confidence in these estimates. Importantly, as we note in CO-3 and in the Subtask 4A reports, chemical classification, anchor chemical selection, and data curation all play key roles in the development of class-based exposure estimates and class-based TRV/CSF estimation. The risk estimation relies on the quality of underlying data and confidence in the estimations performed in these previous steps. The confidence ratings assigned during risk assessment relate solely to the degree of uncertainty, while the distributions themselves capture both interindividual variability and uncertainty. That is, the width of the distribution can be narrowed by new or refined data that would change the confidence rating. Specifically, when limited data exist and values are derived from generic chemical class or generic default data, the distribution is wide. This distribution can be narrowed by replacing these generic values with chemical specific information for either TRV or exposure to better estimate risk. However, the distribution will never be a single point estimate because not every individual has the same level of exposure or physiological characteristics that drive hazard, as seen by distributions for higher confidence data. This means that uncertainty in the distribution can be reduced, but some variability will always remain because of inherent variability of exposed populations.

4. Data Availability

The first step to class-based risk assessment is to describe the availability of the data for the group of chemicals for which risk assessment is being conducted. Sources of data are briefly described below for exposure and point of departure/TRV information. A more complete description of sources and the pros/cons of using each can be found in the CO-3 report and Subtask 4A report.

4.1. Hazard Data

To perform class-based risk assessment, TRVs or PODs are needed. Class-based dose-response assessment should be performed prior to initiating the class-based risk assessment (Subtask 4A). Briefly, this process consists of deriving PODs/TRVs for each chemical. Available TRVs can be directly extracted from agency websites. If a chemical does not have a POD, one will need to be derived. Sources for PODs, or dose-response information that can be used to derive a POD, include peer reviewed literature search, U.S. EPA's ToxValDB, U.S. EPA's CompTox Dashboard, Gene Expression Omnibus (GEO), National Toxicology Program's (NTP) Integrated Chemical Environment (ICE), and Health Canada's automated workflow for prioritization (HAWPr). Additionally, the Chiu Dashboard (Aurisano et al., 2023; Kvasnicka et al., 2024), which is based on collated animal data, contains PODs derived *in silico* for 804,572 chemicals (as of the writing of this report). Importantly, these additional sources may include PODs based on lower quality studies or based on unpublished data that are not included in peer reviewed literature. This means that additional assessment and verification of the quality of these data may be needed beyond that required for TRVs from an agency website or developed in the peer reviewed literature. This is more fully described in the Subtask 4A report

4.2. Exposure Data

In addition to TRVs/PODs, class-based risk assessments also require estimates of exposure, which can be accomplished using various approaches. Table 1 provides descriptions and examples of input data needed for each approach. Depending on the nature of the input data, the source of the input data may be from peer-reviewed literature (e.g., empirical measurements, [bio]monitoring data), grey literature/authoritative sources (e.g., migration rates, mouthing durations), handbooks (e.g., exposure factors), databases (e.g., CompTox physicochemical properties, National Health and Nutrition Examination Survey [NHANES] biomonitoring data), or professional judgement (e.g., frequency and duration of use parameters). The availability of the input data determines whether an approach can be used to estimate exposure.

For class-based exposure assessments, semi-quantitative estimates of exposure can also be determined by extrapolating doses from data-rich chemicals to data-poor chemicals. These semi-quantitative estimates are usually denoted as "approximately," "greater than," or "less than," where a data-rich chemical serves as the reference chemical. For example, chemicals with use cases comparable to that of data-rich chemical "A" and a vapor pressure lower than the vapor pressure of data-rich chemical "A" are assumed to have inhalation doses lower than that of "A".

Table 1. Approaches for Estimating Exposure

Approach	Description	Example Input Data Needed
Mechanistic Models	• Uses mechanistic models based on well-established mechanistic processes informed by physicochemical properties to estimate indoor environmental concentrations and/or doses associated with consumer products used in indoor or quasi-indoor environments • Exposure scenarios (i.e., source-pathway-receptor) are first developed to describe how exposure occurs • Example consumer exposure models include CEM, SHEDS-HT, and ConsExpo	• Physicochemical properties, product information (e.g., chemical concentration), environmental inputs, exposure factors

Empirical Measurements	• Includes:
	  ○ Chemical migration rates from products to people to estimate contact (direct) consumer exposure
	  ○ Chemical emissions measurements from products to indoor environments to estimate mediated (indirect) consumer exposure
	  ○ Chemical concentrations in various media for contact or mediated consumer exposure
	• Measurements are used with simple equations
		• Migration rates into saliva/sweat
		• Chemical-specific dermal inputs (e.g., permeability coefficient)
		• Chemical concentrations from skin wipes, indoor dust, or personal air
		• Product emission/mass transfer parameters and environmental inputs
		• Exposure factors

Reverse Dosimetry	• Uses human biomonitoring and chemical-specific toxicokinetic data to estimate dose	• Chemical concentrations in human biomatrices (e.g., blood, urine) • Chemical-specific toxicokinetic parameters (e.g., fractional urinary excretion)

5. Probability Distribution Parameterization and Derivation

Following data collation, data should be categorized into tiers based upon the ability to parameterize a probability distribution (sometimes referred to simply as a "distribution"). This tiering process informs how distributions account for uncertainty in the estimates of hazard and exposure. Importantly, distributions will be narrower for chemicals for which the data are higher quality. The assumption is that most variability observed in the distribution of higher quality data is a function of population variability, while those data in which we have less confidence in will have wider distributions, as the distributions need to account for population level variability as well as greater uncertainty in the estimate of the central tendency and variability. Note that, although the tiers are related to the confidence in the hazard and exposure estimates, there is not a direct 1:1 relationship, particularly at the lower confidence levels. The tiers discussed in this section are strictly for the purpose of describing the types of data available, which determines the mathematical approach for describing the distribution.

- Tier I: Probability distribution is defined and given by a central tendency (e.g., geometric mean, arithmetic mean) and standard deviation (e.g., geometric standard deviation, arithmetic standard deviation).
- Tier II: Probability distribution is not defined but the shape of the distribution can be assumed based on existing exposure/hazard knowledge. The central tendency, 5^{th}, and 95^{th} percentile are estimated.
- Tier III: Probability distribution is not defined, but the shape of the distribution can be assumed based on existing exposure/hazard knowledge. In contrast to Tier II, only the 5^{th} and 95^{th} percentile are estimated.

