



Review article

Critical review and probabilistic health hazard assessment of cleaning product ingredients in all-purpose cleaners, dish care products, and laundry care products

Zhen Wang^{a,b}, Dan Dinh^a, W. Casan Scott^a, E. Spencer Williams^a, Michael Ciarlo^c, Paul DeLeo^d, Bryan W. Brooks^{a,b,e,*}

^a Department of Environmental Science, Baylor University, Waco, TX, USA

^b School of Environment, Jinan University, Guangzhou, China

^c EA Engineering, Science & Technology, Inc., Baltimore, MD, USA

^d American Cleaning Institute, Washington, DC, USA

^e Institute of Biomedical Studies, Baylor University, Waco, TX, USA



ARTICLE INFO

Handling Editor: Robert Letcher

Keywords:

Cleaning product use category
 Probabilistic health hazard assessment
 Chemical toxicity distribution
 Threshold of toxicological concern
 Uncertainty factor
 Acute-to-chronic ratio

ABSTRACT

Though numerous chemical ingredients are used in cleaning products, empirical mammalian toxicology information is often limited for many substances. Such limited data inherently presents challenges to environmental health practitioners performing hazard and risk assessments. Probabilistic hazard assessment using chemical toxicity distributions (CTDs) is an alternative approach for assessments of chemicals when toxicity information is lacking. The CTD concept allows for derivation of thresholds of toxicological concern (TTCs) to predict adverse effect thresholds for mammalian species. Unfortunately, comparative health hazard assessment of cleaning product ingredients in common use categories such as all-purpose cleaners (APC), dish care products (DCP) and laundry care products (LCP) has not been well studied. However, APC, DCP, and LCP are used routinely for household and industrial applications, resulting in residential and industrial occupational exposures. Therefore, we reviewed and then examined hazard information (median lethal dose (LD50), lowest-observed-adverse-effect level (LOAEL), and no-observed-adverse-effect level (NOAEL)) from different types of standard mammalian toxicity studies for oral toxicity in the rat model from the unique Cleaning Product Ingredient Safety Initiative mammalian toxicology database. Probabilistic distributions (CTDs) were subsequently constructed using LD50, NOAEL and LOAEL data from a specific toxicity study type for all available ingredients in these three use categories. Based on data availability, product type-specific and chemical category-specific CTDs were also generated and compared. For each CTD, threshold concentrations (TCs) and their 95% confidence intervals (95% CIs) at 1st, 5th, 10th, 50th, 90th, 95th and 99th percentiles were calculated using the log-normal model. To test whether the common default uncertainty factor (UF) approach (e.g., 3, 10) in mammalian health risk assessment provides sufficient protection, UFs were also derived for LOAEL-to-NOAEL and exposure duration (e.g., subchronic-to-chronic) extrapolations. Relationships between CTDs of acute LD50s and sublethal LOAELs/NOAELs were also examined for acute-to-chronic ratio calculations, which may be useful in extreme circumstances. Results from our critical review and meta-analysis appear particularly useful for hazard and risk practitioners when identifying TTCs for ingredients in product use categories, and other chemical classes. This approach can also support development of regulatory data dossiers through read across, chemical substitutions and screening-level health risk assessments when limited or no empirical toxicity information exists for industrial chemicals.

1. Introduction

Human are routinely exposed to diverse chemicals from food, cosmetics, household products, medicines and other environmental

matrices. Cleaning products (e.g., soaps, detergents, personal care products), like other consumer products, are sold in large quantities around the world, and consequently, the main chemical ingredients in these products are often high production volume (HPV) chemicals.

* Corresponding author at: Department of Environmental Science, Baylor University, Waco, TX, USA.

E-mail address: Bryan_Brooks@baylor.edu (B.W. Brooks).

<https://doi.org/10.1016/j.envint.2019.01.079>

Received 9 December 2018; Received in revised form 29 January 2019; Accepted 29 January 2019

Available online 08 February 2019

0160-4120/ © 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

Given the large number of products and possible associated consumer exposure scenarios, a priority setting process (e.g., screening-level risk assessment) is needed to identify consumer products and use scenarios for which more detailed hazard evaluation and/or exposure assessment may be needed to adequately characterize consumer risks and then to identify those substances that represent lower or higher concern to public health and the environment (ACI, 2016).

Current hazard assessment approaches tend to rely on the deterministic hazard quotient (HQ) approach and occasionally on probabilistic hazard assessment (PHA) approaches. The HQ approach often uses a worse-case scenario and point estimate, and is generally considered to provide (overly) conservative protection and hence is useful for lower tiers of hazard or risk assessment (Suter II, 1995). As part of the Cleaning Product Ingredient Safety Initiative (CPISI; <http://www.cleaninginstitute.org/CPISI/>), a critical review of product inventory and associated mammalian toxicology data was performed (DeLeo et al., 2018), and then a screening-level risk assessment based on the deterministic HQ approach was conducted to characterize the mammalian and human health safety for each ingredient of concern (588 in total). It is important to note that ingredients of concern examined by DeLeo et al. (2018) have a wide range of uses; for example, 588 ingredients are relevant to three main use categories including all-purpose cleaners (APC), dish care products (DCP) and laundry care products (LCP). Unfortunately, human health hazard assessment of all cleaning product ingredients on a product use category basis (APC, DCP, or LCP) has remained elusive. Moreover, extensive safety data exist for some of these ingredients, yet for others only limited data are currently available.

To support health risk assessment of these exposures when there are insufficient chemical-specific data, and to prioritize those most likely to present health risks, other approaches such as threshold of toxicological concern (TTC) or read across need to be applied to estimate potential human health impact and to make informed risk management decisions. The TTC approach can benefit from PHA tools establishing human TTCs for chemicals below which there apparently is no significant risk to human health (US FDA, 1995; JECFA, 2006). TTCs can be used to assess the likelihood that a particular level of exposure to a chemical having no toxic effects in the absence of chemical-specific data, relying on available toxicity data for a wide range of chemicals (Kroes et al., 2004). The TTC approach has been favorably reviewed for regulatory use in PHA of food contact materials and flavoring substances (see Table 1 for oral TTC values in current global regulatory uses).

A TTC-like approach was initially proposed by the United States Federal Food, Drug and Cosmetic Act (US FD&C Act, 1958) for chemical in food contact material (and their components) in the United States, in combination with the development of more sensitive and discriminating analytical methods to handle putative toxicological risks of low exposures. Frawley (1967) first presented an analysis to establish a generic threshold value (threshold of regulation; ToR) or ranges values with the aim to reduce extensive toxicity studies and safety evaluations, and to address, within the available capacity, those substances for which the potential or actual intake is substantial. Rulis (1986) conducted a similar probabilistic assessment of carcinogenicity data (TD50 values) for 130 compounds from a carcinogenic potency database (CPBD) developed by Gold et al. (1984) and LD50s of 159 compounds (rats, oral; subchronic or chronic toxicity) contained from the Registry of Toxic Effects of Chemical Substances (RTECS) database, with a linear extrapolation to a 10^{-6} risk; a general TTC value of 0.5 ppb (equivalent to $1.5 \mu\text{g}/\text{person}/\text{day}$ or $0.025 \mu\text{g}/\text{kg bw}/\text{day}$) was derived (Rulis, 1986; Flamm et al., 1987; Rulis, 1989). An estimated TTC value of $1.5 \mu\text{g}/\text{person}/\text{day}$ was then implemented by the United States Food and Drug Administration (US FDA, 1995) as the ToR for food contact materials. The TTC concept (TTC-based limits) has also been adopted as by the European Medicines Agency (EMA, 2006) and US FDA (2008) for genotoxic impurities in pharmaceutical products and by EMA (2008)

for genotoxic constituents herbal substances/preparations, despite the limits varied based on the potency of carcinogens (i.e., direct or indirect mechanisms of DNA damage).

For non-carcinogenic endpoints, Munro et al. (1996) evaluated the implication of TTC using structural information based on a decision tree approach developed by Cramer et al. (1978). Human TTCs of 1800, 540, 90 and $90 \mu\text{g}/\text{person}/\text{day}$ were proposed for Cramer class I, II, and III, respectively using the 5th percentile of the distributions and a safety factor of 100. These distributions were developed using non-carcinogenic data of no-observed-effect levels (NOELs) for chronic, subchronic and reproductive toxicity for 613 organic substances. The European Food Safety Authority (EFSA) adopted those values for the regulation of flavoring substances used in food (EC, 2000). The Joint FAO/WHO Expert Committee on Food Additives (JECFA) also adopted the TTC principle in its evaluations of flavoring substances (JECFA, 1995) in the combination of the general TTC values of 90, 540, $1800 \mu\text{g}/\text{person}/\text{day}$ for non-carcinogenic responses with the TTC of $1.5 \mu\text{g}/\text{person}/\text{day}$ based on carcinogenic information.

This TTC concept has also been commonly used for safety evaluation of chemicals, such as substances present in foods (Kroes et al., 2004), flavoring agents (Renwick, 2004), natural flavor complexes (essential oils) (Smith et al., 2004; Smith et al., 2005), consumer products (personal and household care products) (Blackburn et al., 2005), cosmetics (Kroes et al., 2007), metabolites of plant protection products (pesticides) (Brown et al., 2009), fragrance ingredients (Api et al., 2015). It has further been considered for exposure-based waiving of toxicity tests under REACH (Bernauer et al., 2008; Rowbotham and Gibson, 2011). Our recent study identified unique TTCs for cleaning product ingredients from different common chemical classes using rodent (oral) hazard data (Wang et al., 2018); however, health hazard assessment of cleaning product ingredients in different use categories such as APC, DCP and LCP has not been well studied. Therefore, it may be useful to identify TTCs for cleaning product ingredients in a specific use category (APC, DCP or LCP) by employing a probabilistic approach such as chemical toxicity distributions (CTDs) commonly used in PHA.

Another consideration is that since higher volume ingredients are potentially more data rich it may be possible to build read across quantitative structure-activity relationship (QSAR) arguments supporting one or more of the toxicological endpoints for a lower volume chemical by using data from the higher volume one (Api et al., 2015). In principle, read across utilizes common endpoint information, including physiochemical properties and toxicity for chemicals to make a prediction on the same endpoint for another chemical. This process can help to avoid generating specific test data on every substance for every endpoint. To address the potential uncertainties associated with read across, uncertainty factors (UFs; often synonymous with assessment or adjustment or safety factors) are commonly applied for a more conservative concern (e.g., default UF approach). In cases where only a LOAEL is available, and a benchmark dose (BMD) approach is not applicable or available, then selecting an appropriate UF is, therefore, important to the practice of performing screening-level health risk assessments, including for cleaning product ingredients. However, whether default UFs (e.g., 10 or 100) are sufficient for various chemical uses remains understudied, particularly for cleaning products ingredients in different product use categories.

By leveraging a unique mammalian toxicology database (DeLeo et al., 2018), the objectives of present study were primarily designed: 1) to review and identify threshold concentration (TCs) for all available ingredients in each common cleaning product use category (APC, DCP, LCP) and specific product types/chemical categories using the CTD approach; 2) to compare relative sensitivities of ingredients in different product types or common categories within a given use category or among the three use categories; and 3) to derive UFs including acute-to-chronic ratios (ACRs) for each product use category by using a two probability distributions approach (i.e., CTD comparisons and individual UF probability distribution approaches), accompanied by

Table 1
Summary of threshold of toxicological concern (TTC) values (oral) and approaches as pragmatic risk assessment or prioritization tools, identified in areas of chemical hazard and risk assessment within a regulatory context.

| | US FD&C Act | Frawley, 1967 ^c | Rulis, 1986 | US FDA | EMA & US FDA | EMA | Cheeseman et al., 1999 |
|--|--|---|---|--|--|---|---|
| Chemical | Food additives ^b | Heavy metals, pesticides and all other compounds | Food contact articles | Food contact articles | Genotoxic impurities in pharmaceuticals | Genotoxic constituents in herbal medicinal products | Food contact articles ^g |
| Terminology Approach | TTC-like | ToR or range values | ToR | ToR | ToR | ToR | ToR |
| Endpoint | | Chronic (2-year) rodent studies on 220 substances | ToR and de minimis Carcinogens: TD50 ^d | ToR and de minimis Carcinogens: TD50 | | | Tiered SAR Carcinogens: TD50 (short-term); non-carcinogens: NOEL |
| Thresholds (mg/kg bw/day) ^a | | | 1.5 | 1.5 | 0.15 | TTC-like limits ^f | 1.5; 15; 45 ^b |
| Assessment factors | | | (Rulis, 1986) | (US FDA, 1995) | (EMA, 2006) | (EMA, 2008) | (Cheeseman et al., 1999) |
| TTC values (µg/person/day) | (US FD&C Act, 1958) | (Frawley, 1967) | (Gold et al., 1984) ^f | | (US FDA, 2008) | | |
| Sources | | | | | | | |
| | Munro et al., 1996 | EFSA | JECFA | ECETOC | ILSI Europe | | |
| Chemical | Industrial organic substances | Flavoring substances | Flavoring substances | Industrial chemicals | Chemicals present in foods in low concentrations | | |
| Terminology | TTC | TTC | TTC | GLEV | TTC | | ToR |
| Approach | Cramer decision tree | Cramer decision tree | Cramer decision tree | Hazard category | Cramer decision tree | | ToR and de minimis |
| Endpoint | Non-carcinogens; NOEL ^l | Non-carcinogens; NOEL | Non-carcinogens; NOEL | EU criteria for classification of substances for repeated dose toxicity (R48) ^j | Potentially sensitive endpoints n: NOELs | | Carcinogens: TD50 |
| Thresholds (mg/kg bw/day) ^a | III: 0.15; II: 0.91; I: 3.0 | III: 0.15; II: 0.91; I: 3.0 | III: 0.15; II: 0.91; I: 3.0 | High: 0.5; medium: 5; low: 50 | III: 0.15; II: 0.91; I: 3.0 | | |
| Assessment factors | 100 | 100 | 100 | High: 150; medium: 1500; low: 15,000 ^k | 100 | | |
| TTC values (µg/person/day) | III: 90; II: 540; I: 1800 ^l | III: 90; II: 540; I: 1800 ^l | III: 90; II: 540; I: 1800 ^l | (ECETOC, 2004) | III: 90; II: 540; I: 1800 ^l ; organophosphates: 18 ^m | | 1.5 |
| Sources | (Munro et al., 1996) | (EC, 2000) | (JECFA, 1995) | (Cramer et al., 2004) | (Barlow, 2005) | | |
| | (Cramer et al., 1978) | (Cramer et al., 1978) | (Cramer et al., 1978) | (Kroes et al., 2004) | (Cramer et al., 1978) | | |

Notes: US FD&C Act: United States Federal Food, Drug, and Cosmetic Act; US FDA: United States Food and Drug Administration; EMA: European Medicines Agency; EFSA: European Food and Safety Authority; JECFA: Joint FAO/WHO Expert Committee on Food Additives; ECETOC: European Center for Ecotoxicology and Toxicology of Chemicals; ILSI: International Life Sciences Institute; ToR: Threshold of Regulation; SAR: Structure-Activity Relationships; GLEV: General Lowest Effect Value.

