Probabilistic reference and 10% effect concentrations for characterizing inhalation non-cancer and developmental/reproductive effects for 2,160 substances.

Nicolò Aurisano\textsuperscript{a}, Peter Fantke\textsuperscript{a*}, Weihsueh A. Chiu\textsuperscript{b}, Richard Judson\textsuperscript{c}, Suji Jang\textsuperscript{b}, Aswani Unnikrishnan\textsuperscript{c}, Olivier Jolliet\textsuperscript{a,d}

\textsuperscript{a} Quantitative Sustainability Assessment, Department of Environmental and Resource Engineering, Technical University of Denmark, Produktionstorvet 424, 2800 Kgs. Lyngby, Denmark
\textsuperscript{b} Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, Texas, USA
\textsuperscript{c} National Center for Computational Toxicology, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA
\textsuperscript{d} Department of Environmental Health Sciences, School of Public Health, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109, USA

\textsuperscript{*}Corresponding author: Tel.: +45 45254452, e-mail: pefan@dtu.dk

ORCID:
Nicolò Aurisano: https://orcid.org/0000-0003-3651-1307
Peter Fantke: https://orcid.org/0000-0001-7148-6982
Weihsueh A. Chiu: https://orcid.org/0000-0002-7575-2368
Richard Judson: https://orcid.org/0000-0002-2348-9633
Suji Jang: https://orcid.org/0000-0003-1424-2337
Aswani Unnikrishnan: https://orcid.org/0000-0001-9285-8714
Olivier Jolliet: https://orcid.org/0000-0001-6955-4210
Abstract

Chemical management and risk assessment frameworks rely on regulatory toxicity values, which are based on points of departure (POD) identified following dose-response assessment of available human epidemiological or experimental animal studies. Yet, due primarily to the lack of available data, regulatory toxicity values for inhalation exposure are available for less than 200 chemicals out of the tens of thousands currently commercialized. To address this gap, it is the aim of the present study to expand the coverage of chemicals for which inhalation toxicity values can be derived by applying a workflow to determine surrogate inhalation route PODs, and corresponding toxicity values, where regulatory assessments are lacking. We curated and selected in vivo data from the U.S. EPA’s Toxicity Value Database, and adjusted reported effect values to chronic human equivalent benchmark concentrations (BMCₜ) using the WHO/IPCS framework. We tested the hypothesis that the 25th percentile of the BMCₜ distribution (POD₀.25BMCₜ) for a chemical could serve as a surrogate POD by analyzing the correlation against available regulatory PODs (Q²≥0.76, RSE≤0.82 log₁₀ units). We derived from the curated dataset POD₀.25BMCₜ for 2,095 substances for general non-cancer effects and for 638 substances for reproductive/developmental effects, yielding a total substance coverage of 2,160. From these POD₀.25BMCₜ we estimated probabilistic reference concentrations (1% incidence level at 95% confidence) for application in risk assessment and human population effect concentrations (best estimate of 10% incidence level) for use in life cycle impact assessment and other comparative assessments. In providing inhalation PODs calibrated to regulatory values and deriving corresponding toxicity values for more than two thousand substances, we have substantially expanded by a factor 13 for general non-cancer and a factor 45 for reproductive/developmental effects the coverage of chemicals for which inhalation exposure can be addressed in various chemical assessment and management frameworks.
Highlights

- Curated inhalation experimental animal toxicity dataset with 15,219 records
- PODs calibrated to regulatory values for 2,160 substances
- Inhalation toxicity values for general non-cancer and reproductive/developmental effects
- Application in high-throughput risk screening, alternatives assessment, and LCIA
- Probabilistic RfCs and human population effect concentrations consistently derived
1. Introduction

Chemical assessment and management frameworks, including life cycle impact assessment (LCIA), comparative risk screening, and chemical substitution, evaluate potential risks and toxicological impacts from chemical exposures using chemical-specific points of departure (PODs) (Fantke et al., 2021, 2020b; Isaacs et al., 2022; Pham et al., 2019). These PODs represent the point on the dose-response curve typically used for low-dose extrapolation for risk assessment (Pradeep et al., 2020). If the available toxicity data are suitable for dose-response modeling, the statistically-derived benchmark concentration or dose lower confidence limit (BMCL) is typically modeled and considered as a candidate POD for toxicity value derivation; otherwise, the lowest-observed-adverse-effect concentration (LOAEC) or the no-observed-adverse-effect concentration (NOAEC) are used instead (Blessinger et al., 2020; Pham et al., 2020). In addition, many frameworks require PODs based on regulatory assessments and thus can be derived from a comprehensive and systematic dose-response assessment process of available toxicity studies. These include peer-reviewed human health toxicity values from, for example, the U.S. EPA’s Provisional Peer Reviewed Toxicity Values (PPRTV), the Office of Pesticide Programs (OPP), and the California EPA’s Office of Environmental Health Hazard Assessment (OEHHA). Yet, human health assessment relevant data sources currently only provide PODs for a small fraction of the tens of thousands of chemicals used worldwide (Jolliet et al., 2021; Judson et al., 2009; Wang et al., 2020) since conducting such assessments is highly data-, time-, and resource-intensive (Wignall et al., 2018).