It is also important to note that generally, estimation methods for higher quality data will result in narrower distributions relative to estimates for lower quality data. Additionally, the estimation of a distribution can be changed at any time based on new data that are collected. Below we highlight different methods by which the distribution can be defined for risk assessment purposes. It is important to note that this is not an exhaustive list, and there are other options for estimating distributions. However, the options described cover basic ways in which distributions may be estimated for the hazard and exposure data identified. Special consideration for parameterizing for cancer risk distribution are described throughout.

5.1. Distribution Parameterization

5.1.1. Hazard Parameterization

During hazard characterization, both the POD and TRV estimation should be considered. For class-based risk assessment, a shared hazard (e.g. shared health effect, target organ/system, MOA, mechanism of effect) is necessary. This is because hazard extrapolation from high and medium confidence estimates to low confidence estimates assumes that there is a similar toxicity pathway at play. Tiers for estimating distributions for PODs and TRVs are described below briefly and found in Table 2. In Subtask 4A, we describe important caveats to the confidence of estimates for each tier of data in more detail. These considerations include agency review, study quality, literature concordance, and data type (e.g., *in vivo*, *in vitro, in silico*).

Table 2. Example Data Used to Parameterize a Hazard Distribution

Hazard Tier	Distribution Type	Central Tendency	Standard Deviation	5th Percentile	95th Percentile
Tier I	Normal	• "Best" TRV • Median or mean agency TRV	Calculated based on parameterized distribution	Select the most relevant[a]: • TRV/3 • Lowest agency TRV	Select the most relevant[b]: • TRV*3 • Highest agency TRV
Tier II	Log Normal	• Newly identified/derived *in vivo* or *in vitro* TRV • *In vivo, in vitro*, or *in silico* derived RPF/RAX-TRV • Empirical NAMs-based TRV (e.g., tPODs, non-traditional animals) • *In silico* predicted TRV • Anchor chemical/10 • TTC • Most sensitive TRV/10	Calculated based on parameterized distribution	Select the most relevant[a]: • TTC • Most sensitive TRV/10 • Anchor chemical TRV/10	Select the most relevant[b]: • Select the highest of the available PODs from any available data stream including *in silico*-derived values • Anchor chemical TRV*100
Tier III	Log Normal	• Calculated based on parameterized distribution	Calculated based on parameterized distribution	Select the most relevant[a]: • TTC • Most sensitive TRV/10 • Anchor chemical TRV/100	• Anchor chemical TRV*100

TRV = toxicity reference value; RPF = relative potency factor; RAX = read across; NAMs = new approaches/alternative methods; tPOD = transcriptomic point of departure; TTC = threshold of toxicological concern; PODs = points of departure

[a] The most relevant for lower bounds may be the lowest value but is not necessarily the lowest value.

[b] The most relevant for upper bounds may be the highest value but is not necessarily the highest value.

5.1.1.1. Tier I

Generally, Tier I hazard characterization estimates are associated with an agency assessment and derivation of a TRV, peer-reviewed publication of a TRV, or a POD from either source in the context of an MOE assessment. This group encompasses only a few hundred chemicals. Multiple agencies may have assessed the same chemical and the POD/TRV estimates may differ among assessments. For these cases, the competing PODs/TRVs can be used to define the bounds of the distribution. Alternatively, a TRV can be selected and used to define an anchor chemical based upon health effects shared with other chemical members of the same class. Importantly, depending on confidence considerations, some TRVs may not be included (Subtask 4A). Despite these issues, there is generally a higher degree of certainty in these estimates compared to those from other sources, and as such, the distributions of these PODs/TRVs will be narrower. Additionally, Tier I data are parameterized by a normal distribution, as the assumption is that hazard is proportionally distributed around the central tendency. The central tendency is used as the mean of the distribution and the standard deviation is given by the TRV/3 and the TRV*3 to reflect a one order of magnitude degree of uncertainty.

For cancer, higher confidence estimates will be based on one or more high-quality bioassays. In addition, the MOA by which the chemical of interest causes cancer, whether mutagenic or non-mutagenic, should be well understood. Situations where there is substantial evidence pointing towards a non-mutagenic MOA, but not enough to meet regulatory criteria, would also be evaluated using Tier I methods using the CSF calculated with standard methods as the central tendency. However, additional considerations apply to the estimation of the bounds in such a case, as discussed in Section 6.2.

5.1.1.2. Tier II

Generally, PODs/TRVs independently derived by the assessor from any data stream are considered Tier II hazard characterization estimates. PODs/TRVs derived using these methods could include modeling of data from a traditional animal model, derivation based on empirical NAMs data (e.g., tPODs, non-traditional animal models, *in vitro* data), or use of relative potency factors (RPFs) or read across methods (RAX) that utilize *in vivo, in vitro,* or *in silico* data. While TRVs derived using each of these methods are associated with different confidence levels, they all provide a central tendency estimate that can be used to parameterize a distribution. In addition, depending on the specifics of the data available for a given subclass, central tendency estimates of low-confidence TRVs can be derived (and hence Tier II methods can be applied) using methods such as the threshold of toxicological concern (TTC), applying the TRV for the most sensitive member of the subclass, or the anchor chemical TRV/10. Given the varying confidence in these estimates, different choices may be made in how distributions are parametrized based

on the confidence assigned to the estimate during class-based dose-response (Subtask 4A).

However, for Tiers II data, the assumption is that hazard is log normally distributed, as the log normal distribution is right skewed. A right skewed distribution means that the probability of selecting a value closer to zero is higher than seen with a normal distribution with the same parameters (Fayers, 2011). For the distribution, the central tendency (TRV) is used as the value for the geometric mean, and the 5th and 95th percentiles are used to define the geometric standard deviation. Additionally, estimates of standard deviation are calculated based on the parameterized distributions. Data availability and confidence should play a key role in choosing options for parameterizing toxicity distributions.

For mutagenic cancer, Tier II CSF estimates will result when the CSF is derived from a high or medium quality bioassay and the MOA is not known. In addition, empirical data not meeting the standard of a high quality bioassay can also be used to establish a Tier II CSF estimate, regardless of whether the MOA is known. Given that the goal of class-based assessments is to provide estimates for class members based on other chemical members of the class, data that would not generally be suitable for risk assessment can be incorporated, but cautious interpretation of the findings is needed. If the MOA is demonstrated to be non-mutagenic, standard approaches are applied to develop a TRV and associated bounds.