^a 5th percentile of the NOELs distributions.
^b Food contact materials and their components.
^c First to present an analysis to establish a generic threshold value (threshold of regulation) or ranges values.
^d 159 compounds with subchronic or chronic toxicity data, LD50 values from 18,000 oral rodent studies contained in the Registry of Toxic Effects of Chemical Substances (RTECS) database, and TD50 values for 130 compounds found in the carcinogen potency database of Gold et al. (1984).
^e Gold et al. (1984): carcinogenic potency database: 3000 chronic (long term) experiments of 770 test compounds.
^f Permitted levels of a genotoxic compound in herbal preparations can be calculated based on a TTC value of 1.5 µg/person/d; Higher limits could be justified under certain conditions such as short-term exposure, or if the applicant submits additional data and a toxicologically plausible argumentation.
^g Food packaging migrants, contact materials, flavoring substances and unintended chemicals in food.
^h 1.5 (general; for compounds with positive genotoxicity testing and/or no nitroso, bezidione, hydrazine, azo, endocrine disruptor SAS, or no testing data available); 15 (for chemicals without structural alerts for carcinogenicity (no SAS) or with negative mutagenicity tests (Ames test)); 45 (for chemicals without structural alerts for carcinogenicity (no SAS) or with negative mutagenicity tests (Ames test), and appropriate acute toxicity test with LD50 > 1000 mg/kg bw).
ⁱ Chronic and subchronic reproductive toxicity: 2941 NOELs; 613 organic substances; carcinogenic and mutagenic endpoints were not considered.
^j Average individual body weight: 60 kg/person.
^k Average individual body weight: 70 kg/person.
^l Inclusion in hazard categories depends on the substance's specific classification (or no classification) according to Directive 67/548/EC (EC, 1967); for consumers, the Generic Lowest Effect Values (GLEVs) are based on EU criteria for the classification of the substance for repeated dose toxicity (R48 "danger of serious damage to health by prolonged exposure") after adjustment with an assessment factor.
^m Subchronic, chronic, reproductive and developmental toxicity studies for rats and rabbits on 613 chemicals with a wide range of structures and uses; potentially sensitive endpoints: immunotoxicity, developmental toxicity, neurotoxicity and developmental neurotoxicity, endocrine active compounds, and allergenicity.

Monte Carlo simulation approach. This effort thus extended our recent report of novel TTCs and UFs for cleaning products (Wang et al., 2018) to examine specific categories of common product uses.

2. Methods

2.1. Data mining

Health hazard information for the rat model (oral; mg/kg bw/day) including median lethal dose (LD50s), lowest-observed-adverse-effect level (LOAEL), and no-observed-adverse-effect level (NOAEL) following a defined standard mammalian toxicology study type (Test Guideline) were collected and reviewed from the CPISI database (<http://www.cleaninginstitute.org/CPISI/>). These test types include: Repeated Dose 28-Day Oral Toxicity Study in Rodents (e.g., OECD Test Guideline (TG) Number 407), Repeated Dose 90-Day Oral Toxicity Study in Rodent (OECD TG 408), Prenatal Development Toxicity Study (OECD TG 414), Reproduction and Fertility Effects (including One-, Two- and Multi-generation Reproduction Study; OECD TG 415 and 416), Chronic Toxicity Studies (OECD TG 452), Carcinogenicity Studies (OECD TG 451), Combined Chronic Toxicity/Carcinogenicity Studies (OECD TG 453), and two developmental/reproductive screening tests such as Reproduction/Developmental Toxicity Screening Test (OECD TG 421) and Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test (OECD TG 422). Because some ingredients of cleaning products in a product use category contained little to no sublethal toxicology information (e.g., chronic LOAEL/NOAELs), acute LD50s (Acute Oral Toxicity; OECD TG 401, 420, 423 and 425) were still examined in the current study. Information for each ingredient corresponding to its use category (APC, DCP, or LCP), product type and chemical category were also collected.

2.2. Chemical toxicity distributions constructions and threshold concentrations calculations

Probabilistic hazard assessment (PHA) using the CTDs approach was conducted following Wang et al. (2018). Briefly, outliers detected by the Grubb's test (Grubbs, 1969) or Tietjen-Moore test (Tietjen and Moore, 1972) were excluded for each dataset. A geometric mean was used when there were multiple data from a common study type for an ingredient. To increase the robustness of our analyses, datasets for an CTD must have contained a minimum of 5 data points (ingredients). Hazard data were ranked in ascending order, and percentiles were assigned from the Weibull formula (Eq. (1)).

$$\text{percentile} = i/(n + 1) \times 100\% \quad (1)$$

where i is the rank of the datum in ascending order and n is the total number of data points.

Normality of residuals of each dataset (log-transformed) was checked by the Shapiro-Francia test (Shapiro and Francia, 1972), and goodness of fit at the lower tail of each distribution was assessed by Anderson-Darling test (Anderson and Darling, 1954). Each CTD was then fitted by the log-normal model (SigmaPlot, version 13.0, San Jose, CA, USA), and the TC values and 95% confidence intervals (95% CIs) were determined by log-normal regression function and Monte Carlo simulation (resampled 5000 times; SAS, version 9.4, Cary, NC, USA). When there were sufficient data points (≥ 5) for a product type or chemical category in a dataset (use category), product type- or chemical category-specific CTDs were also constructed and compared.

To be consistent, the relative chemical sensitivities between two CTDs were compared based on the calculated TC5s and their 95% CIs; if the 95% CIs of the both TC5s did not overlap, then the two CTDs were significantly different (Sokal and Rohlf, 1995). One-way analysis of variance (ANOVA) and Tukey's post hoc tests were performed to compare the relative sensitivities among different use categories, or among different product types or chemical categories (≥ 3 CTDs) within a use

category on the basis of TC5 values and their 95% CIs (significance level $\alpha = 0.05$; Graph Pad Prism™, version 5.00, San Diego, CA, USA).

2.3. Uncertainty factors derivations

Uncertainty factors were identified using both CTD comparisons and individual UF probabilities proposed by Wang et al. (2018). First, the UFs for a given use category were derived from corresponding TC ratios of the pairwise datasets using all and similar ingredients (hereafter, "all" indicates all hazard data were used for pairwise CTDs constructions, while "similar" indicates both datasets contain common ingredients). Relative TC ratios and their 95% CIs were computed by Monte Carlo simulation approach as described above. One way-analysis of covariance (ANCOVA) was also conducted to compare the slope and/or intercept parameters of the two log-normal fitted CTDs (SPSS, version 23, Chicago, IL, USA). To further quantitatively compare the difference between the both CTDs, TC5 ratio and its 95% CI was also applied for each comparison. If 95% CIs did not overlap with unity, then the ratio was significantly different from 1 (Sokal and Rohlf, 1995). Second, UFs were calculated separately for individual ingredient within a use category from the respective pairwise datasets consisting of similar ingredients. All calculated individual UFs values were then ranked and assigned percentiles following the Weibull formula (Eq. (1)) for a probability distribution. Afterwards, overall UFs (95% CIs) covering 1%, 5%, 10%, 50%, 90%, 95% and 99% of all ingredients from each log-normal fitted distribution were computed by an inverse pre-diction method.

Concurrent meta-analyses were conducted, respectively, using LOAELs and corresponding NOAELs for pairwise CTD comparisons and UFs derivation especially for cases with little or no LOAEL data. Relationships between CTDs using acute LD50s and sublethal LOAELs/NOAELs were also examined, and corresponding extrapolation factors (ACRs) were also derived. Whether a default UF of 10 or 100 would be protective for various distributions of cleaning product ingredients within a given use category were also evaluated. This study was thus not intended to examine UFs associated with extrapolation from animals to humans or variation of sensitivities within or among human populations.

3. Results

3.1. Chemical toxicity distributions and thresholds concentrations for three common use categories

Among all datasets used for CTDs constructions, a few outliers were identified and removed from our analysis (see Table S1 in Supplementary information for datasets). Using sufficient toxicity data ($n \geq 5$) for rats (oral), CTDs using LD50s, LOAELs (Fig. 1), and NOAELs (Fig. S1) were generated for each toxicity type. Calculated TC values and their 95% CIs at different percentiles are listed in Table 2 (for LD50s and LOAELs) and Table S2 (for NOAELs). These derived TCs and their 95% CIs (e.g., TC1 or TC5) can serve as TTCs for particular datasets and thus can be useful during future hazard and risk assessment for ingredients in different use category cleaning products.

Among three use categories (Table 2), the rat model tended to be more acutely sensitive (Acute Oral Toxicity) to APC (lower TC5: 442 (406, 479) mg/kg bw/day) compared to DCP and LCP (602 (529, 682) and 530 (491, 571) mg/kg bw/day, respectively), while they were less sensitive against APC (higher TC5: 6.7 (4.4, 9.0) mg/kg bw/day) than DCP and LCP (2.6 (1.1, 4.9) and 1.7 (0.7, 3.5) mg/kg bw/day, respectively) when considering Repeated Dose 90-Day Oral Toxicity. Nonetheless, there were no significant differences of TC5s among three use categories for other types of subacute and subchronic studies (ANOVA; $p > 0.05$). For example, TC5s of 17 (9.0, 29), 9.3 (3.0, 20), 14 (7.6, 25) mg/kg bw/day for APC, DCP, LCP, and 1.6 (0.3, 5.0), 12 (0.3, 60) mg/kg bw/day for APC and LCP were estimated when