The World Health Organization’s International Programme on Chemical Safety (WHO/IPCS) developed a consistent and transparent framework for dose-response assessment that results in derivation of reference doses (RfDs) and reference concentrations (RfCs) from probabilistic modeled PODs, for both health-based risk assessment as well as
comparative risk screening (Chiu et al., 2018; Chiu and Paoli, 2021; Chiu and Slob, 2015; WHO, 2014). In addition, for LCIA purposes, the WHO/IPCS probabilistic framework was adopted for deriving human dose-response factors for non-cancer endpoints, using population effect concentrations with an incidence response level $I = 10\%$ (Fantke et al., 2021). Even though this framework can be applied to derive both RfDs and RfCs, it has mainly been applied to the evaluation of health risks via the oral route of exposure. Specifically, Chiu et al. (2018) derived probabilistic RfDs for 608 substances with regulatory data, and only 1 probabilistic RfC was derived by Blessinger et al. (2020) for acrolein. Fantke et al. (2021) derived human population effect doses ($I = 10\%$) for 115 organic chemicals, and in a recent study, Aurisano et al. (2023) derived probabilistic RfDs and human population effect doses for 10,145 substances. However, no sets of human population effect concentrations for inhalation exposure have been derived yet, mainly due to the much lower data availability of inhalation toxicity studies compared to oral exposure (e.g., Wignall et al., 2018) as well as the low substance coverage across regulatory sources for inhalation exposure with RfCs available for $n < 200$ chemicals.

The availability of toxicity values for thousands of chemicals for inhalation exposure is nevertheless crucial, especially for comparing chemicals across exposure routes, as highlighted by recent high-throughput exposure and risk screening studies (Aurisano et al., 2022, 2021a; Huang et al., 2022), and for assessing chemicals in a variety of product applications where inhalation often is the predominant exposure route (e.g., volatile solvents in indoor paints). To address this need, we can take advantage of the increasing availability of experimental animal data housed in databases, such as the U.S. EPA’s Toxicity Value Database (ToxValDB), where $in\ vivo$ toxicity data are available for thousands of chemical substances (Judson, 2019; Nelms and Patlewicz, 2020).

In the present study, we propose to implement the probabilistic risk assessment workflow developed for oral exposures (Aurisano et al., 2023) to the derivation of inhalation...
PODs, probabilistic RfCs, and human population effect concentrations, using ToxValDB as starting point. We focus on the following four specific objectives:

(i) to compile from ToxValDB a harmonized dataset of inhalation exposure-response toxicity data covering multiple non-cancer endpoints,

(ii) to develop an approach to derive inhalation PODs from in vivo data by comparing curated toxicity data and available regulatory values,

(iii) to derive surrogate inhalation PODs with quantified uncertainties for a wide range of chemicals, differentiating between two main effect categories, and

(iv) to use the surrogate inhalation PODs to determine human population effect concentrations at 10% incidence response and probabilistic RfCs using the WHO/IPCS framework and compare the latter against available regulatory RfCs.

We consider two different effect categories: reproductive/developmental effects and non-reproductive/developmental effects (the latter hereafter referred to as “general non-cancer effects”). This choice is dictated by the large difference between these two effect categories in severity to affect human lifetime loss (Fantke et al., 2021; Huijbregts et al., 2005) and is fully consistent with previous work conducted on oral exposure (Aurisano et al., 2023). The provided set of inhalation PODs, as well as their corresponding probabilistic RfCs and human population effect concentrations, are suitable for implementation into various chemical assessment and management frameworks, including LCIA, chemical alternatives assessment, and high-throughput risk screening for chemical substitution and prioritization (Fantke et al., 2021, 2016; Jolliet et al., 2015).
2. Methods

We propose a workflow that aims to derive surrogate PODs in a quantitative high-throughput approach, building on the assumption that for substances for which regulatory PODs are not available, *in vivo* toxicity data can be used to estimate a POD that most closely mimics one that would be selected in a regulatory assessment context (Aurisano et al., 2023).

Since regulatory PODs intend to be protective of all potential adverse effects, surrogate POD should be at the lower end of the distribution of available toxicity values (Paul Friedman et al., 2020), following careful data curation where needed (Fantke et al., 2020a). Indeed, for a given chemical, multiple studies might be available reporting various effect-level types (e.g., BMC, NOAEC), observed critical effects (e.g., mortality, developmental), and tested species (e.g., rabbit, mice), with the consequence of reported effect-level values (i.e., experimental values of toxicity from individual studies) varying over orders of magnitude (Chiu et al., 2018; Pradeep et al., 2020; Zeise et al., 2013). Thus, for a correct development of the above-mentioned approach, the numerous challenges in using experimental animal databases such as ToxValDB need to be overcome by applying methods for data selection and harmonization for human toxicity information (Aurisano et al., 2021b; Smith et al., 2019), similar to those proposed for physico-chemical properties and freshwater ecotoxicity information (Aurisano et al., 2019; Aurisano and Fantke, 2022).

To compile a harmonized dataset of inhalation exposure-response toxicity data and derive related probabilistic POD values, our proposed workflow, adapted from Aurisano et al. (2023), is composed of six main stages (Fig. 1). The first stage is the curation and selection of the relevant experimental animal toxicity data, and their allocation in one of the two considered health effect categories, i.e., general non-cancer and reproductive/developmental effects (Fig. 1A). Next, we compiled a dataset of reported POD values for inhalation from various regulatory sources (POD\textsubscript{reg}), again distinguishing between the two considered effect
categories (Fig. 1B). Third, we investigated the potential correlation between curated and selected inhalation experimental animal toxicity data and the collected POD$_{\text{reg}}$ (Fig. 1C).

Based on this analysis, we then systematically determine a surrogate POD for each substance in the two curated datasets (Fig. 1D) and characterize the uncertainty around the determined value (Fig. 1E). Finally, we derived probabilistic RfCs and human population effect concentrations ($I = 10\%$) with related uncertainty using the WHO/IPCS framework (Fig. 1F).