5.1.1.3. Tier III

Tier III estimation methods are used rarely since toxicity studies typically provide a single number for the POD (e.g., LOAEL, NOAEL, benchmark dose [BMD], BMDL) when calculating a TRV, rather than a range of numbers. However, a POD typically is not available for data poor chemicals. In such cases, a range of potential TRVs for the data-rich chemical members of the class can be used to bound the distribution of TRVs for data-poor chemicals. For example, the TTC may be chosen as the 5th percentile but other data for extrapolating from the target chemical may be insufficient for estimating a central tendency. Generally, these upper and lower bound estimates are likely to result in very broad distributions reflecting the large degree of uncertainty in the estimates.

As for Tier II data, it is assumed that hazard is log normally distributed for Tier III data, as the log normal distribution is right skewed (Section 5.1.1.2). Estimates of standard deviation and central tendency are calculated based on the parameterized distributions.

For cancer, there is no obvious way to create Tier III estimates of the CSF. This is related to the difficulty of using linear extrapolation to accurately estimate risk at doses orders

of magnitude below the doses tested in the animal bioassay. Lower confidence CSF estimates would involve a high degree of uncertainty, which may or may not be health protective. As such, we recommend only estimating CSFs for Tiers I and II data.

5.1.2. Exposure Parameterization

Exposure estimates can vary depending on the approach used and the source of the input data (Section 4.2). Confidence ratings on the quality of the exposure estimate are based on three criteria:

1. relative agreement among exposure estimates from different approaches,
2. relative amount of input data available, and
3. reasonableness of the estimate (i.e., the dose falls between 10^{-7} to 10^{-2} mg/kg/day).

Regarding the type of approach used, reverse dosimetry is the preferred approach if the input data are available. Reverse dosimetry uses human biomonitoring (a direct measure of the chemical in people) and toxicokinetic data to provide the total dose for all sources and pathways to which a person is exposed. However, depending on the dataset, the biomonitoring data available may not be nationally representative or, in the case of NHANES, may not capture potentially under-sampled population groups.

Doses estimated using empirical measurements and/or environmental monitoring data reflect the actual migration/emissions/concentration measured and are therefore 'real world samples'. However, similar to human biomonitoring data, these migrations/emissions/concentrations may be dependent on the product/experimental conditions, and monitoring data may be dependent on the microenvironment/geographic location.

Due to the relative scarcity of environmental monitoring, source characterization, and human biomonitoring data for chemicals overall, modeling exposure scenarios may sometimes be the only approach available for estimating exposure. However, careful selection of the input parameters is needed to avoid calculating unrealistic doses.

Some ways in which exposure distributions may be characterized are described in Table 3. Generally, Tier I data will have a geometric mean and standard deviation that are predefined by a given population(s) or model estimation(s). As with hazard data, Tier II data will have a predefined estimate of the central tendency and 5th and 95th percentile, while Tier III data will have estimates associated with only the 5th and 95th percentile. When considering which option to select for parameterizing exposure distributions, data quality should play a role based on the description in Section 4. Specifically, there should

be some degree of certainty in the underlying studies before selecting any particular method or combination of methods for estimating the exposure level(s).

Table 3. Example Data Used to Parameterize an Exposure Distribution

Exposure Tier	Distribution Type	Central Tendency	Standard Deviation	5th Percentile	95th Percentile
Tier I	Log Normal	GM	GSD	N/A	N/A
Tier II	Log Normal	Measured or modeled from population(s)	Calculated based on parameterized distribution	Select the most relevant: • Lower limit of detection for biomonitoring method • 5th percentile of a low-exposure population • Modeled estimate	Select the most relevant: • Upper limit of detection for biomonitoring method • 95th percentile of a highly-exposed population • Modeled estimate
Tier III	Log Normal	Calculated based on parameterized distribution	Calculated based on parameterized distribution	Select the most relevant: • *De-minimis*[a] exposure • Sum of a subset of pathways (e.g., key pathways missing) • Background exposure only	Select the most relevant: • Aggregate total exposure from modeled estimates, assuming all products used in household contain chemical of interest • Maximal exposure estimate[b]

GM = geometric mean; GSD = geometric standard deviation; N/A = not applicable

[a] *De-minimis* exposure = 10^{-7} mg/kg/day. While exposures can be quantified below this level, it is unlikely that they will result in appreciable differences in aggregate exposure estimates for a population. The *de-minimis* exposure for this assessment may differ from *de-minimis* exposures used for other chemicals and other scenarios.

[b] Maximal exposure estimate = 10^{-2} mg/kg/day. While exposures can exceed 10^{-2} mg/kg/day in some scenarios, it is likely that most exposures will fall below this level. Individual consideration of highly exposed individuals within a population can always be considered on a case-by-case basis. Exposure for these individuals may exceed the maximal exposure estimate. The maximal exposure for individual assessments may differ from maximal exposures used for other chemicals and other scenarios. A more global maximal default exposure estimate of 1 mg/kg/day may be considered as it unlikely that most aggregate chronic doses will exceed this level for most chemicals.

5.1.2.1. Tier I

Exposure estimates are considered to be Tier I estimates when there is agreement between doses calculated from at least two approaches. Doses that are within a factor of 2–3 between approaches provide greater confidence that:

- If reverse dosimetry was used, the biomonitoring data used is representative of the general population and not specific to a population subgroup.
- If empirical measurements were used and the resulting doses are summed to determine an aggregate dose, the empirical data available captures the key sources and pathways.
- If mechanistic modeling was used, the exposure scenarios developed represent the key consumer product uses in the general population, and the input data used and assumptions made were reasonable.

The amount of input data used to generate the exposure estimates should also be considered. For reverse dosimetry, both NHANES and non-NHANES individual cohort studies should be used to estimate doses, and those values should be compared. For exposures estimated from empirical measurements, the total dose is the sum of individual doses from multiple pathways (i.e., total dose = inhalation dose + ingestion dose + dermal dose). Within each pathway, there could also be multiple contributions (e.g., ingestion dose = dust ingestion + mouthing + drinking water/soil/food ingestion). In general, exposure estimates derived from four or more datasets per data type (e.g., four datasets on indoor dust concentrations) are considered higher confidence. In certain cases, such as for mouthing, the migration rates needed to calculate exposure are often not measured and, in these instances, one empirically measured value is considered higher confidence than using an empirical model-predicted value (e.g., consumer exposure model's [CEM] suggested default migration rates).