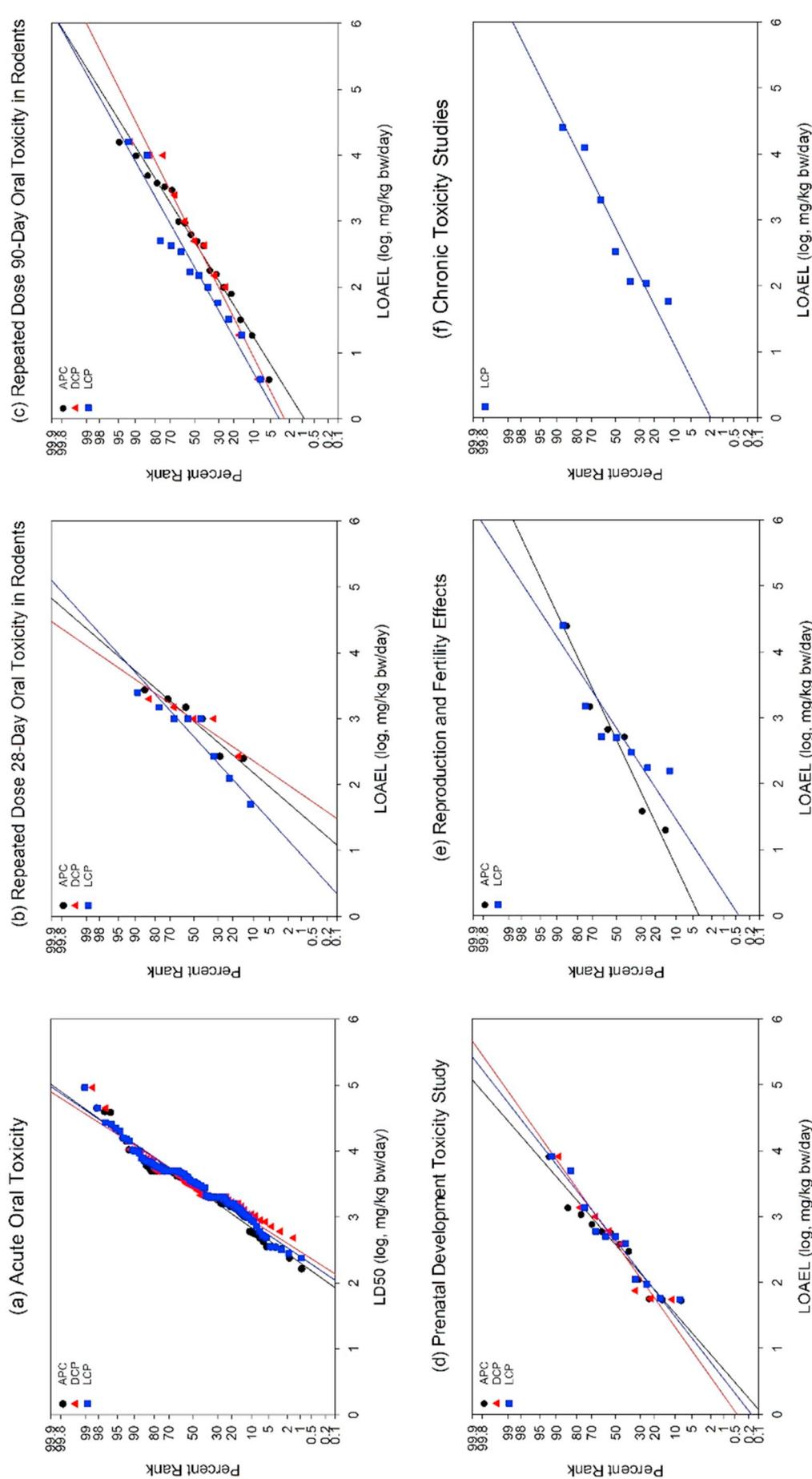
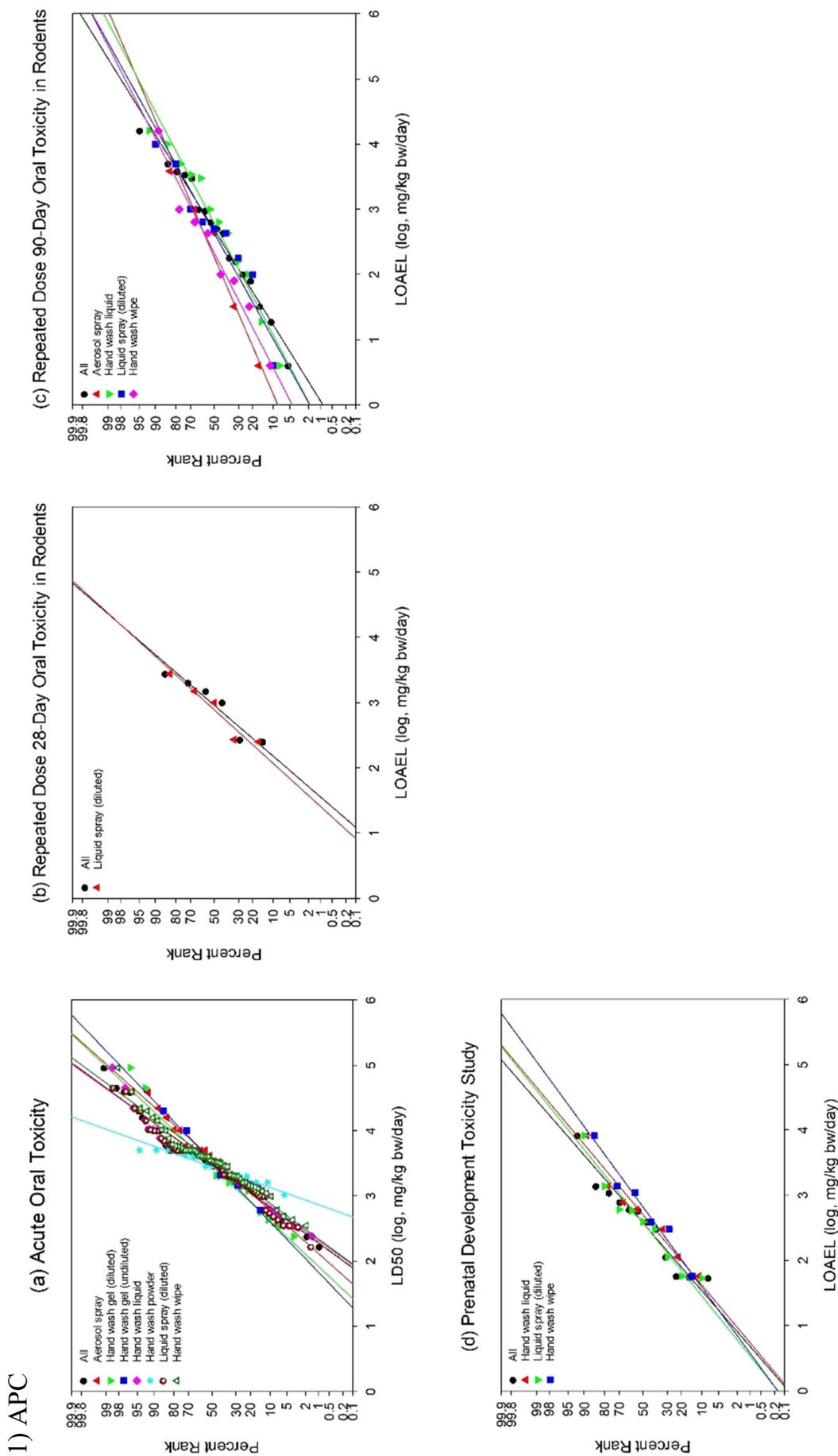


Fig. 1. Chemical toxicity distributions (CTDs) for all purpose cleaners (APC), dish care products (DCP) and laundry care products (LCP) using all available median lethal doses (LD50) and lowest-observed-adverse-effect level (LOAEL) values from different types of mammalian toxicity studies.

Table 2
 Threshold concentrations (TCs; mg/kg bw/day) and their 95% confidence intervals (95% CIs) at 1st, 5th, 10th, 50th, 90th, 95th, and 99th percentiles for each chemical toxicity distribution (CTD) using all available median lethal doses (LD50) and lowest-observed-adverse-effect level (LOAEL) values.

| CTDs | Use category | n | $y = bx + a$ | | TC1 (95% CI) | TC5 (95% CI) | TC10 (95% CI) | TC50 (95% CI) | TC90 (95% CI) | TC95 (95% CI) | TC99 (95% CI) | |
|---|--------------|-----|--------------|-------|--------------|-----------------|------------------|----------------|-------------------|--------------------------|-----------------------------|---------------------------------|
| | | | b | a | | | | | | | | R ² |
| Acute Oral Toxicity (LD50) | APC | 109 | 56.8 | -147 | 0.879 | 202 (180, 225) | 442 (406, 479) a | 671 (626, 719) | 2933 (2813, 3052) | 12,813 (11,969, 13,771) | 19,463 (17,934, 21,187) | 42,633 (38,245, 47,706) |
| | DCP | 68 | 64.7 | -177 | 0.869 | 298 (250, 351) | 602 (529, 682) b | 875 (789, 973) | 3275 (3097, 3501) | 12,255 (11,203, 13,912) | 17,815 (15,997, 20,769) | 35,939 (31,157, 43,992) |
| | LCP | 106 | 59.9 | -160 | 0.891 | 251 (227, 277) | 530 (491, 571) b | 789 (740, 840) | 3211 (3097, 3334) | 13,071 (12,310, 13,907) | 19,459 (18,096, 20,952) | 41,049 (37,264, 45,353) |
| Repeated Dose 28-Day Oral Toxicity Study in Rodents (LOAEL) | APC | 6 | 57.7 | -121 | 0.922 | 35 (6.3, 74) | 92 (25, 155) | 152 (53, 231) | 907 (638, 1191) | 5414 (3111, 12,333) | 8983 (4594, 25,675) | 23,224 (9649, 101,410) |
| | DCP | 5 | 72.0 | -165 | 0.828 | 72 (6, 174) | 154 (28, 286) | 230 (61, 386) | 959 (609, 1318) | 3991 (2166, 15,178) | 5980 (2850, 31,798) | 12,767 (4706, 132,156) |
| Repeated Dose 90-Day Oral Toxicity Study in Rodents (LOAEL) | LCP | 8 | 43.7 | -69.0 | 0.889 | 8.6 (2.1, 22) | 29 (11, 58) | 55 (24, 97) | 531 (378, 751) | 5131 (2993, 12,961) | 9761 (5069, 30,299) | 32,607 (13,428, 149,342) |
| | APC | 18 | 27.9 | -24.2 | 0.972 | 1.1 (0.7, 1.7) | 6.7 (4.4, 9.0) a | 17 (12, 22) | 485 (407, 561) | 13,665 (10,243, 18,227) | 35,202 (25,079, 50,115) | 207,719 (132,796, 339,297) |
| | DCP | 11 | 23.4 | -13.9 | 0.964 | 0.3 (0.1, 0.7) | 2.6 (1.1, 4.9) b | 8.5 (4.0, 14) | 532 (369, 715) | 33,313 (18,345, 62,925) | 107,637 (53,173, 236,427) | 971,434 (380,021, 2,844,389) |
| Prenatal Development Toxicity Study (LOAEL) | LCP | 12 | 25.6 | -8.88 | 0.913 | 0.2 (0.1, 0.6) | 1.7 (0.7, 3.5) b | 4.9 (2.3, 8.9) | 200 (143, 289) | 8165 (4544, 17,620) | 23,358 (11,452, 59,174) | 167,780 (64,266, 586,754) |
| | APC | 12 | 40.2 | -53.3 | 0.929 | 4.9 (2.0, 9.6) | 17 (9.0, 29) | 34 (20, 52) | 372 (298, 481) | 4056 (2739, 7094) | 7984 (4949, 15,858) | 28,436 (14,903, 72,220) |
| | DCP | 8 | 34.1 | -38.6 | 0.939 | 2.0 (0.4, 5.4) | 9.3 (3.0, 20) | 21 (8.7, 40) | 397 (278, 589) | 7398 (4039, 19,230) | 16,955 (8198, 54,455) | 80,334 (29,850, 391,205) |
| Reproduction and Fertility Effects (LOAEL) | LCP | 11 | 36.5 | -46.4 | 0.945 | 3.5 (1.4, 7.2) | 14 (7.6, 25) | 31 (18, 48) | 436 (347, 573) | 6201 (4217, 11,452) | 13,163 (8218, 27,644) | 54,023 (28,674, 144,474) |
| | APC | 7 | 20.8 | 0.199 | 0.959 | 0.2 (0.01, 0.7) | 1.6 (0.3, 5.0) | 5.7 (1.4, 14) | 465 (274, 814) | 38,211 (15,630, 137,760) | 133,361 (44,710, 665,123) | 1,390,894 (315,296, 13,308,318) |
| Chronic Toxicity Studies (LOAEL) | LCP | 7 | 31.3 | -38.9 | 0.78 | 2.2 (0.01, 18) | 12 (0.3, 60) | 29 (1.6, 115) | 696 (359, 2269) | 16,595 (6293, 543,005) | 40,776 (12,511, 2,950,883) | 220,192 (43,224, 71,008,283) |
| | APC | 7 | 24.7 | -21.3 | 0.935 | 0.4 (0.03, 2.3) | 3.9 (0.6, 14) | 13 (2.6, 36) | 767 (431, 1467) | 46,717 (18,320, 255,649) | 149,744 (48,466, 1,221,620) | 1,331,354 (288,686, 22,693,880) |



2) DCP

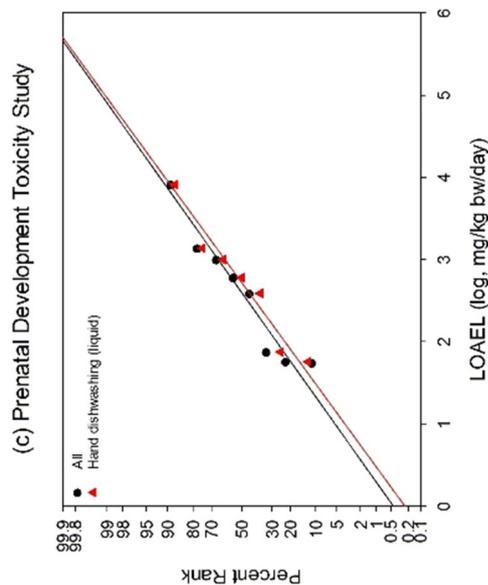
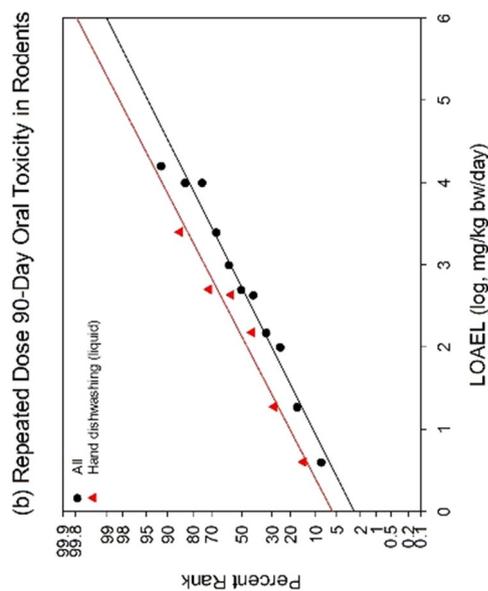
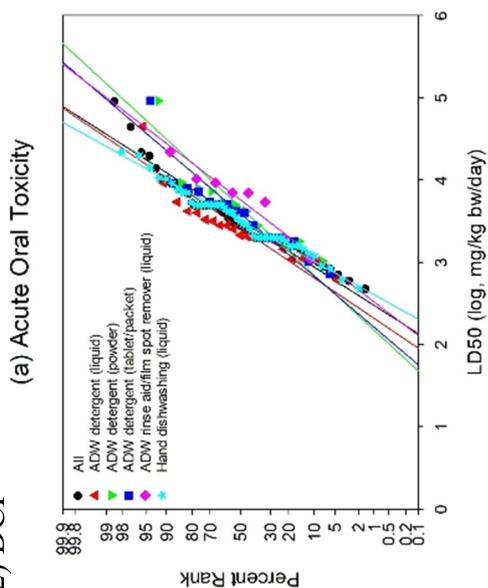


Fig. 2. (continued)

considering LOAEL values generated from Prenatal Development Toxicity Study and Reproduction and Fertility Effects studies, respectively. There was only one chronic CTD generated for LCP using LOAEL data from Chronic Toxicity Studies (Fig. 1f), and a chronic TC5 of 3.9 (0.6, 14) was derived accordingly. Comparatively, there were more CTDs generated using NOAEL data generated from additional sublethal toxicology study types such as two reproductive/developmental screening tests, and three chronic toxicity tests (except for Chronic Toxicity Studies for LCP) (Fig. S1). Corresponding TC values were also estimated and listed in Table S2. For example, APC were found to be less toxic to rats (oral) than LCP when considering NOAELs of Chronic Toxicity Studies (DCP shared similar sensitivity to APC and LCP, respectively), while APC were the most toxic category to rats when considering NOAELs derived from Carcinogenicity Studies (that is, we did not examine carcinogenicity endpoints, but analyzed NOAEL for other responses reported from these studies), followed by LCP and DCP.