The following sections present each of these stages in further detail.

**Fig. 1.** Overview of the proposed workflow: (A) data curation and selection applied to the collected *in vivo* data from ToxValDB, (B) collection and extrapolation of regulatory PODs, (C) analysis of the correlation between ToxValDB and regulatory POD data, (D) systematic derivation of inhalation PODs from the curated datasets, differentiating between general non-cancer (non-reproductive/developmental) and reproductive/developmental effects, (E) quantification of the substance-specific uncertainty of the derived PODs, and (F) derivation of probabilistic reference concentrations (RfC) and human population effect concentrations at 10% incidence level. The workflow is adapted from Aurisano et al. (2023).
2.1. Description of the in vivo input data set

We used the U.S. EPA’s Toxicity Value Database (ToxValDB) as a source for the experimental animal toxicity data. ToxValDB is a database collecting toxicity data from more than forty publicly available sources (Judson, 2019), including—among others—the Toxicity Reference Database (ToxRefDB) (Martin et al., 2009; Watford et al., 2019), the Integrated Risk Information System (IRIS), the Provisional Peer-Reviewed Toxicity Value (PPRTV), the European Chemicals Agency’s eChem Portal and the European Food Safety Authority’s Chemical Hazards Database. ToxValDB is accessible at https://comptox.epa.gov/dashboard via the CompTox Chemicals Dashboard (Williams et al., 2021, 2017). The entire ToxValDB was downloaded for subsequent filtering and processing. The reported data used in the present study were retrieved in March 2021, including all major updates until that date.

2.2. Input data curation and selection

ToxValDB is reporting toxicity data from diverse sources. Such data are often of varying quality, developed for specific applications using different methods, following different guidelines, and available in formats that are not always easily integrated (Aurisano and Fantke, 2022; Kosnik et al., 2022; Li et al., 2022). For this reason, the collected toxicity data went through a curation and selection process based on criteria derived from the WHO/IPCS framework to simplify the data processing and select only specific records (e.g., excluding all records not targeting inhalation exposure) (Fig. 1A) (Chiu et al., 2018; Chiu and Slob, 2015; WHO, 2014). For example, via this process, the reported effect value of each record was extrapolated to a chronic human equivalent benchmark concentration (BMC$_h$) with a consistent effect value unit expressed in mg/m$^3$. The curation and selection process is presented in detail in the Supplementary Material (SM, Section 1). After the curation and selection process, the curated data were split into two distinct datasets covering general non-
cancer and reproductive/developmental effects based on each record study type and reported critical effects (Fig. 1A).

2.2, Regulatory data

To build a regulatory dataset for inhalation exposure, we used as a starting point the work conducted by Wignall et al. (2018, 2014) and their published database, collecting peer-reviewed toxicity values reported in various public sources, such as IRIS. In addition, the collected toxicity values were cross-checked with the November 2019 release of the U.S. EPA Regional Screening Levels (U.S. EPA 2019) and incorporated new substances with available PODs. In our study, a POD\text{reg} is defined as a NOAEC, LOAEC, or BMCL associated with a reported reference concentration. As done for the in vivo input data set, POD\text{reg} values were extrapolated to chronic human equivalent benchmark concentrations (POD\text{reg,BMC\text{h}}), distinguishing between the two considered effect categories (Fig. 1B).

2.3, Approach for deriving inhalation PODs

To derive inhalation PODs, we followed and tested the hypothesis that the lower end of the effect values distribution is a suitable proxy for POD\text{reg} across different chemicals (Paul Friedman et al., 2020). Thus, for chemicals for which both POD\text{reg,BMC\text{h}} and in vivo data were available, we assumed a lognormal distribution across BMC\text{h} and derived an inhalation POD as the 25\textsuperscript{th} %-ile of the fitted lognormal distribution (i.e., POD\textsubscript{p25BMC\text{h}}). For consistency, we directly selected and tested the 25\textsuperscript{th} %-ile, since this percentile was identified as the most suitable for estimating a POD in a previous study focusing on oral exposure (Aurisano et al., 2023). To test the appropriateness of the selected percentile for inhalation toxicity data, the resulting POD\textsubscript{p25BMC\text{h}} were compared against the respective available POD\textsubscript{reg,BMC\text{h}} (Fig. 1C).
The comparison was performed distinguishing between general non-cancer and reproductive/developmental effects.

The two function moments used for fitting the lognormal distribution to $BMC_h$ values are mu ($\mu$) and sigma ($\sigma$), where $\mu$ represents the log-scale population median, and $\sigma$ is the standard deviation of the available effect values for a substance (Posthuma et al., 2019).

Following our previous approach, we differentiate between substances with $\geq$10 records available (data-rich chemicals) and $<$10 records available (data-poor chemicals). For fitting the lognormal distribution, $\mu$ was calculated from the available $BMC_h$ for all substances, and $\sigma$ was calculated from the available $BMC_h$ only for data-rich chemicals, whereas for data-poor chemicals, we applied a fixed standard deviation ($\sigma_{\text{fixed}}$). The $\sigma_{\text{fixed}}$ is derived from the average across $\sigma$ of data-rich chemicals. This choice was based on the highly unstable estimates of $\sigma$ for data-poor chemicals; thus, an average-shaped distribution was applied instead of relying on the few available effect values to avoid bias introduced by too few data points.