Finally, exposure estimates from all approaches should be evaluated for reasonableness particularly when modeling (e.g., mechanistic modeling, simpler modeling with empirical measurements). Higher confidence estimates are values that fall within the expected range. Values that are too high indicate that the input parameters or equations need to be reassessed or there was potential overinclusion of sources. Similarly, values that are too low (e.g., inhalation doses of 10^{-9} mg/kg/day for chemicals with very high vapor pressure) also require additional follow-up.

5.1.2.2. Tier II

Generally, Tier II exposure estimates are developed when the central tendency can be calculated and empirical information about the range of exposure is available to derive doses for some individual pathways that are supplemented with modeled estimates for

some of the remaining pathways to give a complete picture of exposure overall. Estimated values should fall within the typical dose range. There may be some pathways that are unaccounted for but there is still a fair degree of confidence in the estimate of exposure across the population. Thus, the distribution estimate will likely be wider than those for Tier I chemicals.

5.1.2.3. Tier III

Tier III exposure estimates are used:

- For empirical measurements when estimates are based on only one or two exposure pathways but it is likely that other important exposure pathways may be missing, and therefore total dose is not captured.
- For empirical measurements when estimates are derived using only one or two datasets (e.g., one indoor air monitoring study).
- For modeling data when values are derived using an existing model that was trained on a different set of chemicals, and there is no mention that the chemicals of interest are within the domain of applicability.
- For all approaches when exposure values are outside the typical dose range of 10^{-7} to 10^{-2} mg/kg/day

In this case, the exposure estimates are used to draw boundaries for the 5th and 95th percentile with limited ability to estimate the central tendency.

5.2. Distribution Calculations

5.2.1. Tier I

5.2.1.1. Hazard

For some data estimates, there is a good estimation of the central tendency and defined 5th and 95th percentiles, with the assumption of a normal distribution rather than a log normal distribution. A normal distribution shape is chosen to estimate the TRV distribution as we assume that the TRV distribution should be equally sized on either side because the TRV is defined as having a one order of magnitude uncertainty surrounding it (Krithikadatta, 2014; U.S. EPA, 2002). In these cases, the population mean is estimated with the TRV central tendency estimate ($\hat{\mu}$), and the 5th ($q_{0.05}$) and 95th ($q_{0.95}$) percentile of the distribution are estimated as spanning one order of magnitude centered on the central tendency (i.e., the 5th percentile is equal to the mean divided by 3 and the 95th percentile is equal to the mean multiplied by 3). This distribution is based on the phrase in the definition of the reference dose (RfD) that it is a value with "uncertainty spanning perhaps an order of magnitude." This quoted phrase has been interpreted to mean that

the RfD is at the upper end, the lower end, or the middle of the range of an order of magnitude (U.S. EPA, 2002). More nuanced considerations recognize that larger uncertainty factors are related to greater uncertainty (Dourson, Felter, & Robinson, 1996). Other options exist for bounding these estimates, such as using different agency TRVs or different factors to establish the confidence limits.

The α-percentile equation for a normal distribution with mean μ and standard deviation σ is:

$$q_\alpha = \mu + \sigma \times \Phi^{-1}(\alpha)$$

Where α is the cumulative probability and Φ is the cumulative distribution function (CDF) for the standard normal distribution. From this equation, we solve for μ and σ:

$$\mu = q_\alpha - \sigma \times \Phi^{-1}(\alpha) \quad (1)$$

$$\sigma = \frac{q_\alpha - \mu}{\Phi^{-1}(\alpha)} \quad (2)$$

Two estimates of σ are calculated from Equation (2) as:

$$\hat{\sigma}_1 = \frac{q_{0.05} - \hat{\mu}}{\Phi^{-1}(0.05)}$$

$$\hat{\sigma}_2 = \frac{q_{0.95} - \hat{\mu}}{\Phi^{-1}(0.95)}$$

A third estimate of σ is calculated by plugging $q_{0.05}$ and $q_{0.95}$ into Equation (1), setting the two values equal to each other, and solving for σ.

$$\hat{\sigma}_3 = \frac{q_{0.95} - q_{0.05}}{\Phi^{-1}(0.95) - \Phi^{-1}(0.05)}$$

The average of these three estimates was used as the final estimate $\hat{\sigma}$.

5.2.1.2. Exposure

For some higher quality exposure data, the geometric mean (GM), geometric standard deviation (GSD), and 95th percentile ($q_{0.95}$) can be extracted. For these data, the distribution is characterized with log normal distribution where the log-transformed exposure values are normally distributed with population mean (μ) and standard deviation (σ). The GM and GSD data are used to estimate μ and σ using a log transformation (Limpert, Stahel, & Abbt, 2001).

$$\hat{\mu} = \log(GM)$$

$$\hat{\sigma} = \log(GSD)$$

The $q_{0.95}$, $\hat{\mu}$, and $\hat{\sigma}_1$ are used to estimate a second set of parameters based on the cumulative distribution function (CDF) for the log normal distribution:

$$P(X < x) = \Phi\left(\frac{\log(x) - \mu}{\sigma}\right) \quad (3)$$

where Φ is the CDF for the standard normal distribution (Johnson, Kotz, & Balakrishnan, 1994). In Equation (3), the value of x corresponding to the desired probability α is the percentile q_α such that:

$$\alpha = \Phi\left(\frac{\log(q_\alpha) - \mu}{\sigma}\right)$$

which can be rearranged as:

$$\Phi^{-1}(\alpha) = \Phi^{-1}\left(\Phi\left(\frac{\log(q_\alpha) - \mu}{\sigma}\right)\right)$$

$$\Phi^{-1}(\alpha) = \frac{\log(q_\alpha) - \mu}{\sigma}$$

$$\log(q_\alpha) = \sigma \times \Phi^{(-1)}(\alpha) + \mu \quad (4)$$

Equation (4) can be rearranged to solve for μ and σ as:

$$\mu = \log(q_\alpha) - \sigma \times \Phi^{(-1)}(\alpha) \quad (5)$$

$$\sigma = \frac{\log(q_\alpha) - \mu}{\Phi^{(-1)}(\alpha)} \quad (6)$$

With equations (5) and (6), a second set of parameters is then estimated as:

$$\hat{\mu}_2 = \log(q_{0.95}) - \hat{\sigma}_1 \times \Phi^{-1}(0.95)$$

$$\hat{\sigma}_2 = \frac{\log(q_{0.95}) - \hat{\mu}_1}{\Phi^{-1}(0.95)}$$

The final distribution parameter estimates, $\hat{\mu}$ and $\hat{\sigma}$, are then calculated as the average of the two estimates.