3.2. Product type- and chemical category-specific chemical toxicity distributions

Product type-specific CTDs and chemical category-specific CTDs (Figs. 2 and 3 using LD50s and LOAELs, and Figs. S2 and S3 using NOAELs) were developed and can be viewed as sub-distributions within most of general datasets for a given type of toxicity test. Significant differences between or among product type-specific and chemical category-specific CTDs were also detected by checking whether their 95% CIs of TC5s were overlapping ($n = 2$) or by conducting one-way ANOVA and Tukey's post hoc test ($n \geq 3$; Table S3 using LD50s and LOAELs, and Table S4 using NOAELs).

Using APC as an example, cleaning product ingredients used in products of aerosol spray, hand wash gel (diluted) and hand wash gel (undiluted) were more acutely toxic to rodents compared to the general CTD consisting of all acute data (one-way ANOVA: $p < 0.05$, Table S3; left shift at the lower tails of the CTDs, Fig. 2a for APC), while the ingredients used in hand wash powder were the least acutely toxic. Significant differences in sensitivities among product type-specific CTDs were also found using LOAELs generated from Repeated Dose 90-Day Oral Toxicity in Rodents, and aerosol spray, and hand wash wipe were the two most two toxic types of cleaning products to rats compared to other product type-specific and global CTDs (Fig. 2c for APC, and Table S3).

Due to limited toxicity data, sub-distributions were only generated for 4 specific chemical categories (aliphatic acids and salts, chelants, ethers, inorganic acids and salts) to APC, 1 (inorganic acids and salts) within the DCP product use category, and 3 (aliphatic acids and salts, chelants inorganic acids and salts) for cleaning products used in LCP considering Acute Oral Toxicity (LD50s; Fig. 3). Similarly, rodents (oral) tended to be less acutely sensitive to aliphatic acids and salts in APC, inorganic acids and salts in DCP, and chelants in LCP, respectively compared to the CTD generated using all LD50s, and other chemical categories (Table S3). Ethers- and inorganic acids and salts-specific CTDs were also generated using subchronic NOAELs (Fig. S3), and rats (oral) were less subchronically sensitive (Prenatal Development Toxicity Study) to ingredients of inorganic acids and salts in DCP than the global CTD, while more subchronically sensitive to ingredients of inorganic acids and salts in LCP (Table S4).

3.3. Uncertainty factors including acute-to-chronic ratios

Most CTD comparisons using either all (Figs. S4 and S5) or similar (Figs. S6 and S7) ingredients were visually diverged, indicating different sensitivities. This was supported by significantly different slope and/or intercept parameters for the both CTDs fitted by log-normal model (Table S5). UFs were also estimated using the CTD comparison approach with all (Tables 3, S6 and S7) and similar ingredients of both datasets (e.g., Acute Oral Toxicity (LD50) vs Reproduction and Fertility

3) LCP

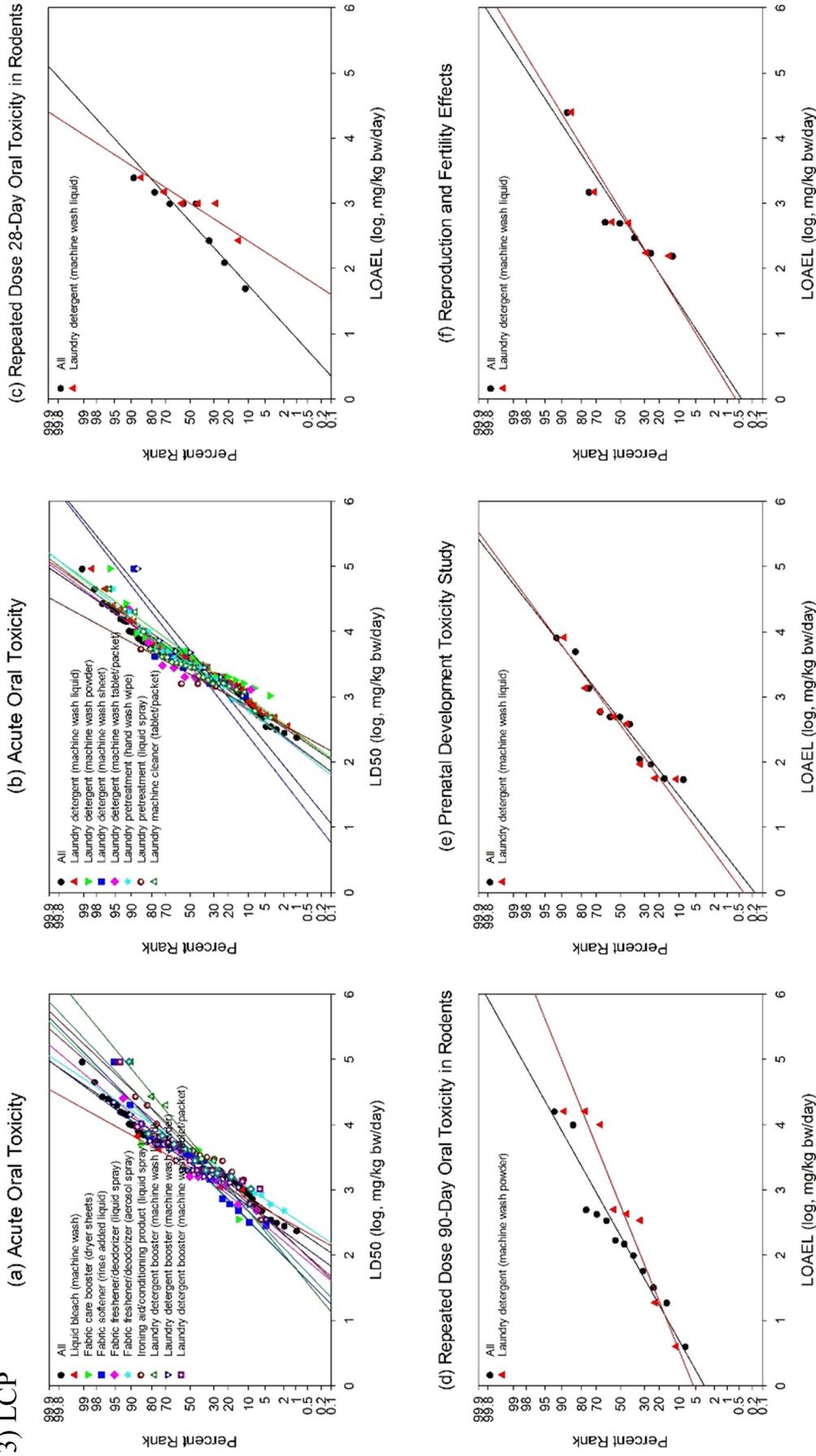


Fig. 2. (continued)

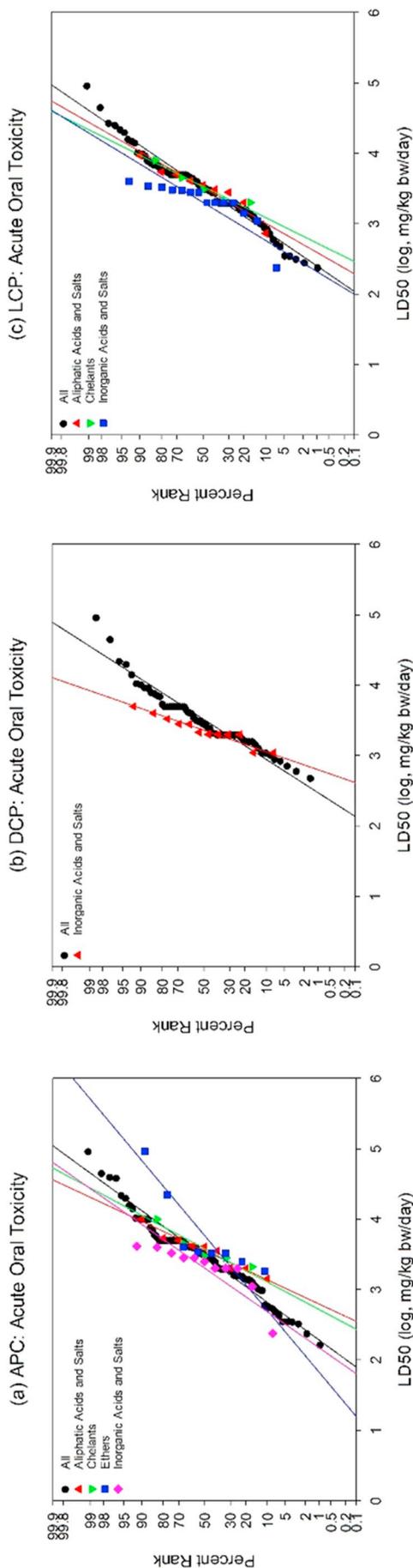


Fig. 3. Chemical category-specific acute (LD50) chemical toxicity distributions (CTDs) generated using all available median lethal doses (LD50) and lowest-observed-adverse-effect level (LOAEL) values.

Effects (LOAEL); Tables 3, S8 and S9). Using available pairwise data for each ingredient from the two dataset (Acute Oral Toxicity (LD50) vs Reproduction and Fertility Effects (LOAEL)), individual UFs were also calculated and ranked for construction of probability distributions (Figs. 4 and S8). UFs covering 1%, 5%, 10%, 50% 90%, 95%, and 99% of ingredients were computed from each distribution using log-normal function (Tables 3, S10 and S11). Corresponding TC values (95% CIs) for the CTD comparisons with similar ingredients were also computed (Tables S12 and S13).

For LOAEL-to-NOAEL ratios (UF_{L-N}), we identified 4 (3 with similar ingredients), 3 and 5 UFs, respectively, for the APC, DCP, and LCP use categories (Tables 3, S6 and S8). Taking TC5 ratios for example (Table 3), only 1 UF_{L-N} using CTD comparison with all ingredients was larger than 1 (i.e., 14 (2.7, 32)) for DCP using data of Repeated Dose 28-Day Oral Toxicity Study in Rodents, while the rest were all equal to 1 for APC, DCP and LCP, respectively. Comparatively, most UF_{L-N} 's (3 for APC, 2 for DCP, and 3 for LCP, respectively) derived from CTD comparisons with similar ingredients and all UF_{L-N} 's estimated from individual UF probability distribution (90th centile) were larger than 1. For example, for the Prenatal Development Toxicity Study type, UF_{L-N} 's of 1.1 (0.6, 1.9)–9.8 (7.7, 14), 0.8 (0.3, 1.8)–12 (8.7, 22), 1.1 (0.5, 1.9)–8.6 (7.0, 13) were estimated for APC, DCP, and LCP use categories, respectively.

For exposure duration extrapolations, 3 larger than 1 subacute-to-subchronic UFs were estimated using CTD comparisons consisting of all available LOAEL data from Repeated Dose 28-Day Oral Toxicity Study and Repeated Dose 90-Day Oral Toxicity Study in Rodents only (i.e., 14 (3.4, 27) for APC, 58 (10, 179) for DCP, and 17 (4.5, 54) for LCP; Table 3). Three subacute-to-subchronic UFs using corresponding NOAELs from the two toxicity types were also estimated as 1.1 (0.2, 15)–3.8 (3.4, 5.2) for APC, 0.7 (0.1, 8.0)–3.3 (2.5, 7.4) for DCP, 1.0 (0.7, 1.4)–5.3 (4.3, 8.9) for LCP, respectively (Tables S7, S9 and S11). For subchronic-to-chronic extrapolations, only 2 UFs were estimated using LOAEL between Prenatal Development Toxicity Study or Reproduction and Fertility Effects and Chronic Toxicity Studies for LCP, whereas the both were equal to 1 (1.0 (3.7 (0.9, 27) and 3.0 (0.06, 45), respectively; Table 3). Corresponding subchronic-to-chronic extrapolations using NOAELs from Prenatal Development Toxicity Study or Reproduction and Fertility Effects to Chronic Toxicity Studies and additional two chronic toxicity study types (Carcinogenicity Studies and Combined Chronic Toxicity/Carcinogenicity Studies) were also computed and listed in Tables S7, S9 and S11. Using LCP, for example, UF_{S-C} 's of 4.7 (2.2, 11)–31 (13, 138) and 0.6 (0.2, 1.6)–53 (28, 178) were estimated by extrapolating from Prenatal Development Toxicity Study or Reproduction and Fertility Effects to Chronic Toxicity Studies, respectively.