2.4. Deriving PODs per substance

After confirming that the 25th %-ile of the fitted lognormal distribution is suitable to derive surrogates of regulatory values also for inhalation exposure, we systematically estimated $\mu$ and $\sigma$ for each substance from the available records and then derived related POD$_{25BMC_h}$ (Fig. 1D). This was done separately for the two datasets. For substances with curated toxicity records available in both datasets, two distinct POD$_{25BMC_h}$ values were derived separately, one for general non-cancer effects and one for reproductive/developmental effects.
2.5. Quantifying uncertainty around the derived points of departure

We characterized the uncertainty around the derived POD$_{p25BMC_h}$ from the residual standard error (RSE) of the comparison carried out between POD$_{reg}$ and POD$_{p25BMC_h}$ (Fig. 1E). This uncertainty is expressed as the squared geometric standard deviation (GSD$_{p25→reg}$). GSD$_{p25→reg}$ describes the spread of data around their geometric mean, and more specifically indicates that 95% of the data fall within the range of POD$_{p25BMC_h}/$GSD$_{p25→reg}$ and POD$_{p25BMC_h} \times$ GSD$_{p25→reg}²$ (Fantke et al., 2012; Hong et al., 2010; Slob, 1994; Stylianou et al., 2021). For example, a GSD$_{p25→reg}²$ = 10 indicates that the 95% confidence interval of POD$_{p25BMC_h}$ span over two orders of magnitude.

2.6. Deriving probabilistic reference and human effect concentrations

For each of the derived POD$_{p25BMC_h}$ we calculated probabilistic reference concentrations (RfCs) and human population effect concentrations (at incidence level $I = 10\%$) for application in risk screening and LCIA or other comparative assessments, respectively (Fig. 1F).

We implemented the approximate approach by Chiu et al. (2018) for the calculation of the probabilistic RfCs. Briefly, probabilistic RfCs were derived from the lower 95% confidence bound of HC$_M^{1\%}$, i.e., the daily human concentration at which 1% of the population shows a level of effect $M$ corresponding to the effect level type reported in the database and the type of endpoint (e.g., continuous, quantal deterministic, or stochastic deterministic). For each chemical, HC$_M^{1\%}$ was calculated from POD$_{p25BMC_h}$ by dividing it by an extrapolation factor of $P50 = 9.7$ to account for variability in sensitivity between the median and the 1st %-ile of human population response (WHO, 2014). The lower 95% confidence bound of HC$_M^{1\%}$ was derived by combining probabilistically GSD$_{p25→reg}$ and the uncertainty distribution (i.e., $P95/P50 = GSD_h^{1.65} = 4.3$) assigned to the human variability at
1\textsuperscript{st} %-ile (WHO, 2014). Note that $\text{GSD}^{1.65} = (\text{GSD}^2)^{1.65}$, indicating a one-sided (i.e., lower) confidence interval range. We then compared the derived lower 95\% confidence bound of $\text{HC}_M^{1\%}$ against the related regulatory RfC (if available) to investigate the potential influence of the database uncertainty factor ($UF_d$). $UF_d$ is applied when deriving regulatory RfCs but it is not directly included in the WHO/IPCS framework (Chiu et al., 2018). This comparison will help us understand whether the derived toxicity values are consistent with regulatory RfCs and identify potential biases.

For LCIA purposes, recent updates of the globally recommended approach for deriving human dose-response factors for non-cancer endpoints proposed using the human population effect concentration with an incidence response level $I = 10\%$ (Fantke et al., 2021). Hence, we also derived $\text{HC}_M^{10\%}$ from the provided $\text{POD}_{p25\text{BMCh}}$, accounting for the human variability between 50\% and 10\% incidence level by dividing the $\text{POD}_{p25\text{BMCh}}$ by the best estimate factor of $P50 = 3.49$ (WHO, 2014). $\text{HC}_M^{10\%}$ is also referred to in LCIA as EC10 (e.g. Fantke et al. 2021). $\text{HC}_M^{10\%}$ related uncertainty was calculated by combining probabilistically $\text{GSD}_{p25\rightarrow\text{reg}}^2$ of $\text{POD}_{p25\text{BMCh}}$ and the uncertainty distribution assigned to the human variability at 10\textsuperscript{th} %-ile, i.e., $P97.5/P50 = 2.67$ (WHO, 2014).

Finally, the derived RfCs and $\text{HC}_M^{10\%}$s were compared against the results of a previous study focusing on oral toxicity data to investigate potential trends across exposure routes (Aurisano et al., 2023). In addition, we matched our results with exposure estimates from the SEEM meta-model (Ring et al., 2019) to examine the fraction of assessed substances with population median chemical intake rates above our derived probabilistic RfCs.
2.7. Data analysis

The gathered toxicity data from ToxValDB were processed using the open source statistical software R version 3.6.1 (R Core Team, 2020), and the package “ggplot2” was used to generate all results figures (Wickham, 2016).

3. Results

3.1. Curated toxicity test datasets

The downloaded version of ToxValDB listed 427,506 records for more than 30,000 chemicals and reported a wide range of toxicity information for different species, effects, and exposure routes. Following the described data curation and selection process applied to ToxValDB, the resulting curated dataset compiled inhalation toxicity information for 2,160 substances covered by 15,219 records. We split this curated dataset into two distinct datasets covering the two considered effect categories, i.e., 2,095 substances (11,767 records) for general non-cancer and 638 substances (3,452 records) for reproductive/developmental effects. Records were available in both datasets for 573 substances. The curated datasets are provided in the SM, separately for general non-cancer effects (Table S1) and for reproductive/developmental effects (Table S2).