5.2.2. Tier II

For moderate quality data for both exposure and hazard data, the estimate of the central tendency as well as the estimates for the 5th and 95th percentiles are known. This situation also often applies to low confidence hazard data. For both hazard and exposure, a log normal distribution is assumed. For exposure, log normality is a replicated trend across populations and is the general distribution shape for most environmental exposures (Flynn, 2004). For hazard, a health protective approach in the face of uncertainty for the actual TRV value is to assume that the probability of hazard is log normally distributed, as the distribution is right skewed. This would mean that the probability of choosing a number closer to 0 is higher than would be the case with a normal distribution that is symmetrically distributed around the central tendency. Bounds should be selected based upon the options found in Table 2 and Table 3 for hazard and exposure, respectively. These bounds can then be used to estimate the parameter σ by plugging the values into Equation (5), setting the estimates equal to each other, and solving for σ :

$$\hat{\mu}_{q05} = \log(q_{0.05}) - \hat{\sigma}_1 \times \Phi^{-1}(0.05) = \log(q_{0.95}) - \hat{\sigma}_1 \times \Phi^{-1}(0.95) = \hat{\mu}_{q95}$$

$$\hat{\sigma}_1 \times \Phi^{-1}(0.95) - \hat{\sigma}_1 \times \Phi^{-1}(0.05) = \log(q_{0.95}) - \log(q_{0.05})$$

$$\hat{\sigma}_1 = \frac{\log(q_{0.95}) - \log(q_{0.05})}{\Phi^{-1}(0.95) - \Phi^{-1}(0.05)}$$

This value is used to estimate the parameter μ using Equation (5) from Section 5.2.1.2:

$$\hat{\mu}_1 = \log(q_{0.05}) - \hat{\sigma}_1 \times \Phi^{-1}(0.05)$$

The sample mean of a log normal distribution for random variable X can be expressed in terms of the parameters μ and σ (Johnson, Kotz, & Balakrishnan, 1994):

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right) \quad (7)$$

Using Equation (7), a second estimate of μ is calculated by plugging in $\hat{\sigma}_1$ and the TRV estimate (E(TRV)) as the log normal mean, E(X):

$$E(TRV) = \exp\left(\hat{\mu}_2 + \frac{\hat{\sigma}_1^2}{2}\right)$$

$$\log\big(E(TRV)\big) = \hat{\mu}_2 + \frac{\hat{\sigma}_1^2}{2}$$

$$\hat{\mu}_2 = log\big(E(TRV)\big) - \frac{\hat{\sigma}_1^2}{2}$$

The final estimate $\hat{\mu}$ is calculated as the average of the two μ estimates.

For chemicals where $\log(E(TRV)) > \hat{\mu}_1$, a second estimate of σ is calculated by plugging $\hat{\mu}_1$ and E(TRV) into Equation (7):

$$E(TRV) = \exp\left(\hat{\mu}_1 + \frac{\hat{\sigma}_2^2}{2}\right)$$

$$\log[E(TRV)] = \hat{\mu}_1 + \frac{\hat{\sigma}_2^2}{2}$$

$$\frac{\hat{\sigma}_2^2}{2} = \log\big(E(TRV)\big) - \hat{\mu}_1$$

$$\hat{\sigma}_2^2 = 2 \times \left[\log\big(E(TRV)\big) - \hat{\mu}_1\right]$$

$$\hat{\sigma}_2 = \sqrt{2 \times [\log\big(E(TRV)\big) - \hat{\mu}_1]}$$

In cases where there are two estimates of σ, the average of the two estimates is used as the final estimate $\hat{\sigma}$, otherwise $\hat{\sigma}_1$ is selected.

5.2.3. Tier III

For some estimates, only the 5th percentile ($q_{0.05}$) and the 95th percentile ($q_{0.95}$) are defined. These are generally lower confidence estimates. As described in Section 5.1 for both exposure and hazard, the shape of the distribution is assumed to be log normal. Empirical evidence shows that exposures in the general population usually follow a log normal distribution. For hazard, we assume a log normal distribution to be more health protective than a normal distribution. For this data type, the first step is to estimate the log normal parameter $\hat{\sigma}$ as described for $\hat{\sigma}_1$ in Section 5.2.2

The log normal parameter μ is estimated by plugging values into Equation (6), setting the estimates equal to each other, and solving for $\hat{\mu}$:

$$\frac{\log(q_{0.95}) - \hat{\mu}}{\Phi^{-1}(0.95)} = \frac{\log(q_{0.05}) - \hat{\mu}}{\Phi^{-1}(0.05)}$$

$$\Phi^{-1}(0.05) \times (\log(q_{0.95}) - \hat{\mu}) = \Phi^{-1}(0.95) \times (\log(q_{0.05}) - \hat{\mu})$$

$$\Phi^{-1}(0.05) \times \log(q_{0.95}) - \Phi^{-1}(0.05) \times \hat{\mu} = \Phi^{-1}(0.95) \times \log(q_{0.05}) - \Phi^{-1}(0.95) \times \hat{\mu}$$

$$\Phi^{-1}(0.95) \times \hat{\mu} - \Phi^{-1}(0.05) \times \hat{\mu} = \Phi^{-1}(0.95) \times \log(q_{0.05}) - \Phi^{-1}(0.05) \times \log(q_{0.95})$$

$$\hat{\mu} \times \left(\Phi^{-1}(0.95) - \Phi^{-1}(0.05)\right) = \Phi^{-1}(0.95) \times \log(q_{0.05}) - \Phi^{-1}(0.05) \times \log(q_{0.95})$$

$$\hat{\mu} = \frac{\Phi^{-1}(0.95) \times \log(q_{0.05}) - \Phi^{-1}(0.05) \times \log(q_{0.95})}{\Phi^{-1}(0.95) - \Phi^{-1}(0.05)}$$

6. Calculating Risk Probabilities from Distributions

Once distribution shapes are defined, exposure distributions must be compared to hazard distributions. Generally, the region of concern is where the probability distributions overlap between hazard and exposure (i.e., the region where estimated exposure is higher than the estimated TRV). The goal of the risk calculations is to identify the cumulative probability of exposure at a level of concern (i.e., the proportion of the population that is contained within this overlap). Below is a description of how this can be done for both TRVs and CSFs.