Relationships among CTDs generated using LD50s from Acute Oral Toxicity and LOAELs/NOAELs from subacute, subchronic and chronic mammalian toxicology studies were also compared. Considering LD50-to-LOAEL ratios, for example, ACRs of 3.1 (0.1, 54)–169 (79, 637) for LCP were estimated using acute LD50s and chronic LOAELs (Table 3). And UFs of 25 (25, 47)–46 (35, 77) for APC, 37 (13, 129)–50 (28, 308) for DCP, 37 (21, 71)–67 (50, 95) for LCP, and 273 (87, 1564)–625 (67, 11,902) for APC, 11 (1.6, 33)–45 (9.0, 1806) for LCP were identified by comparing Acute Oral Toxicity study and Prenatal Development Toxicity Study and Reproduction and Fertility Effects study, respectively. Corresponding LD50-to-NOAEL ratios and additional ACRs considering different study types were also computed (Tables S7, S9 and S11). We then employed a PHA approach using these CTDs to determine probability of encountering an UF for cleaning products above or below common UF of 10 or 100. Herein, the percentile of cleaning product ingredients in a use category accounted for a given distribution was computed (Tables 3 and S11). For example, if acute hazard information (e.g., LC50s or TC5s) are used as surrogates and divided by a factor of 10, these extrapolated reproductive LOAELs can only provide 39% for APC and 39% for LCP of protection, respectively (Table 3), while the

Table 3 Uncertainty factors (UFs) for median lethal dose (LD50)-to-lowest-observed-adverse-effect level (LOAEL), LOAEL-to-no-observed-adverse-effect level (NOAEL), exposure duration (e.g., subacute LOAEL-to-subchronic LOAEL) extrapolations using both chemical toxicity distribution (CTD) comparisons (5% threshold concentration (TCS) ratios) with all and similar ingredients, and individual UF probability distributions (90th centile). The percentiles of ingredients being protected under default UF of 10 and 100 were also estimated from each log-normal fitted distribution.

| UFs | CTD comparisons: TCS ratios (95% CIs) | | | | Individual UF probability distribution: 90th centile (95% CIs) | The likelihood (%) of ingredients being protected under a default UF | |
|-------------------------|---|---|---------------------|------------------|--|--|-------|
| | All ingredients | | Similar ingredients | | | 10 | 100 |
| | | | | | | | |
| 1) APC LD50-to-LOAEL | Acute Oral Toxicity (LD50) | Repeated Dose 28-Day Oral Toxicity Study in Rodents (LOAEL) | 4.8 (2.8, 18) | 9.8 (3.2, 37) | 14 (10, 21) | 83% | > 99% |
| | | Repeated Dose 90-Day Oral Toxicity Study in Rodents (LOAEL) | 66 (49, 99) | 73 (21, 222) | 367 (131, 8150) | 50% | 79% |
| | | Prenatal Development Toxicity Study (LOAEL) | 25 (15, 47) | 38 (21, 75) | 46 (35, 77) | 61% | 96% |
| | | Reproduction and Fertility Effects (LOAEL) | 273 (87, 1564) | 625 (67, 11,902) | 481 (295, 1018) | 39% | 74% |
| LOAEL-to-NOAEL | Repeated Dose 28-Day Oral Toxicity Study in Rodents (LOAEL) | Repeated Dose 28-Day Oral Toxicity Study in Rodents (NOAEL) | 3.5 (0.9, 7.1) | 4.4 (1.3, 9.8) | 7.9 (6.5, 10) | 95% | > 99% |
| | Repeated Dose 90-Day Oral Toxicity Study in Rodents (LOAEL) | Repeated Dose 90-Day Oral Toxicity Study in Rodents (NOAEL) | 0.9 (0.6, 1.6) | 4.7 (2.9, 8.1) | 46 (23, 245) | 65% | 95% |
| | Prenatal Development Toxicity Study (LOAEL) | Prenatal Development Toxicity Study (NOAEL) | 1.1 (0.6, 1.9) | 7.0 (2.5, 21) | 9.8 (7.7, 14) | 90% | > 99% |
| | Reproduction and Fertility Effects (LOAEL) | Reproduction and Fertility Effects (NOAEL) | 0.3 (0.1, 1.2) | NA | NA | NA | NA |
| Subacute-to-subchronic | Repeated Dose 28-Day Oral Toxicity Study in Rodents (LOAEL) | Repeated Dose 90-Day Oral Toxicity Study in Rodents (LOAEL) | 14 (3.4, 27) | NA | NA | NA | NA |
| 2) DCP LD50-to-LOAEL | Acute Oral Toxicity (LD50) | Repeated Dose 28-Day Oral Toxicity Study in Rodents (LOAEL) | 3.9 (2.0, 21) | NA | NA | NA | NA |
| | | Repeated Dose 90-Day Oral Toxicity Study in Rodents (LOAEL) | 229 (120, 587) | 95 (19, 289) | 807 (225, 60,923) | 45% | 73% |
| | | Prenatal Development Toxicity Study (LOAEL) | 65 (30, 194) | 37 (13, 129) | 50 (28, 308) | 2.0% | 18% |
| LOAEL-to-NOAEL | Repeated Dose 28-Day Oral Toxicity Study in Rodents (LOAEL) | Repeated Dose 28-Day Oral Toxicity Study in Rodents (NOAEL) | 14 (2.7, 32) | 3.0 (0.5, 20) | 8.3 (6.5, 14) | 93% | > 99% |
| | Repeated Dose 90-Day Oral Toxicity Study in Rodents (LOAEL) | Repeated Dose 90-Day Oral Toxicity Study in Rodents (NOAEL) | 0.3 (0.1, 1.5) | 3.0 (1.0, 11) | 23 (13, 231) | 75% | 98% |
| | Prenatal Development Toxicity Study (LOAEL) | Prenatal Development Toxicity Study (NOAEL) | 0.8 (0.3, 1.8) | 5.6 (1.8, 16) | 12 (8.7, 22) | 87% | > 99% |
| Subacute-to-subchronic | Repeated Dose 28-Day Oral Toxicity Study in Rodents (LOAEL) | Repeated Dose 90-Day Oral Toxicity Study in Rodents (LOAEL) | 58 (10, 179) | NA | NA | NA | NA |
| 3) ICP LD50-to-LOAEL | Acute Oral Toxicity (LD50) | Repeated Dose 28-Day Oral Toxicity Study in Rodents (LOAEL) | 18 (8.9, 52) | 4.2 (0.1, 17) | 8.4 (6.7, 12) | 94% | > 99% |
| | | Repeated Dose 90-Day Oral Toxicity Study in Rodents (LOAEL) | 309 (151, 790) | 145 (31, 664) | 3773 (2173, 13,492) | 3.0% | 27% |
| | | Prenatal Development Toxicity Study (LOAEL) | 37 (21, 71) | 55 (27, 142) | 67 (50, 95) | 56% | 93% |
| | | Reproduction and Fertility Effects (LOAEL) | 45 (9.0, 1806) | 11 (1.6, 33) | 25 (19, 37) | 39% | > 99% |
| | | Chronic Toxicity Studies (LOAEL) | 135 (40, 882) | 3.1 (0.1, 54) | 169 (79, 637) | 39% | 83% |

(continued on next page)

Table 3 (continued)

| UFs | | CTD comparisons: TC5 ratios (95% CIs) | | Individual UF probability distribution: 90th centile (95% CIs) | The likelihood (%) of ingredients being protected under a default UF | | |
|------------------------|---|---|---|--|---|---|--|
| | | All ingredients | Similar ingredients | | 10 | 100 | |
| | | LOAEL-to-NOAEL | Repeated Dose 28-Day Oral Toxicity Study in Rodents (LOAEL) Repeated Dose 90-Day Oral Toxicity Study in Rodents (LOAEL) Prenatal Development Toxicity Study (LOAEL) Reproduction and Fertility Effects (LOAEL) | | Repeated Dose 28-Day Oral Toxicity Study in Rodents (NOAEL) Repeated Dose 90-Day Oral Toxicity Study in Rodents (NOAEL) Prenatal Development Toxicity Study (NOAEL) Reproduction and Fertility Effects (NOAEL) | 1.8 (0.6, 4.3) 0.1 (0.04, 1.2) 1.1 (0.5, 1.9) 1.2 (0.1, 6.8) | 2.8 (0.2, 7.8) 2.5 (0.2, 23) 6.5 (2.5, 21) 14 (4.2, 77) |
| Subacute-to-subchronic | Chronic Toxicity Studies (LOAEL) Repeated Dose 28-Day Oral Toxicity Study in Rodents (LOAEL) | Chronic Toxicity Studies (NOAEL) Repeated Dose 90-Day Oral Toxicity Study in Rodents (LOAEL) | 1.3 (0.2, 5.9) 17 (4.5, 54) | 21 (1.2, 4399) NA | 33 (15, 192) NA | 72% NA | 98% NA |
| Subchronic-to-chronic | Prenatal Development Toxicity Study (LOAEL) Reproduction and Fertility Effects (LOAEL) | Chronic Toxicity Studies (LOAEL) Chronic Toxicity Studies (LOAEL) | 3.7 (0.9, 27) | NA | NA | NA | NA |
| | | Chronic Toxicity Studies (LOAEL) | 3.0 (0.06, 45) | NA | NA | NA | NA |

derived reproductive NOAELs were predicted to provide estimations of 61%, 2.0%, 56% the time for APC, DCP, and LCP, respectively (Table S11). In contrast, subacute NOAELs were predicted to be > 93% for the three use categories, respectively, if they were derived from corresponding LOAEL values from the Repeated Dose 28-Day Oral Toxicity Study in Rodents.

4. Discussion

4.1. Chemical toxicity distributions and thresholds of toxicological concern

Probabilistic modeling approaches have been incorporated in hazard and risk assessments, and management paradigms in an effort to estimate potential adverse effects to human health or the environment (Hanson and Solomon, 2002). For example, PHA using the species sensitive distribution (SSD) approach is based on a continuum of acute or chronic toxicity data for species from an array of taxonomic groups (Solomon and Takacs, 2002). When there is insufficient toxicity information, intraspecies endpoint sensitivity distribution (IESD) can be developed using a range of endpoints for a given species, and leveraged in PHA (Hanson and Solomon, 2002). Similarly, the CTD concept, utilizing toxicity information of a group of chemicals (e.g., a use category or common category) to the same species and same experimental endpoints, has also been used in PHA especially when data availability is limited (Brain et al., 2006). Based on this concept, CTDs can be assembled for various model species (e.g., *Daphnia magna* and *Pimephales promelas*) (Williams et al., 2011), test methods (e.g., in vivo and in vitro) (Dobbins et al., 2008), exposure durations (e.g., acute and chronic) (Berninger and Brooks, 2010), mode of action (MOA; e.g., inert chemicals, less inert chemicals, reactive chemicals, and specifically acting chemicals) (de Wolf et al., 2005), or chemical classes (e.g., Cramer class I, II, and III) (Munro et al., 1996), and plotted on the same axes for relative sensitivities comparisons. This approach provides insights to which organism, method, exposure duration, MOA, or chemical classes are more sensitive than others for a group of chemicals of interest and thus provide a robust mechanism for prioritization of testing. We recently expanded this approach to identify novel UFs and to derive TTCs for different chemical classes of cleaning product ingredients (Wang et al., 2018). In the present critical review and analysis of a unique mammalian toxicology dataset, we extended the application of CTDs using toxicity data of different study types for rats (oral) in three different cleaning product use categories (APC, DPC, LCP) and for ingredients from common product types or chemical categories within each use category, respectively (Figs. 1–3 and Figs. S1–S3).

The utility of a CTD, like an SSD or IESD, is largely dependent on the quantity and quality of data. Requirements on the minimum number of data points (from individual or multiple studies) for a distribution vary considerably. For instance, OECD (1992) requires at least of five data points, whereas the US EPA (Stephan et al., 1985) and ANZECC and ARMCANZ (2000) recommended eight. In this study, in order to be consistent with our previously published methods (Dobbins et al., 2008; Williams et al., 2011; Wang et al., 2018) and to make full use of datasets from the CPISI database, a dataset containing a minimum of five data points was considered to be suitable for an CTD. Nonetheless, the CTD approach is more robust with increasing data, and is thus more accurate with narrower variances (95% CIs) especially for the lower ends (e.g., TC1 and TC5) of a distribution. Therefore, further study is warranted when additional new hazard data are available, which can serve to update and refine the observations presented here.

Use of CTDs (or SSD/IESD) in PHA allows health and ecological risk assessors to generate a criterion threshold value to protect a desired level of effect to a particular endpoint (e.g., mortality, development or reproduction) for chemicals in a particular chemical class or a use category (e.g., APC, DCP, LCP). Considering SSDs in lower tiered risk assessment, for example, threshold levels (known as hazardous concentrations; HC) are selected based on what is considered an acceptable