Fig. 2 summarizes the statistics of the two curated datasets, presenting results for the general non-cancer and reproductive/developmental effects on the left and right, respectively. More specifically, Fig. 2A-B presents the distribution of the extrapolated effects values (BMCₜ) across all records, differentiating between underlying effect-level and study types information. NOAEC is the most common effect-level type in both datasets (~75%), followed by LOAEC (~24%) and BMCL (~1%). For the study types distribution across records in the general non-cancer effects dataset (Fig. 2A), 33%, 58%, and 8% of the records were reported as chronic, subchronic, and subacute, respectively. This differentiation is not performed for
records covering reproductive/developmental effects as we assumed that the relevant window of susceptibility was always covered for this effect category.

Fig. 2C-D gives an overview of the number of curated records available per substance in the two datasets, highlighting a limited number of records available for most substances. For example, only one or two records are available for around half of the covered substances, and for both datasets, only 15% of all substances are considered as data-rich chemicals, for which 10 or more records are available (Fig. 2C-D right of the red dashed line). Concerning the tested species, the majority of records report rat followed by mouse in both datasets (Fig. 2E-F). The statistics of the curated datasets are in line with data for oral exposure (Aurisano et al., 2023) and with other studies using the same database to develop quantitative structure-activity relationship (QSAR) and new approach methodologies (NAMs) models (e.g., Pradeep et al., 2020; Pham et al., 2020).
Fig. 2. Distribution of the effect values (chronic human equivalent benchmark concentration, $BMC_h$) across curated records and underlying effect-level and study types for the general non-cancer effects (A) and for the reproductive/developmental effects dataset (B). Number of records for each chemical in the general non-cancer effects dataset (C) and in the reproductive/developmental effects dataset (D), and distribution of the tested species in the
general non-cancer effects (E) and reproductive/developmental effects dataset (F). By default, not reported data are treated similarly to rat data – the dominant species in available tests.

Fig. 3A-B visualizes the extrapolated effect values (BMC<sub>E</sub>) for all the records in the two datasets, differentiating between originally reported effect-level types. In addition, collected POD<sub>reg,BMC</sub> are represented as black triangles for the substances for which a regulatory RfC was available. Across the records in the general non-cancer effects, BMC<sub>E</sub> values range from $2.5 \times 10^{-4}$ to $3.7 \times 10^{7}$ mg/m<sup>3</sup> with a median value of 713 mg/m<sup>3</sup> (Fig. 3A), while in the reproductive/developmental effects dataset, they range from $3.5 \times 10^{-3}$ to $3.4 \times 10^{6}$ mg/m<sup>3</sup> with a median value of 7875 mg/m<sup>3</sup> (Fig. 3B). The BMC<sub>E</sub> values across the records available for the same substance can span over several orders of magnitude. In general, this variability might be related to different factors, such as different critical effects studied or species tested in various environmental conditions (i.e., biological variability), as well as systematic errors, including errors in the measurements, differences in experimental protocols, or measurement tools (Browne et al., 2015; Kleinstreuer et al., 2016; Pradeep et al., 2020).
Fig. 3. Inhalation effect values adjusted to chronic human equivalent benchmark concentrations (BMC_h) for all the records in the general non-cancer effects (A) and reproductive/developmental effects (B) dataset, together with the corresponding POD_reg (black ▼, when available) and derived PODs (POD_{p25BMC_h}, grey data points). Chemicals are ranked by derived POD_{p25BMC_h}.

3.2. Points of departure and comparison with regulatory toxicity values

We derived inhalation PODs for all the substances for which toxicity information were available in the curated datasets as the 25th %-ile of the fitted lognormal distribution to the available records per substance (POD_{p25BMC_h}). To fit the lognormal distribution for data-rich chemicals (≥10 records), we directly used the available effect values (BMC_h) to derive a chemical-specific standard deviation, assuming that the available records are sufficient to represent and cover different potential adverse effects. In contrast, we derived an average standard deviation across data-rich chemicals of log_{10} \sigma_{Rxed} = 0.6 for both general non-cancer and reproductive/developmental effects (Fig. S1). We then applied \sigma_{fixed} to all data-
poor chemicals (<10 records) for calculating POD_{p_{25BMCh}}, assuming that the chemical-specific σ are not reliable for data-poor chemicals.

Following this approach, we systematically derived surrogate POD_{p_{25BMCh}} for 2,095 substances for general non-cancer effects and for 638 substances for reproductive/developmental effects, yielding a total substance coverage of 2,160. For 573 substances, we derived two distinct POD_{p_{25BMCh}} as toxicity values were available for both effect categories. The derived POD_{p_{25BMCh}} are presented in Fig. 3 (grey triangles), ranging from $9.8 \times 10^{-5}$ to $1.5 \times 10^{7}$ mg/m$^3$ for general non-cancer effects, with a median POD_{p_{25BMCh}} value of 117 mg/m$^3$. For reproductive/developmental effects, POD_{p_{25BMCh}} are on average more than one order of magnitude higher (average value of 1747 mg/m$^3$), ranging from $1.9 \times 10^{-3}$ to $9.4 \times 10^{5}$ mg/m$^3$. The substances with the highest toxicity (lowest POD_{p_{25BMCh}} values) include the insecticides disulfoton (CAS: 298-04-4) and parathion (CAS: 56-38-2), the chemical warfare agent sulfur mustard (CAS: 505-60-2), the metal beryllium (CAS: 7440-41-7), and melamine (CAS: 108-78-1). All derived PODs and the number of underlying effect values are provided in the SM (Table S3).