Additionally, it is important to note that while the 95th percentile of the exposure estimate corresponds to higher exposure, the 95th percentile of the TRV corresponds to lower toxicity. Thus, the highest risk is with the 95th percentile of exposure and 5th percentile of the TRV. This inversion of the 5th and 95th percentile means that in a low risk exposure scenario the distributions for hazard and exposure should minimally overlap with exposure being to the left of hazard on the graph. In high risk scenarios, the overlap between the two distributions will be more significant.

6.1. Threshold-based Analysis (HQs and MOEs)

To conduct risk analysis for TRVs, we perform a HQ analysis, calculating a cumulative probability assessment of the HQ.

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

HQ is interpreted relative to a value of 1. If HQ ≤ 1, no adverse effects are expected. If HQ > 1, then sensitive populations may begin to experience mild effects, with the likelihood of an effect and potential severity of the effect increasing as the HQ increases. For cases where both exposure and TRV are assumed to be log normally distributed, the cumulative probability that HQ > 1 (CP(HQ > 1)) is calculated from the cumulative distribution function for a log normal distribution with:

$$\mu = \hat{\mu}_e - \hat{\mu}_{TRV} \text{ and } \sigma = \sqrt{\hat{\sigma}_e^2 + \hat{\sigma}_{TRV}^2}.$$

For cases where exposure is assumed to be log normally distributed and TRV is assumed to be normally distributed, 100,000 exposure values are randomly sampled from $LogNormal(\hat{\mu}_e, \hat{\sigma_e})$ and 100,000 TRV values are randomly sampled from $N(\hat{\mu}_{TRV}, \hat{\sigma_{TRV}})$. The values were divided to generate a simulated sample of 100,000 HQs. CP(HQ > 1) is estimated as the proportion of sampled HQ values greater than 1.

Given that a HQ > 1 indicates sensitive populations may begin to experience mild effects, one interpretation of the cumulative probability of HQ > 1 is that (at relatively low probabilities) it represents the probability of mild effects developing in sensitive populations based upon the given exposure scenario. Higher probabilities of HQ > 1 (and thus presumably higher exceedances) may result in moderate or worse effects in a broader population. The relationship among the probability of HQ > 1, the percentage of the population affected, and severity of effect is not well known, and depends on the chemical-specific shape of the dose-response curve, the effects observed, and the magnitude of the composite UF. In light of these considerations, there is not an explicit and straightforward way to interpret the implications of the percent of the population with HQ > 1. Based on the approach of Chiu and Slob (2015), we are considering cumulative probability values below 0.01 (i.e., 1 out of 100, or 1%) to be not of concern, values between 0.01-0.05 to be of marginal concern, and values above 0.05 to be of concern (Chiu & Slob, 2015). These levels of concern reflect the definition of the TRV in protecting sensitive populations, and the expectation that small exceedances of the TRV would result in mild effects.

For a MOE calculation as given by

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

the same approach described above could be undertaken. In this case, the POD would be used in place of the TRV. Similarly, the cumulative probability could be calculated for a MOE < 100, which represents a scenario where the margin between the POD for a given health effect and exposure is insufficient, taking into account interspecies extrapolation and human variability. The interpretation of the cumulative probability for a MOE < 100 would be similar to that applied to the HQ > 1, with cumulative probability of <0.01 being not of concern, 0.01-0.05 being of marginal concern, and >0.05 being of concern. Similarly, the central tendency values for these data should also be estimated.

In addition, to the probabilistic values based on HQ and/or MOE, the central tendencies of the exposure and TRV (or POD) should be compared using the traditional risk framework. This can help with both risk interpretation as well as identifying chemicals that should be prioritized based on whether the central tendency of exposure exceeds the central tendency of a TRV (or whether the POD is less than 100x the exposure in the context of an MOE).

6.2. Cancer Considerations

When considering cancer, different mathematical assumptions are made. For cancers caused by a mutagenic MOA (or if the MOA is unknown), extrapolation to lower doses is done by drawing a straight line from the POD (i.e., the BMDL10) to the origin to determine the CSF. In other words, a policy decision is made assuming that there is no dose that is without risk. For the purposes of ease of interpretation, we suggest a CSF probability distribution is calculated based upon the assumption that a single response level (such as 10%) is used. In theory it is possible to replace this with a distribution but the interpretation of such a replacement becomes much more complex and difficult to understand. Using this approach, a CSF distribution can be parametrized by the same framework described for estimating the hazard distribution (Table 2).

Importantly, as mentioned in Section 5.1.1.3 and in the Subtask 4A report, actual data should be used in any cancer risk assessment and the use of Tier III methods for estimating CSFs should be avoided. This is because the degree of uncertainty in Tier III estimation does not support a health protective approach. Further, given the desire to be health protective, in general we recommend the use of a log normal distribution shape, as the right skew will lead to a higher probability of choosing lower CSFs, which are associated with higher potency. This is in contrast to a normal distribution, which would lead to an equal probability of picking above or below the central tendency of the distribution.

As noted in Section 5.1.1.1, situations where there is substantial evidence pointing towards a non-mutagenic MOA, but not enough to meet regulatory criteria, pose an additional challenge in characterizing the probability distribution. In these cases, the central tendency would be the CSF calculated using linear extrapolation. However, since substantial professional judgement points to a non-mutagenic MOA, this would suggest that more of the weight should be for a lower potency (higher CSF). This would suggest that a right skew best describes the likely distribution of CSFs. However, since a right-skewed distribution becomes mathematically more complex and is not part of the methods described here, a normal distribution (symmetric around the central tendency estimate) is a health-protective approach to this situation. Additional research would be useful regarding the best approach for describing the probability distribution for the types of chemicals addressed in this paragraph.