1) APC

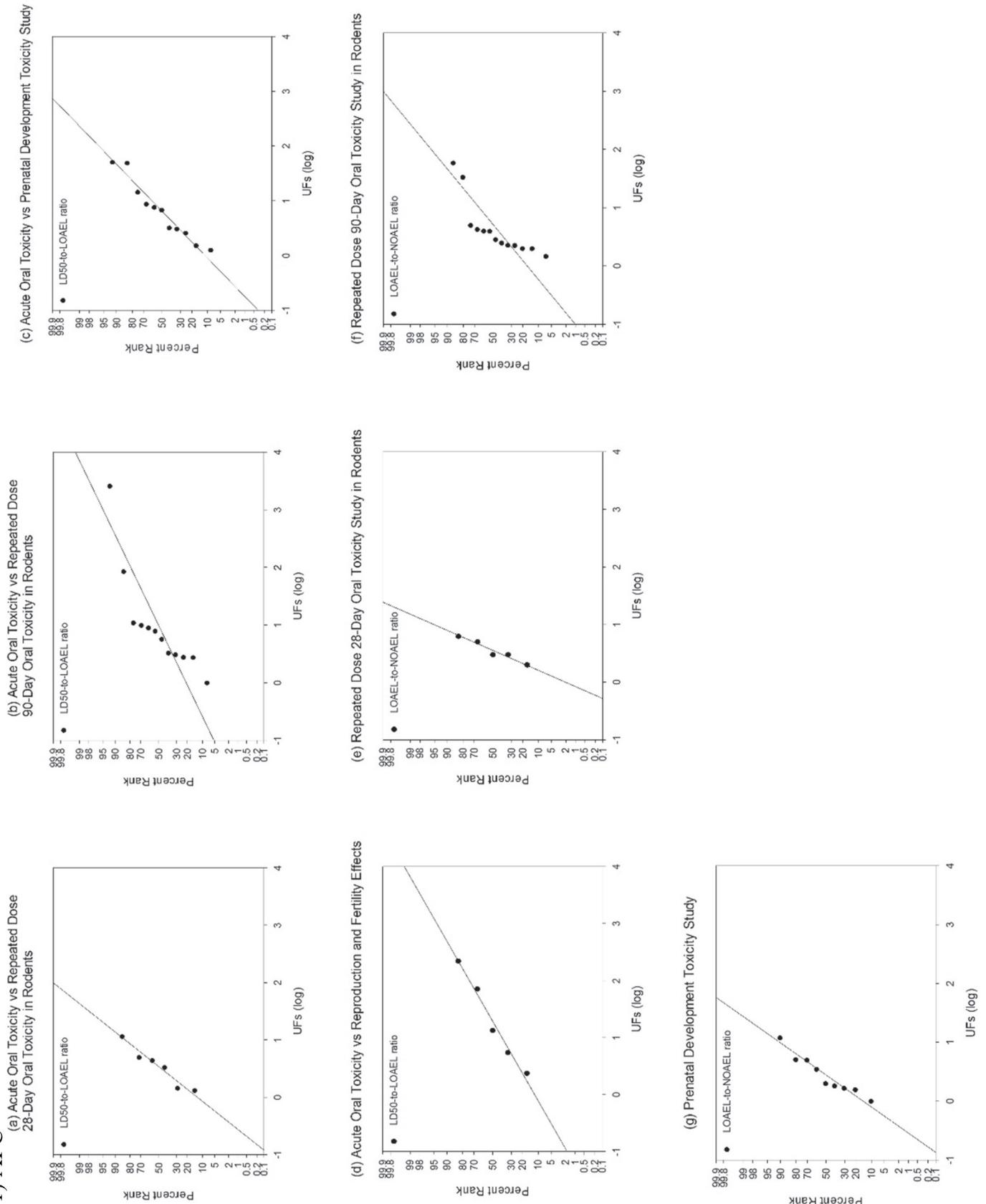


Fig. 4. Individual uncertainty factor probability distributions for median lethal dose (LD50) -to- lowest-observed-adverse-effect level (LOAEL) and LOAEL-to-no-observed-adverse-effect level (NOAEL) extrapolations.

3) LCP

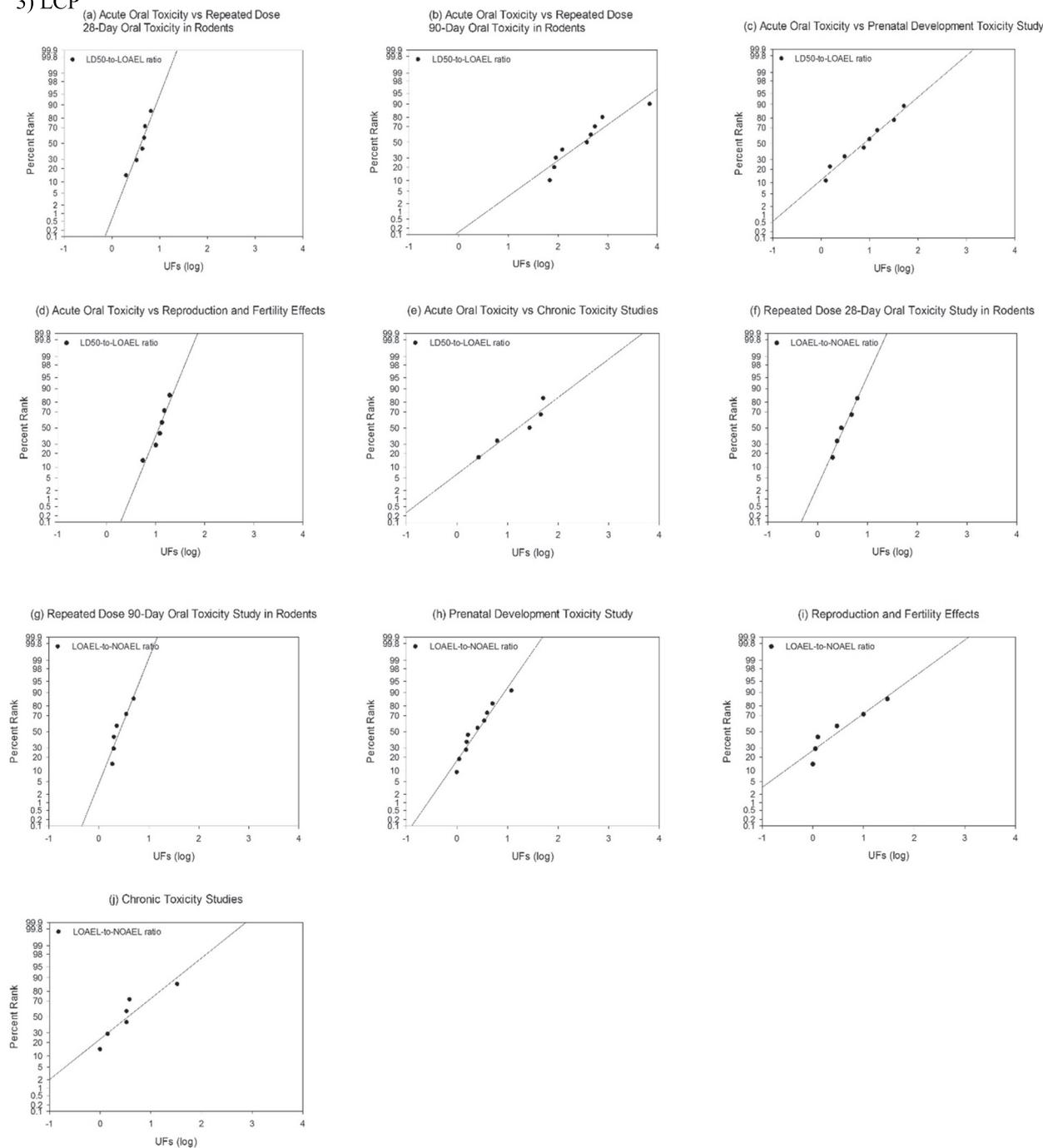


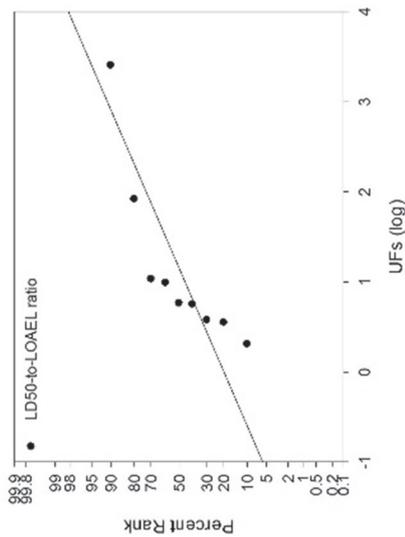
Fig. 4. (continued)

concentration of the substance, thereby protecting most organisms (typically 95%) in an assemblage of species (Wagner and Lokke, 1991; Aldenberg and Slob, 1993). The centile is normally set as 5% (i.e., HC5) by regulatory agencies and other practitioners, and is routinely used to establish water quality guidelines/criteria (Stephan et al., 1985; ANZECC and ARMCANZ, 2000; CCME, 2007) and to conduct ecological risk assessments (US EPA, 1998; Solomon and Takacs, 2002; ECHA, 2003). Other centiles have also been considered such as HC10 (Wang et al., 2014; Wang and Leung, 2015) or HC20 (Saari et al., 2017) in SSD comparisons due to data availability. Hanson and Solomon (2002) defined thresholds of toxicity as a low centile (0.1%) from EC10 distributions (IESD) and then used these centiles as toxicological benchmark concentrations, though they also suggested a higher centile (1%

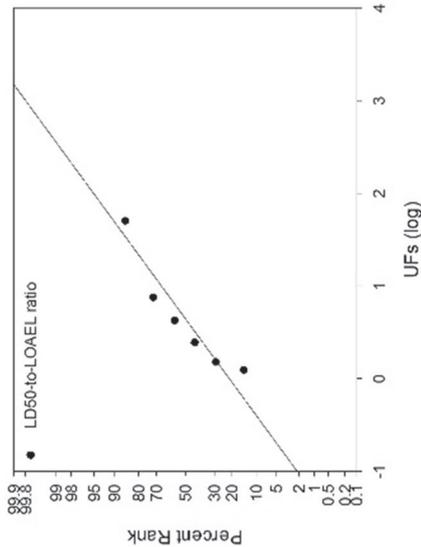
or 10%) may be more appropriate as these tend to be within the observed values with greater confidence in their representativeness. For CTDs, TC5s have been used to evaluate the relative sensitivities among different types of in vitro and in vivo assays (Dobbins et al., 2008; Dreier et al., 2015), between and among different distributions (Solomon et al., 2000; Wang et al., 2018), and identifications of TTCs for industrial organic chemicals (Munro et al., 1996) and for substances present at low levels in diet (Kroes et al., 2004). In this study, TCs (95% CIs) at different percentiles for a CTD were computed (Tables 2 and S2), which were further used to identify TTCs (e.g., TC1 or TC5), compare relative sensitivities between or among CTDs (i.e., TC5), and derive novel UFs through TCs ratios of pairwise CTDs comparisons (e.g., TC5-to-TC5 ratios). Additionally, the linear regression function for a

2) DCP

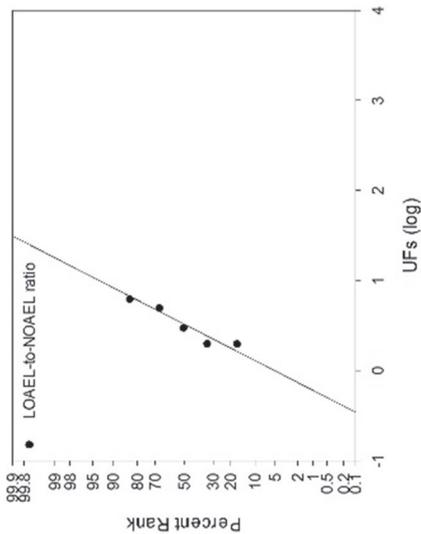
(a) Acute Oral Toxicity vs Repeated Dose 90-Day Oral Toxicity in Rodents



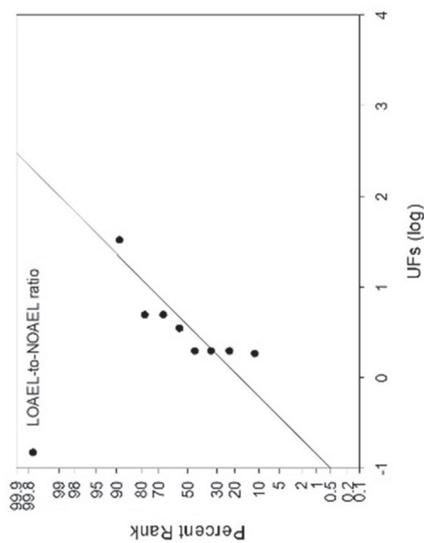
(b) Acute Oral Toxicity vs Prenatal Development Toxicity Study



(c) Repeated Dose 28-Day Oral Toxicity Study in Rodents



(d) Repeated Dose 90-Day Oral Toxicity Study in Rodents



(e) Prenatal Development Toxicity Study

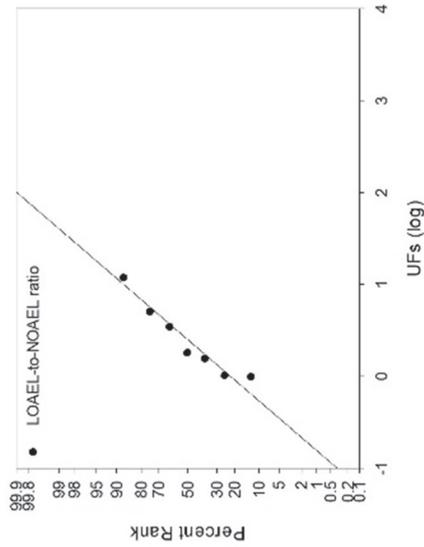


Fig. 4. (continued)

particular distribution was also derived in this study ($y = bx + a$; Table 2 and S2). These functions can be used by practitioners in future PHA to estimate the probability of finding a compound at or below a measured or predicted exposure threshold for a common endpoint (e.g., mortality, development or reproduction) of a defined type of toxicity test for a particular group of ingredients, such as cleaning product ingredients in a use category (APC, DCP or LCP) or a product types or chemical categories within a given product use category.

Representativeness and compositions of ingredients in a dataset may be another factor influencing CTDs comparisons. Thus, in the present study the general datasets consisting of all available LD50s and LOAELs/NOAELs were separated in major groups (e.g., product types or chemical categories), and the relative sensitivities between or among different product types (~3 order of magnitude differences) or chemical categories (same order of magnitude) were also examined (see Tables S3 and S4 for more details). With this information, it is feasible to further examine chemical effects on a particular group and minimize noise due to differences in ingredients compositions in a general CTD (Brix et al., 2001; Wang et al., 2014). For instance, the addition of acute LD50s data of more sensitive ingredients of ethers into a general acute LD50s dataset for APC would weigh the general acute LD50s to the lower end, while adding numerous less sensitive data points for aliphatic acids and salts could shift the distribution to the upper end, causing more or less sensitivity, depending on the data (Fig. 3a). The information on the relative sensitivities of ingredients from different product types within a use category are also important because of consumer interest level in what specific ingredients are used or safer; interest among stakeholders in sustainable alternative ingredients; and the potential for use of frame formulas in determining ingredient concentration as part of exposure assessment.