Fig. 4 compares the derived POD_{p_{25BMCh}} against the respective available POD_{regBMCh} for both studied effects, further differentiating between data-rich and data-poor chemicals. The comparison was carried out for a total of $n = 168$ substances with available regulatory inhalation data, i.e., $n = 154$ substances for general non-cancer (Fig. S2A) and $n = 14$ for reproductive/developmental effects (Fig. S2B). The estimated POD_{p_{25BMCh}} values correlate well with the derived POD_{p_{25BMCh}}, with a coefficient of "goodness of prediction" of $Q^2 = 1 - \frac{PRESS}{TSS} = 0.76$ and a residual standard error on the log of RSE = 0.82 evaluated on log-scale for the 1:1 line. PRESS is the Predictive Error Sum of Squares, that is the sum of the squares of the differences (residuals) between the predicted and regulatory values and TSS is the Total Sum of Square (Gramatica, 2013). These results support our choice of selecting the
25<sup>th</sup> %-ile of the fitted lognormal distribution as a surrogate for regulatory data and are consistent with the best-suited percentile identified for oral exposure (Aurisano et al., 2023). The collected POD<sub>reg,BMC<sub>h</sub></sub> are summarized in the SM for both general non-cancer and reproductive/developmental effects (Table S4).

Fig. 4. Comparison between estimated POD<sub>p25BMC<sub>h</sub></sub> and POD<sub>reg,BMC<sub>h</sub></sub> for general non-cancer (▲) and for reproductive/developmental effects (■), differentiating between data-rich (light blue, ≥10 records) and data-poor chemicals (dark blue, <10 records). Coefficient of "goodness of prediction" (Q<sup>2</sup>) and residual standard error (RSE) are evaluated on log-scale for the 1:1 line (black dashed line).

The uncertainty factor of GSD<sup>2</sup><sub>p25→reg</sub> = 10<sup>2.02×0.82</sup> = 45 for both general non-cancer and reproductive/developmental effects is derived from this comparison with regulatory values, to reflect the use of POD<sub>p25BMD</sub> as a suitable approximation of POD<sub>reg</sub>. Assigning a factor GSD<sup>2</sup><sub>p25→reg</sub> = 45 indicates that 95% CI of the derived POD<sub>p25BMC<sub>h</sub></sub> fall in the range
between \( \text{POD}_{p25\text{BMC}_h}/45 \) and \( \text{POD}_{p25\text{BMC}_h} \times 45 \). The limited amount of chemicals considered in the comparison against regulatory values for reproductive/developmental effects precluded the characterization of an effect-specific uncertainty; hence, the same uncertainty as for general non-cancer effects is used by default. Fig. S3 presents the distributions of the derived \( \text{POD}_{p25\text{BMD}_h} \) together with their characterized 95% CI.

### 3.3. Probabilistic reference concentrations and human population effect concentrations

From the provided \( \text{POD}_{p25\text{BMC}_h} \), we derived probabilistic reference concentrations (RfCs) and human population effect concentrations (HCM\(^{10\%}\)), following the WHO/IPCS framework (WHO, 2014). Since WHO (2014) focuses on endpoint-specific uncertainties and RfCs, an additional database uncertainty factor \( (U_F_d) \) needed to be included when deriving probabilistic RfCs that are comparable to and consistent with regulatory RfCs. This so-called database deficiency factor accounts for data gaps and is typically equal to 1, 3, and 10 as a function of the data coverage for different endpoints (OEHHA, 2008).

To derive probabilistic RfCs, the following additional \( U_F_d \) were thus applied: the lower 95% confidence bound of HCM\(^{1\%}\) is divided by \( U_F_d = 10 \) for substances with very poor data availability \( (n \leq 3 \text{ records}) \), by \( U_F_d = 3 \) for substances with intermediary data availability \( (3 < n < 10 \text{ records}) \), and by \( U_F_d = 1 \) for data-rich substances \( (n \geq 10 \text{ records}) \). For data-rich chemicals, the probabilistic RfC value is thus equal to the lower 95% confidence bound of HCM\(^{1\%}\). The derived probabilistic RfCs show a good correlation with the regulatory RfCs with a \( Q^2 = 0.59 \) and RSE = 1.11 evaluated on log-scale for the 1:1 line (Fig. S4B). In contrast, neglecting \( U_F_d \) would lead to a systematic overestimation of the RfCs (Fig. S4A, \( Q^2 = 0.54 \), RSE = 1.18). Following the \( U_F_d \)-complemented approach,
probabilistic RfCs for \( n = 2,169 \) substances were derived, differentiating between general non-cancer and reproductive/developmental effects.

For the same substances and the two effect categories, we also derived best estimates of \( HC_{M,10\%} = HC_{M,50\%}/3.49 = POD_{p25BMCMa}/3.49 \) with their uncertainties for application in LCIA as recommended by Fantke et al. (2021). We obtained \( HC_{M,10\%} \) for 2,095 substances for general non-cancer effects and for 638 substances for reproductive/developmental effects. The associated uncertainty characterized by combining probabilistically \( GSD_{p25-reg}^2 \) and the uncertainty distribution assigned to the human variability at 10\%-ile, is equal to \( GSD_{HC_{M,10\%}}^2 = 51 \) (Fig. S5). The derived probabilistic RfCs and \( HC_{M,10\%} \)s with related uncertainties are provided in the SM (Table S3).

### 3.4. Comparison of toxicity value ranges across effect categories and exposure routes

Fig. 5 summarizes the derived toxicity values for inhalation and compares their ranges against results for oral exposure provided by Aurisano et al. (2023). Fig. 5A-B presents ranges of inhalation \( HC_{M,10\%} \) and oral \( HD_{M,10\%} \), while Fig. 5C-D presents RfCs and RfDs, separately for the two effect categories. For each combination (e.g., inhalation \( HC_{M,10\%} \) for general non-cancer effects), regulatory values are presented first (darker color), followed by probabilistic values for the same chemicals for which regulatory assessments were available, and finally by the probabilistic values for all covered substances. Regulatory \( HC_{M,10\%} \) and \( HD_{M,10\%} \) are also estimated from \( POD_{reg} \) following the WHO/IPCS framework. In addition, for this comparison, RfCs and \( HC_{M,10\%} \) were converted into a consistent unit of mg/kg-d, assuming an average individual human breathing rate of 13 m\(^3\)/d and body weight of 70 kg.