The probability distribution of hazard and the probability distribution for exposure are merged through Monte Carlo Simulation. Specifically, the probability distributions can be multiplied by each other across 100,000 simulations to derive a probability distribution for LER. This can then be described by its central tendency and 95th percentile.

For cases where both CSF and exposure are assumed to be log normally distributed, the probability distribution for LER is characterized by a log normal distribution with:

$$\hat{\mu} = \hat{\mu}_e + \hat{\mu}_{CSF} \text{ and } \hat{\sigma} = \sqrt{\hat{\sigma}_e^2 + \hat{\sigma}_{CSF}^2}.$$

The GM and GSD are reported as $\exp(\hat{\mu})$ and $\exp(\hat{\sigma})$, respectively. The 95th percentile is calculated from Equation (2) and reported. In general terms, the α percentile of a distribution is the value of the random variable with cumulative probability α. As an example, if the 95th percentile of the LER is 10^{-3}, we expect 95% of the population to have LER ≤ 10^{-3}. Likewise, we would expect 5% of the population to have LER > 10^{-3}. For cases where exposure is assumed to be log normally distributed and CSF is assumed to be normally distributed, 100,000 exposure values are randomly sampled from $LogNormal(\hat{\mu}_e, \hat{\sigma}_e)$ and 100,000 CSF values are randomly sampled from $N(\hat{\mu}_{CSF}, \hat{\sigma}_{CSF})$. The values are multiplied to generate a simulated sample of 100,000 LERs. The mean, standard deviation, and 95th percentile from the simulated data are reported.

Note that the approach described here is a simplification of the true probabilities, particularly in cases where evidence points to a non-mutagenic MOA, but the weight of evidence is insufficient to meet regulatory requirements. In such cases, the lower end of the distribution can be estimated by the POD/100, or the BMDL10/100, corresponding to about a 1 in 1,000 incidence, compared with a 1 in a million incidence that is considered *de minimis* for a mutagenic carcinogen.

As described above, in addition to the probabilistic values, the central tendencies of the exposure and CSF should also be used to calculate the LER using the traditional risk framework. This can provide insight into how the central tendency of the LER probability distribution compares to the traditional metric. The difference between these two may provide insight into data interpretation.

7. Application to PHOPs and Findings

In the Subtask 3B report, we performed class-based risk assessment for the PHOP flame retardants. We did this by first grouping chemicals based on data availability and quality for exposure data (identified or estimated in CO-3) and for TRVs (identified, derived, or estimated in CO-9 Task 1 or Subtask 2B). Next, we created probability distributions that represented the likely level of exposure and the likely value of the TRV. Lastly, we performed a cumulative probability estimate combining these two distributions to determine the likelihood of exposure above a level of health concern for each chemical. Overall, we were able to construct quantitative risk estimates for 22 of the 29 PHOP chemicals. For the remaining 7 chemicals, no estimates were obtained. When considering data quality impacts on our estimations of risk, we found that for Tier I chemicals, there was a relatively low likelihood of risk, Tier II chemicals were mixed between low and high likelihood of risk, and for Tier III chemicals, there appeared to be a high likelihood of risk. However, as discussed in the following paragraphs, for at least some of the Tier III chemicals, this higher likelihood of risk is based on a mixture of the central tendency estimate and the wide confidence limits reflecting high uncertainty.

Of the cumulative probabilities calculated for PHOP subclass members, we found that 5 chemicals, TCIPP (13674-84-5), TCEP (115-96-8), V6 (38051-10-4), TTBNPP (19186-97-1), and TDBPP (126-72-7) are currently in use at an acceptable level (all CP(HQ > 1) ≤ 0.01). One chemical, TDCIPP (13674-87-8), is of marginal concern (CP(HQ > 1) = 0.05). The remaining 16 chemicals had CP(HQ > 1) ≥ 0.2 and as high as 0.6. Importantly, for lower quality estimates, only four were determined to be worthy of particular attention (140-08-9, 1067-98-7, 76649-15-5, 27568-90-7) because for these four chemicals, the central tendency for exposure is higher than the central tendency of the TRV. Therefore, despite the high degree of uncertainty, the available data does suggest that for these four chemicals, the overlap between exposure and the TRV cannot be explained by the wide distributions. Therefore, we recommended additional research on these four chemicals in particular to determine whether the overlap is real, or whether it results from over-conservative central tendency estimates of exposure and the TRV. If the overlap is real, this could indicate that these chemicals are present at levels of exposure that are likely associated with health effects above the level of acceptable risk.

The biggest takeaway from our findings was a relationship between decreasing data availability and increasing cumulative probability of a HQ of concern. This was likely related to both the conservative approaches used for estimating central tendencies of exposure and TRVs when data were limited, and the increased breadth of the exposure and TRV distribution with decreasing amount and/or quality of available data. Specifically, to compensate for increased uncertainty in the estimate of the level of exposure as well

as health effect, the bounds around the central estimate were widened. This led to a higher probability of an overlap between the exposure and TRV distributions and identified an important caveat of the approach we developed: for risk estimates based on lower quality data, the probabilities of HQ > 1 are likely overinflated due to the relatively broad estimate of the bounds associated with our estimation method. Thus, for many chemicals, the only clear conclusion we can draw is that better assessment of exposure and the TRV are needed.

In our case study report for PHOPs, we did not conduct a class-based assessment for cancer data as we found that there were too many differences between chemicals in terms of MOAs and lack of human relevance of identified MOAs (Subtask 3B). Specifically, three PHOPs had existing cancer dose-response data relevant to humans with a common endpoint of kidney cancer. However, one chemical (TCEP, 115-96-8) is thought to act via a nongenotoxic MOA while the other two (TDCIPP, 13674-87-8 and TDBPP, 126-72-7) are thought to act via a genotoxic MOA. While other tumor types were identified, it was not clear that these tumors developed via a human relevant mechanism or had a significant enough weight-of-evidence to justify class-based assessment. Therefore, we are not able to define key considerations in the probabilistic assessment of LER. Future work should concentrate on how class-based cancer dose-response assessments might be conducted for chemicals with cancer data and how the integration of this data with exposure data can be interpreted within a risk assessment framework. Individual cancer-risk estimates for chemicals within the PHOP class with cancer data were generated. Even though these estimates quantify different kinds of cancer-risk for individual chemicals, the class-based approach ensured that these were considered and quantified at the same time.