Fig. 5 confirms that considering chemicals for which regulatory assessments were available, the ranges of derived probabilistic values are well in line with regulatory values across different toxicity values, exposure routes, and effects considered. On the other hand,
when considering all chemicals, the median of the probabilistic toxicity values is higher than the available regulatory values in the majority of the cases. This trend is linked to our probabilistic results covering thousands of substances while the regulatory values only cover a few hundred substances, and suggests that regulatory values tend to be selected among the most toxic substances. We do not observe any trends in ranges across exposure routes.

![Graph showing comparison between derived probabilistic reference concentrations (RfCs) and human population effect concentrations (I = 10%, HC_{M10%}) and oral toxicity values covering probabilistic reference doses (RfDs) and human population effect doses (I = 10%, HD_{M10%})](image)

**Fig. 5.** Comparison between the derived probabilistic reference concentrations (RfCs) and human population effect concentrations (I = 10%, HC_{M10%}) and oral toxicity values covering probabilistic reference doses (RfDs) and human population effect doses (I = 10%, HD_{M10%}) (Aurisano et al., 2023). Probabilistic RfCs and HC_{M10%} were converted to doses with a
consistent unit of mg/kg-d, assuming an average individual human breathing rate of 13 m$^3$/d and body weight of 70 kg. $n$ represents the number of chemicals covered. Reg.: regulatory values; POD$_{p25}$ With reg.: probabilistic values for the same chemicals for which regulatory assessments were available; POD$_{p25}$ All: probabilistic values for all covered substances.

Finally, we compared the derived probabilistic RfCs with population median chemical intake rates estimated via the SEEM meta-model, available for around half of the considered substances (Ring et al., 2019). We identified 33 substances against only 3 substances for oral exposure (Aurisano et al., 2023) despite a five-fold lower number of inhalation data, for which exposure best estimates are higher than derived probabilistic RfCs, highlighting a high potential risk (Fig. S6). When considering the upper 95% confidence bound, median intake rates are 100 times higher than doses calculated from probabilistic RfCs for around 5% (2% for oral exposure) of the substances for which SEEM intake rates are available, that is substances that deserve further scrutiny in priority.

4. Discussion

4.1. Applicability of the derived toxicity values

This study expanded by a factor 13 the coverage of chemicals for which inhalation toxicity values can be derived for general non-cancer effects, and by a factor 45 for reproductive/developmental effects. Combined with our previous effort focused on oral exposure, the presented approach provides a solid basis for consistently assessing toxicity effects across these two exposure routes for thousands of chemicals in various impact assessment and risk screening contexts.

The provided human population effect concentrations with an incidence response level $I = 10\%$ (HC$_M^{10\%}$) can be implemented in LCIA to derive human toxicity effect factors with
direct application in the scientific consensus model USEtox (Hauschild et al., 2008; Rosenbaum et al., 2011, 2008). In the current version of USEtox, human toxicity effect factors covering non-cancer toxicity were only available for less than 500 chemicals, of which only one-tenth is derived from inhalation toxicity data. Once implemented, the provided HC$_{M10\%}$s will increase the chemical coverage for inhalation by a factor of forty. In addition, by providing HC$_{M10\%}$s specific to general non-cancer effects and reproductive/developmental effects, these will be able to reflect the difference in severity when evaluating disability-adjusted life years (DALY) related to chemical exposure (i.e., 2.4 DALY/incidence for general non-cancer effects vs. 44.1 DALY/incidence for reproductive/developmental effects) (Fanteke et al., 2021; Huijbregts et al., 2005).

The provided probabilistic RfCs find direct application to support high-throughput risk screening studies. In such studies, hundreds (if not thousands) of chemicals are assessed in terms of multi-pathway exposure and related effects on humans. Thus the availability of toxicity information is a key factor when, e.g., evaluating exposures and identifying chemicals of concern and potential alternatives to harmful chemicals present in consumer products (Aurisano et al., 2022, 2021a; Huang et al., 2022).

Finally, by estimating surrogate PODs calibrated to regulatory values and deriving corresponding toxicity values also for chemicals with a limited amount of toxicity data available (data-poor chemicals), our results support the work of health assessors at multiple levels, including the cases of chemicals of potential concern not yet tested or reviewed (Wignall et al., 2018). While our results are primarily applicable at screening level and cannot substitute the rigorous assessments of chemicals potentially of concern, they constitute a useful dataset to train in-silico approaches beyond the restricted availability of regulatory values.
4.2. Limitations of the proposed workflow

The presented workflow also comes with limitations. First, the provided PODs (and related toxicity values) are based for 85% of the covered substances on less than ten curated records. For these data-poor chemicals, there is the risk of potentially missing critical effects not covered by the considered studies and thus underestimating the toxicity of specific substances. To minimize and mitigate this risk when deriving PODs for these substances, we fitted a lognormal distribution with a predefined average shape and a fixed standard deviation of $\log_{10}\sigma_{\text{fixed}} = 0.6$ for both general non-cancer and reproductive/developmental effects (value based on data-rich chemicals). Nevertheless, fitting a chemical-specific distribution based on a set of experiments to derive $POD_{p25\text{BMC}_{h}}$ would still be preferred and more accurate. In our approach, data richness is nevertheless considered, but in a simplified way when deriving probabilistic RfCs, where different database uncertainty factors ($UF_d$) are applied deterministically to the lower 95% confidence bound of $HC_{M1\%}$ based on the data availability.