8. Discussion

Here we describe a framework for class-based risk assessment (Figure 1). Before deploying this approach, it is necessary to perform class-based hazard assessment and class-based exposure assessment. Key considerations for data quality for exposure and hazard should be considered when thinking about how to parameterize the distributions. These considerations are described more fully in the CO-3 and Subtask 4A report. Additionally, this pipeline is responsive to new data as parameters can be altered based on the addition of new data. This means that as new data are collected, estimates of risk will continue to improve.

Additionally, the approach presented here is a hybrid approach between the traditional central tendency approach and a probabilistic approach. Incorporating probabilistic methods was of interest because these methods can help to address the different levels of uncertainty associated with different levels of data quality used to estimate exposure and hazard. We chose to incorporate probability distributions into traditional hazard metrics rather than use a HD_M^I given that class-based approaches need to be applied to data-poor chemicals, and the HD_M^I is quite data intensive. Additionally, using metrics that are traditionally used in hazard assessment carries with it familiarity, making it more accessible to risk managers and interpretation by the general public.

Lastly, our findings from our case study on the PHOPs demonstrate an approach to conducting class-based risk assessment. The confidence in and usefulness of the risk estimates are dependent upon the quantity and quality of the data available for each subclass member. Specifically, for all chemicals with a high degree of uncertainty (e.g., Tier II chemicals with less certain central tendency estimates), a high risk estimate was identified based on comparison of the full distributions for hazard and exposure. To determine which of these chemicals are higher priority, we compared the central estimates of hazard and exposure. We also suggested that more data would greatly improve hazard characterization. We demonstrated this by showing that even limited increases in amount and/or quality of data, related both to the TRV and exposure, result in better risk estimates particularly for Tier II chemicals. These findings suggest that undertaking studies that increase information about both toxicity thresholds and exposure estimates is of great utility.

In conclusion, the framework presented here is a fit-for-purpose method that can be utilized to provide estimates both for cancer and noncancer outcomes. Further, it is responsive to changes in data availability and new data can be easily integrated into the assessment paradigm to improve risk estimates. Finally, its utility has been demonstrated

by its application to the generally data-poor PHOPs chemical subclass where risk estimates were generated for 22 out of 29 subclass members.

References

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. *Regulatory Toxicology and Pharmacology*, *129*, 105094. https://doi.org/10.1016/j.yrtph.2021.105094

Aurisano, N., Jolliet, O., Chiu, W. A., Judson, R., Jang, S., Unnikrishnan, A., Kosnik, M. B., & Fantke, P. (2023). Probabilistic points of departure and reference doses for characterizing human noncancer and developmental/reproductive effects for 10,145 chemicals. *Environmental Health Perspectives*, *131*(3), 037016. https://doi.org/doi:10.1289/EHP11524

Bokkers, B. G. H., Mengelers, M. J., Bakker, M. I., Chiu, W. A., & Slob, W. (2017). APROBA-Plus: A probabilistic tool to evaluate and express uncertainty in hazard characterization and exposure assessment of substances. *Food Chem Toxicol*, *110*, 408-417. https://doi.org/10.1016/j.fct.2017.10.038

Chiu, W. A., Axelrad, D. A., Dalaijamts, C., Dockins, C., Shao, K., Shapiro, A. J., & Paoli, G. (2018). Beyond the RfD: Broad application of a probabilistic approach to improve chemical dose-response assessments for noncancer effects. *Environ Health Perspect*, *126*(6), 067009. https://doi.org/10.1289/ehp3368

Chiu, W. A., & Slob, W. (2015). A unified probabilistic framework for dose-response assessment of human health effects. *Environ Health Perspect*, *123*(12), 1241-1254. https://doi.org/10.1289/ehp.1409385

Dourson, M. L., Felter, S. P., & Robinson, D. (1996). Evolution of science-based uncertainty factors in noncancer risk assessment. *Regulatory Toxicology and Pharmacology*, *24*(2), 108-120. https://doi.org/10.1006/rtph.1996.0116

Fayers, P. (2011). Alphas, betas and skewy distributions: Two ways of getting the wrong answer. *Adv Health Sci Educ Theory Pract*, *16*(3), 291-296. https://doi.org/10.1007/s10459-011-9283-6

Flynn, M. R. (2004). The 4-parameter lognormal (SB) model of human exposure. *The Annals of Occupational Hygiene*, *48*(7), 617-622. https://doi.org/10.1093/annhyg/meh071

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous Univariate Distributions* (2nd ed., Vol. 1). Wiley.

Krithikadatta, J. (2014). Normal distribution. *J Conserv Dent*, *17*(1), 96-97. https://doi.org/10.4103/0972-0707.124171

Kvasnicka, J., Aurisano, N., von Borries, K., Lu, E.-H., Fantke, P., Jolliet, O., Wright, F. A., & Chiu, W. A. (2024). Two-stage machine learning-based approach to predict points of departure for human noncancer and developmental/reproductive effects. *Environmental Science & Technology*, *58*(35), 15638-15649. https://doi.org/10.1021/acs.est.4c00172

Limpert, E., Stahel, W. A., & Abbt, M. (2001). Log-normal distributions across the sciences: Keys and clues: On the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and

probability—normal or log-normal: That is the question. *BioScience*, *51*(5), 341-352. https://doi.org/10.1641/0006-3568(2001)051[0341:Lndats]2.0.Co;2

Silva, A. V., Ringblom, J., Lindh, C., Scott, K., Jakobsson, K., & Öberg, M. (2020). A probabilistic approach to evaluate the risk of decreased total triiodothyronine hormone levels following chronic exposure to PFOS and PFHxS via contaminated drinking water. *Environmental Health Perspectives*, *128*(7), 076001. https://doi.org/10.1289/EHP6654

U.S. Environmental Protection Agency (U.S. EPA). (2002). *A review of the reference dose and reference concentration process* (EPA/630/P-02/002F).

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

van der Voet, H., & Slob, W. (2007). Integration of probabilistic exposure assessment and probabilistic hazard characterization. *Risk Anal*, *27*(2), 351-371. https://doi.org/10.1111/j.1539-6924.2007.00887.x

World Health Organization (WHO). (2018). *Guidance document on evaluating and expressing uncertainty in hazard characterization– 2nd edition*. https://www.who.int/publications/i/item/9789241513548