Second, the characterized uncertainty for each POD, expressed as squared geometric standard deviation ($GSD_{p25\rightarrow\text{reg}}^2$), is limited to the uncertainty around the derived POD and directly reflects the use of $POD_{p25\text{BMC}_{h}}$ as a suitable approximation of regulatory values ($POD_{\text{regBMC}_{h}}$). The limited availability of reproductive/developmental effects data precluded the possibility of deriving an effect-specific $GSD_{p25\rightarrow\text{reg}}^2$ instead of a generic uncertainty applied to $POD_{p25\text{BMC}_{h}}$ for both effects. The same is valid for the uncertainty around the derived reproductive/developmental $HC_{M10\%}$ as well as the uncertainty used to define the lower 95% confidence bound of $HC_{M1\%}$.

Third, we did not evaluate the best-fit percentile of the fitted lognormal distribution when comparing against regulatory values ($POD_{\text{regBMC}_{h}}$), but for consistency, directly implemented the 25th %-ile identified in the study for oral exposure as the best approximation...
for both disease categories (Aurisano et al., 2023). However, the observed good correlation
between the derived POD\(_{p25\text{BMC}h}\) against the respective POD\(_{\text{reg.BMC}h}\) support this choice.
Nevertheless, considering the limited amount of substances for which the comparison with
inhalation regulatory toxicity values was carried out for reproductive/developmental effects, it
is necessary to verify in future efforts that the 25\(^{\text{th}}\) %-ile of the fitted lognormal distribution
(POD\(_{p25\text{BMC}h}\)) is the best approximation of POD\(_{\text{reg.BMC}h}\).

Finally, we acknowledge that in our workflow, there is an intrinsic limitation related
to predicting a toxicity value from \textit{in vivo} data. More specifically, even if starting from the
same underlying toxicity dataset, risk estimates can vary across regulatory settings despite the
rigorous scientific judgment involved (Wignall et al., 2018; NRC, 2009).

4.3. Future research needs

Future research should focus on further increasing the exposure route coverage,
prioritizing dermal exposure. Even though the toxicity data availability and related chemical
coverage will be lower for other exposure routes, such as dermal (most likely a few hundreds
of chemicals), route-specific toxicity data are key for assessing chemicals in specific product
applications (e.g., dermal exposure for personal care products) and for frameworks comparing
toxicity effects from chemical exposures across exposure routes. With that, the three main
exposure routes would be covered for the chemicals for which toxicity data are currently
available.

Similarly, in our work, we differentiated between reproductive/developmental effects
and general non-cancer effects due to the difference in the severity of these two disease
categories to affect human lifetime loss (Fantke et al., 2021; Huijbregts et al., 2005). Future
work should further increase this differentiation and provide more critical effect-specific
PODs (and related HC\textsubscript{M}$^{10\%}$ and probabilistic RfCs). Highly relevant critical effects include, for example, endocrine disruption (Emara et al., 2021).

Finally, given the large number of new and existing substances requiring assessment and management, there is a pressing need for cost-effective and rapid non-animal alternatives (Mansouri et al., 2021). In answer to this, the curated dataset compiling inhalation toxicity information provided in this study can be used in future research for training in silico, machine learning-based methods (e.g., random forest algorithms) to construct QSAR models for predicting PODs for substance lacking \textit{in vivo} data (Hou et al., 2020; Wignall et al., 2018). This would cover an even broader range of chemical substances.

5. Conclusions

This study provides inhalation PODs in line with regulatory values and their corresponding probabilistic RfCs and human population effect concentrations ($I = 10\%$) for more than two thousand chemicals, which drastically increases the chemical coverage for inhalation exposure as compared to previous studies. These results, combined with effect results for oral exposure, deliver a solid basis for consistently assessing toxicity effects and impacts of chemical exposures for thousands of chemicals in various chemical assessment and management frameworks, including LCIA, chemical alternatives assessment, and high-throughput risk screening for chemical substitution and prioritization.

Acknowledgments

This research was funded in part, by grants P42 ES027704 and P30 ES029067 from the National Institute of Environmental Health Sciences. This work was supported by the ‘Global Best Practices on Emerging Chemical Policy Issues of Concern under U.N. Environment’s Strategic Approach to International Chemicals Management (SAICM)’ (GEF
project ID 9771, grant no. S1-32GFL-000632), and by the ‘Safe and Efficient Chemistry by Design (SafeChem)’ project funded by the Swedish Foundation for Strategic Environmental Research (grant no. DIA 2018/11).

Conflict of Interest

The authors declare no conflict of interest.

Disclaimer

Dr. Jolliet discloses his role as a member of the USEtox Center scientific advisory board and chair of the project on Global guidance for Life Cycle Impact Assessment a project supported by the Life Cycle Initiative, hosted at UN-environment. The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.
References


Browne, P., Judson, R.S., Casey, W.M., Kleinstreuer, N.C., Thomas, R.S., 2015. Screening


Pradeep, P., Paul Friedman, K., Judson, R., 2020. Structure-based QSAR models to predict

https://doi.org/https://doi.org/10.1016/j.comtox.2020.100139


Foundation for Statistical Computing, Vienna, Austria.


https://doi.org/10.1038/s41370-019-0165-y


Williams, A.J., Grulke, C.M., Edwards, J., McEachran, A.D., Mansouri, K., Baker, N.C.,