Previous studies demonstrated the possibility of using the CTD concept in TTCs developments for diverse industrial chemicals (de Wolf et al., 2005; Kroes et al., 2005; Connors et al., 2014; Belanger et al., 2015) and pharmaceuticals (Berninger and Brooks, 2010). The current study also supports the potential application of the CTD approach in TTCs development to a group of cleaning product ingredients in a use category (APC, DCP, or LAP) and bolsters support for use of this approach for ingredients in different product types or chemical categories within categories of a given product use. This approach is not only focused or limited to the identification of potential hazards, but also provides a quantitative estimate of potency (Veenstra and Kroese, 2005). As mentioned above, the derived TC values and their 95% CIs as listed in Table 2 (for LD50s and LOAELs) and Table S2 (for NOAELs), which rely on the most robust mammalian toxicity dataset currently available for cleaning product ingredients in a use category (APC, DCP, LCP), can represent TTC values for future PHA. In practice, risk assessors or authorities can decide whether TC1 or TC5 or their lower bands of 95% CIs should be used for a certain case of interest. In contrast to using a set of basic test data as the initial starting pointing for a screening-level risk assessment (HQ approach), particularly for chemicals with little or no data, this approach could be beneficial in both an industrial and regulatory setting for avoiding extensive toxicity testing and safety evaluations when human intake or environmental exposure are at or below such thresholds.

The TTC concept proposes a low level of exposure with a negligible risk can be identified for many chemicals, including those without toxicity information based on knowledge of their chemical structures. For cleaning products, previous assessments have shown that human exposure to household cleaning product ingredients is very low for a number of scenarios in which ingredients of interests comprise up to 30% of the product (e.g., aggregated exposure of soap, linear alkylbenzene sulfonate, and alkyl sulfates < 0.006 mg/kg bw/day; www.heraproject.com). In practice, traditional risk assessment scenarios may not be necessary if an extremely large screening-level margin of exposure (lower risk characterization), and the TTC approach, when data is limited, can provide mammalian doses for a chemical that may not

present a safety concern (EFSA, 2016). In the present study, our calculated TTCs were based on current available mammal hazard data for individual ingredients among use categories or product types. This PHA approach thus is not intended to account for effects of chemical mixtures, but can inform substitution of less hazardous chemicals within and among product uses. It further appear that this approach can be extended to other chemical classes and product types, depending on data availability.

It is also important to note that the BMD approach uses all dose-response data to estimate the shape of the overall dose-response relationship. It thus can be selected in an analogous way to the LOAEL/NOAEL approach (EFSA Scientific Committee, et al., 2017), because NOAEL and LOAEL values are influenced by ranges of treatment levels employed in toxicology experiments. In contrast to use of LOAELs/NOAELs, the BMD approach (lower band 95% level of BMD; BMDL) can account for additional toxicological considerations (e.g., dose-response information), and may result in a lower threshold (the lowest BMDL). In the present study we critically examined and performed meta-analyses on a highly unique mammalian database; however, BMDL (lower band 95% CI of BMD) information was sufficiently limited for ingredients in APC, DCP, and LCP categories, which precluded our ability to conduct PHA or to identify UFs using BMDLs for these cleaning product use categories. Further studies should be conducted when additional BMDL values are available, and to refine our current findings (e.g., TTCs and UFs).

4.2. Uncertainty factors including acute-to-chronic ratios

Two probability approaches (CTDs comparisons and individual UF probability distributions) were applied for UFs derivations in the present study. Rather than comparing the individual toxicity of chemicals, CTDs incorporate responses for multiple compounds during development of probability distributions. Therefore, UFs derived from CTDs comparisons (i.e., TCs ratios; Tables 3 and S6–S9) may provide sufficient protection through a more data driven approach. The PHA concept was also applied to estimate an overall UF covering most of ingredients of concern from the distributions of individual UFs for a particular use category (Tables 3, S10 and S11). Again, this approach incorporated UFs for multiple ingredients during development of individual UF probability distributions. Moreover, a Monte Carlo simulation approach was also applied and a range (i.e., 95% CIs) of UFs rather than a point estimate were estimated. This allows hazard and risk assessors to determine whether median (50%), or lower (e.g., 2.5%) or higher (e.g., 97.5%) bounds of these UFs should be used in their practice for certain chemicals of interest. To reduce potential bias introduced by representativeness of ingredients for each of pairwise CTDs comparisons, we also conducted coherent meta-analyses using the both datasets consisting of similar ingredients when there were sufficient data points (≥ 5). Due to data limitation of LOAELs of our critical review, this cannot be resolved currently for some, especially those exposure duration extrapolations. Further study should be conducted for those missing CTDs comparisons using similar ingredients when more hazard data are available.

The UF concept has been integrated in human hazard and risk assessment for the explicit derivation of a human limit value (HLV) for man from experimental or epidemiological toxicity data (e.g., NOAEL or LOAEL) (IPCS, 1994; ECETOC, 1995; US EPA, 2002). A 10-fold default UF can be used to account for various uncertainties, such as inter- and intraspecies variances, subchronic-to-chronic extrapolations, LOAEL-to-NOAEL ratios, adequacy of the total database, and route-to-route extrapolations (JECFA, 1961; JMPR, 1962; US EPA, 1988). The rationale for such default UF of 10 was also examined by Dourson and Stara (1983) for specific areas of uncertainty in risk calculations, so as to protect the majority of human populations from adverse health effects (Bigwood, 1973; Lu, 1979; Vettorazzi, 1980; Calabrese, 1985). However, results from our study present indicated that a factor of 10

might not be sufficient dependent on specific situations for cleaning product ingredients in the APC, DCP, or LCP categories of uses (Table 3).

When chronic NOAELs are not available, HLVs derivations can be established based on LOAEL values with an UF_{L-N} . Previous studies demonstrated that a factor of 10 or below is adequate for UF_{L-N} (Weil and McCollister, 1963; Dourson and Stara, 1983; Naumann and Weideman, 1995), while a range of UF_{L-N} values are commonly applied in hazard or risk assessment due to the severity of the adverse effect at the LOAEL, including values of 1–10 by US EPA (2002) and 3–10 by IPCS (1994). Based on the results in the present study from UF_{L-N} analysis using the CTDs comparison approach (e.g., TC5 ratios; using all or similar ingredients), UF_{L-N} 's were estimated as 0.9 (0.6, 1.6)–4.7 (2.9, 8.1) for APC, and 0.3 (0.1, 1.5)–3.0 (1.0, 11) for DCP, respectively using corresponding LOAEL-to-NOAEL TC5 ratios of Repeated Dose 90-Day Oral Toxicity Study in Rodents (Table 3). Corresponding UF_{S-C} 's (90th percentile) for APC and DPC were estimated as 46 (23, 245) and 23 (13, 231) from the individual UF probability distribution, while a factor of 10 was observed to only be protective 65% and 75% of the time. Considering both this approach and conservatism, UF_{L-N} of 46 and 23 could be more appropriate for APC and DCP, respectively, when using LOAELs as surrogates for NOAELs derivations. Factors of 2.5–5.3 tended to provide adequate protection on LOAEL-to-NOAEL extrapolations for LCP in the type of Repeated Dose 90-Day Oral Toxicity Study in Rodents.

Exposure duration extrapolations especially for subchronic-to-chronic extrapolation have been commonly used in hazard or risk assessments (ECETOC, 1995; Swartout, 1997; Dankovic et al., 2015). Previous studies also supported that a UF_{S-C} of 10 would provide sufficient protection (Weil and McCollister, 1963; Dourson and Stara, 1983; Pieters et al., 1998), despite other studies suggesting a more flexible approach in which 10 could be an upper bound estimate (McNamara, 1976; ECETOC, 1995; Dourson et al., 1996). With limited toxicity LOAEL data, only two UF_{S-C} 's (3.7 (0.9, 27) and 3.0 (0.06, 45)) were computed during our present study for developmental and reproductive responses in rats (oral) for LCP product chemicals. In this case, a factor of 10 would provide sufficient protection for subchronic Prenatal Development Toxicity Study or Reproductive and Fertility Effects to Chronic Toxicity Studies extrapolation, while such values possibly could be decreased from 10 to 3.7 and 3.0, respectively, based on the existing information. However, such UF_{S-C} 's should be further validated in examined and/or refined when more data are available for CTDs comparisons consisting of similar ingredients or individual UF probability distributions, and to derived separate UF_{S-C} 's for DCP and LCP, respectively.

Appropriate selection of ACRs is crucial for read across processes when extrapolating chronic LOAELs/NOAELs from acute LD50s for health hazard or risk assessment (Weil and Wright, 1967; Layton et al., 1987; Kramer et al., 1996). Though not consistently employed in practice, ACRs can be useful; for example, the State of Michigan (USA) uses a total ACR of 1,000,000 when setting a safety value using rodent LD50s from Acute Oral Toxicity studies as a point of departure (State of Michigan, 2017). Recently, we conducted an initial meta-analysis using all available acute LD50s and developmental and reproductive NOAELs for ingredients of cleaning products, and ACRs of 1996 and 317 were recommended when extrapolating developmental and reproductive chronic NOAELs, respectively from rodent acute LD50s (oral) (Wang et al., 2018). Comparatively, for cleaning product ingredients in a specific use category, results from the present study indicated that ACR of 169 could be more appropriate for APC of acute LD50s to chronic LOAELs extrapolations (Table 3). Comparatively, ACRs of 26–399, 124–2336 and 45–990 may be more appropriate for APC, DCP, and LCP, respectively when acute LD50s were used for corresponding chronic NOAELs extrapolations (Tables S7, S9 and S11). In addition, other LD50-to-NOAEL relationships among acute LD50s and subacute, subchronic or chronic LOAELs/NOAELs from other toxicity study types

were also examined in the present study. Taking APC product use chemicals, for example, if acute LD50s were used as a surrogate, ACRs of 25–46, 37–50, 37–67 may be useful for sublethal developmental LOAELs for APC, DCP and LCP, respectively (Table 3), while UFs of 28–66, 40–178, and 24–78 were identified, respectively using LD50s and NOAELs relationships (Tables S7, S9 and S11).

5. Conclusions

Lack of available mammalian toxicity data for most chemicals in commerce inherently challenges human health practitioners performing risk assessments of chemical threats. We reviewed a unique mammalian toxicology database for specific cleaning product use categories and then identified how using the CTDs concept in PHA can support assessments of these chemical classes and specific uses when robust toxicity data for hazard and risk assessment is lacking. We examined available hazard data (acute LD50s and sublethal LOAELs/NOAELs) generated from different types of standard mammalian toxicity studies for rats (oral) used in health risk assessments for APCs, DCPs and LCPs. Probabilistic CTDs were subsequently developed for all available ingredients in each use category, and specific product types or chemical categories, TCs (95% CIs) from each CTD were computed, and potential TTCs were identified. The present study also indicates that sub-distributions exist in terms of different product types and chemical categories within a particular use category with different relative sensitivities. Moreover, we extended a novel approach (Wang et al., 2018) to identify UFs particularly for LOAEL-to-NOAEL and exposure duration extrapolations (e.g., subchronic-to-chronic), and ACRs considering LD50-to-LOAEL/NOAEL relationships for ingredients in the APC, DCP and LCP categories of common household products, respectively. Findings from this study are anticipated to be useful for future mammalian health hazard and risk assessment, such as TTC identification, regulatory data dossier development through read across, chemical substitutions, and screening-level health hazard and risk assessment, especially when limited to no toxicity data is available for industrial chemicals.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2019.01.079>.

Conflict of interest

P DeLeo worked for the American Cleaning Institute during this study. The authors have no other conflicts of interest.

Acknowledgements

This work research was supported by an unrestricted grant through the American Cleaning Institute's Cleaning Product Ingredient Safety Initiative. Additional support was provided by Baylor University. Co-author ES Williams currently works for the National Aeronautic and Space Administration. However, this work was conducted while this author was employed at Baylor University, and as such the findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the National Aeronautic and Space Administration.

References

- ACI, 2016. Work Plan for the American Cleaning Institute Cleaning Product Ingredient Safety Initiative - Task 4: Screening Level Risk Assessment for Cleaning Product Ingredients. EA Engineering, Science and Technology; USA & Baylor University, Hunt Valley, MD; Waco, TX, USA.
- Aldenberg, T., Slob, W., 1993. Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. *Ecotoxicol. Environ. Saf.* 25, 48–63.
- Anderson, T.W., Darling, D.A., 1954. A test of goodness of fit. *J. Am. Stat. Assoc.* 49, 765–769.
- ANZECC & ARMCANZ, 2000. Australian and New Zealand guidelines for fresh and marine water quality. In: National Water Quality Management Strategy Paper No. 4.

- Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, Canberra, Australia.
- Api, A.M., Belsito, D., Bruze, M., Cadby, P., Calow, P., Dagli, M.L., Dekant, W., Ellis, G., Fryer, A.D., Fukayama, M., Griem, P., Hickey, C., Kromidas, L., Lalko, J.F., Liebler, D.C., Miyachi, Y., Politano, V.T., Renskers, K., Ritacco, G., Salviato, D., Schultz, T.W., Sipes, I.G., Smith, B., Vitale, D., Wilcox, D.K., 2015. Criteria for the Research Institute for Fragrance Materials, Inc. (RIFM) safety evaluation process for fragrance ingredients. *Food Chem. Toxicol.* 82 (Suppl), S1–S19.
- Barlow, S., 2005. Threshold of toxicological concern (TTC): a tool for assessing substances of unknown toxicity present at low levels in the diet. In: ILSI Europe Concise Monograph Series (Brussels, Belgium).
- Belanger, S.E., Sanderson, H., Embry, M.R., Coady, K., DeZwart, D., Farr, B.A., Gutsell, S., Halder, M., Sternberg, R., Wilson, P., 2015. It is time to develop ecological thresholds of toxicological concern to assist environmental hazard assessment. *Environ. Toxicol. Chem.* 34, 2864–2869.
- Bernauer, U., Heinemeyer, G., Heinrich-Hirsch, B., Ulbrich, B., Gundert-Remy, U., 2008. Exposure-triggered reproductive toxicity testing under the REACH legislation: a proposal to define significant/relevant exposure. *Toxicol. Lett.* 176, 68–76.
- Berninger, J.P., Brooks, B.W., 2010. Leveraging mammalian pharmaceutical toxicology and pharmacology data to predict chronic fish responses to pharmaceuticals. *Toxicol. Lett.* 193, 69–78.
- Bigwood, E.J., 1973. The acceptable daily intake of food additives. *Crit. Rev. Toxicol.* 2, 41–93.
- Blackburn, K., Stickney, J.A., Carlson-Lynch, H.L., McGinnis, P.M., Chappell, L., Felter, S.P., 2005. Application of the threshold of toxicological concern approach to ingredients in personal and household care products. *Regul. Toxicol. Pharmacol.* 43, 249–259.
- Brain, R.A., Sanderson, H., Sibley, P.K., Solomon, K.R., 2006. Probabilistic ecological hazard assessment: evaluating pharmaceutical effects on aquatic higher plants as an example. *Ecotoxicol. Environ. Saf.* 64, 128–135.
- Brix, K.V., Deforest, D.K., Adams, W.L., 2001. Assessing acute and chronic copper risk to freshwater aquatic life using species sensitivity distributions for different taxonomic groups. *Environ. Toxicol. Chem.* 20, 1846–1856.
- Brown, R., Carter, J., Dewhurst, I., Stephenson, C., Tessier, S., 2009. Applicability of thresholds of toxicological concern in the dietary risk assessment of metabolites, degradation and reaction products of pesticides. In: Scientific/technical Report Submitted to EFSA. Grant Agreement EFSA/PPR/2008/01. Chemical Regulation Directorate (CRD), UK.
- Calabrese, E.J., 1985. Uncertainty factors and interindividual variation. *Regul. Toxicol. Pharmacol.* 6, 190–196.
- CCME, 2007. Protocol for the derivation of water quality guidelines for the protection of aquatic life. In: Canadian Environmental Quality Guidelines. Canadian Council of Ministers of the Environment, Winnipeg, Canada.
- Cheeseman, M.A., Machuga, E.J., Bailey, A.B., 1999. A tiered approach to threshold of regulation. *Food Chem. Toxicol.* 37, 387–412.
- Connors, K.A., Voutchkova-Kostal, A.M., Kostal, J., Anastas, P., Zimmerman, J.B., Brooks, B.W., 2014. Reducing aquatic toxicity: Probabilistic hazard evaluation of sustainable molecular design guidelines. *Environ. Toxicol. Chem.* 33, 1894–1902.
- Cramer, G.M., Ford, R.A., Hall, R.A., 1978. Estimation of toxic hazard - a decision tree approach. *Food Chem. Toxicol.* 16, 255–276.
- Dankovic, D.A., Naumann, B.D., Maier, A., Dourson, M.L., Levy, L.S., 2015. The scientific basis of uncertainty factors in setting occupational exposure limits. *J. Occup. Environ. Hyg.* 12, S55–S68.
- de Wolf, W., Siebel-Sauer, A., Lecloux, A., Koch, V., Holt, M., Feijtel, T., Comber, M., Boeije, G., 2005. Mode of action and aquatic exposure thresholds of no concern. *Environ. Toxicol. Chem.* 24, 479–485.
- DeLeo, P.C., Ciarlo, M., Pacelli, C., Greggs, W., Williams, E.S., Scott, W.C., Wang, Z., Brooks, B.W., 2018. Cleaning product ingredient safety: what is the current state of availability of information regarding ingredients in products and their function? *ACS Sustain. Chem. Eng.* 6, 2094–2102.
- Dobbins, L.L., Brain, R.A., Brooks, B.W., 2008. Comparison of the sensitivities of common in vitro and in vivo assays of estrogenic activity: application of chemical toxicity distributions. *Environ. Toxicol. Chem.* 27, 2608–2616.
- Dourson, M.L., Stara, J.F., 1983. Regulatory history and experimental support of uncertainty (safety) factors. *Regul. Toxicol. Pharmacol.* 3, 224–238.
- Dourson, M.L., Felter, S.P., Robinson, D., 1996. Evolution of science-based uncertainty factors in noncancer risk assessment. *Regul. Toxicol. Pharmacol.* 24, 108–120.
- Dreier, D.A., Connors, K.A., Brooks, B.W., 2015. Comparative endpoint sensitivity of in vitro estrogen agonist assays. *Regul. Toxicol. Pharmacol.* 75, 185–193.
- EC, 1967. Council Directive 67/548/EEC of 27 June 1967 on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labeling of dangerous substances.
- EC, 2000. Commission regulation (EC) no. 1565/2000 of 18 July laying down the measures necessary for the adoption of an evaluation programme in application of regulation (EC) no. 2232/96 of the European Parliament and of the Council. In: Official Journal of the European Communities 19.7.2000, Annex I, L180/8–L180/16.
- ECETOC, 1995. Assessment factors in human health risk assessment. In: Technical Report, No. 68. European Center for Ecotoxicology and Toxicology of Chemicals (ECETOC), Brussels, Belgium.
- ECETOC, 2004. Targeted risk assessment. In: Technical Report No.93. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium.
- ECHA, 2003. Technical Guidance Document on Risk Assessment-Part II. Institute for Health and Consumer Protection, Ispra, Italy.
- EFSA, 2016. Review of the threshold of toxicological concern (TTC) approach and development of new TTC decision tree. In: EFSA (European Food Safety Authority) Supporting Publication 2016: EN-1006, (50 pp).
- EFSA Scientific Committee, Hardy, A., Benford, D., Halldorsson, T., Jeger, M.J., Knutsen, K.H., More, S., Mortensen, A., Naegeli, H., Noteborn, H., Ockleford, C., Ricci, A., Rychen, G., Silano, V., Solecki, R., Turck, D., Aerts, M., Bodin, L., Davis, A., Edler, L., Gundert-Remy, U., Sand, S., Slob, W., Böttex, B., Abrahamantes, J.C., Marques, D.C., Kass, G., Schlatter, J.R., 2017. Update: guidance on the use of the benchmark dose approach in risk assessment. *EFSA J.* 15, 4658 (41 pp).
- EMA, 2006. Guideline on the Limits of Genotoxic Impurities, CPMP/SWP/5199/02, EMEA/CHMP/QWP/251344/2006. European Medicines Agency (EMA), Committee for Medicinal Products for Human Use (CHMP), London, UK.
- EMA, 2008. Guideline on the Assessment of Genotoxicity of Herbal Substances/preparations. EMEA/HMPC/107079/2007. European Medicines Agency (EMA), Committee on Herbal Medicinal Products (HMPC), London, UK.
- Flamm, W.G., Lake, L.R., Lorentzen, R.J., Rulis, A.M., Schwartz, P.S., Troxell, T.C., 1987. Carcinogenic potencies and establishment of a threshold of regulation for food contact substances. In: De Minimis Risk, Whipple. Plenum, New York, NY, USA.
- Frawley, J.P., 1967. Scientific evidence and common sense as a basis for food-packaging regulations. *Food Cosmet. Toxicol.* 5, 293–308.
- Gold, L.S., Sawyer, C.B., Magaw, R., Backman, G.M., de Veciana, M., Levinson, R., Hooper, N.K., Havender, W.R., Bernstein, L., Peto, R., Pike, M.C., Ames, B.N., 1984. A carcinogenic potency database of the standardized results of animal bioassay. *Environ. Health Perspect.* 58, 9–319.
- Grubbs, F.E., 1969. Procedures for detecting outlying observations in samples. *Technometrics* 11, 1–21.
- Hanson, M.L., Solomon, K.R., 2002. New technique for estimating thresholds of toxicity in ecological risk assessment. *Environ. Sci. Technol.* 36, 3257–3264.
- IPCS, 1994. Assessing human health risks of chemicals: derivation of guidance values for health-based exposure limits. In: Environmental Health Criteria No. 170. International Programme on Chemical Safety (IPCS), World Health Organization (WHO), Geneva, Switzerland.
- JECFA, 1961. Evaluation of the carcinogenic hazards of food additives. In: 55th Report of the Joint FAO/WHO Expert Committee on Food Additives. FAO Nutrition Meetings Report Series, No. 29; WHO Technical Report Series, No. 220 WHO, Geneva, Switzerland.
- JECFA, 1995. Evaluation of certain food additives and contaminants. In: 44th Report of the Joint FAO/WHO Expert Committee on Food Additives. WHO Technical Report Series, No. 859 WHO, Geneva, Switzerland.
- JECFA, 2006. Evaluation of certain food additives and contaminants. In: 65th Report of the Joint FAO/WHO Expert Committee on Food Additives. WHO Technical Report Series, No. 934 WHO, Geneva, Switzerland.
- JMPR, 1962. Principles governing consumer safety in relation to pesticide residues. Report of a meeting of a WHO Expert Committee on Pesticide Residues held jointly with the FAO Panel of Experts on the use of pesticides in agriculture. In: FAO Plant Production and Protection Division Report, No. PL/1961/11. WHO Technical Report Series, No. 240.
- Kramer, H.J., van den Ham, W.A., Slob, W., Pieters, M.N., 1996. Conversion factors estimating indicative chronic no-observed-adverse-effect levels from short-term toxicity data. *Regul. Toxicol. Pharmacol.* 23, 249–255.
- Kroes, R., Renwick, A.G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., Schilter, B., Schlatter, J., van Schothorst, F., Vos, J.G., Würtzen, G., 2004. Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. *Food Chem. Toxicol.* 42, 65–83.
- Kroes, R., Kleiner, J., Renwick, A., 2005. The threshold of toxicological concern concept in risk assessment. *Toxicol. Sci.* 86, 226–230.
- Kroes, R., Renwick, A.G., Feron, V., Galli, C.L., Gibney, M., Greim, H., Guy, R.H., Lhuguenot, J.C., van de Sandt, J.J., 2007. Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. *Food Chem. Toxicol.* 45, 2533–2562.
- Layton, D.W., Mallon, B.J., Rosenblatt, D.H., Small, M.J., 1987. Deriving allowable daily intakes for systemic toxicants lacking chronic toxicity data. *Regul. Toxicol. Pharmacol.* 7, 96–112.
- Lu, F.C., 1979. Assessment at an international level of health hazards to man of chemicals shown to be carcinogenic in laboratory animals. In: Coulston, E.F. (Ed.), *Regulatory Aspects of Carcinogenesis*. Academic Press, New York, NY, USA, pp. 315–328.
- McNamara, B.P., 1976. Concepts in health evaluation of commercial and industrial chemicals. In: Mehlmán, M.A., Shapiro, R.E., Blumenthal, H. (Eds.), *Advances in Modern Technology, 1, Part 1, New Concepts in Safety Evaluation*. Hemisphere, Washington, DC, USA, pp. 61–140.
- Munro, I.C., Ford, R.A., Kennepohl, E., Sprenger, J.G., 1996. Correlation of structural class with no-observed-effect levels: a proposal for establishing a threshold of concern. *Food Chem. Toxicol.* 34, 829–867.
- Naumann, B.D., Weideman, P.A., 1995. Scientific basis for uncertainty factors used to establish occupational exposure limits for pharmaceutical active ingredients. *Hum. Ecol. Risk Assess.* 1, 590–613.
- OECD, 1992. Report of the OECD Workshop on the Extrapolation of Laboratory Aquatic Toxicity Data to the Real Environment. *Env. Monograph NO. 59. OCDE/GD (92)169*. (Paris, France).
- Pieters, M.N., Kramer, H.J., Slob, W., 1998. Evaluation of the uncertainty factor for subchronic-to-chronic extrapolation: statistical analysis of toxicity data. *Regul. Toxicol. Pharmacol.* 27, 108–111.
- Renwick, A.G., 2004. Toxicology databases and the concept of thresholds of toxicological concern as used by the JECFA for the safety evaluation of flavouring agents. *Toxicol. Lett.* 149, 223–234.
- Rowbotham, A.L., Gibson, R.M., 2011. Exposure-driven risk assessment: applying exposure-based waiving of toxicity tests under REACH. *Food Chem. Toxicol.* 49, 1661–1673.

- Rulis, 1986. De minimis and the threshold of regulation. In: Felix, C.W. (Ed.), *Food Protection Technology*. Chelsea/Lewis, MI, USA, pp. 29–37.
- Rulis, 1989. Establishing a threshold of regulation. In: Bonin, J., Stevenson, D. (Eds.), *Risk Assessment in Setting National Priorities*. Plenum, New York, NY, USA, pp. 221–278.
- Saari, G.N., Wang, Z., Brooks, B.W., 2017. Revisiting inland hypoxia: diverse exceedances of dissolved oxygen thresholds for freshwater aquatic life. *Environ. Sci. Pollut. Res.* <https://doi.org/10.1007/s11356-017-8908-6>.
- Shapiro, S.S., Francia, R.S., 1972. An approximate analysis of variance test for normality. *J. Am. Stat. Assoc.* 67, 215–216.
- Smith, R.L., Adams, T.B., Cohen, S.M., Doull, J., Feron, V.J., Goodman, J.I., Hall, R.L., Marnett, L.J., Portoghese, P.S., Wadell, W.J., Wagner, B.M., 2004. Safety evaluation of natural flavour complexes. *Toxicol. Lett.* 149, 197–207.
- Smith, R.L., Cohen, S.M., Doull, J., Feron, V.J., Goodman, J.I., Marnett, L.J., Portoghese, P.S., Wadell, W.J., Wagner, B.M., Hall, R.L., Higley, N.A., Lucas-Gavin, C., Adams, T.B., 2005. A procedure for the safety evaluation of natural flavor complexes used as ingredients in food: essential oils. *Food Chem. Toxicol.* 43, 345–363.
- Sokal, R.R., Rohlf, F.J., 1995. *Biometry*, 3rd ed. W.H. Freeman and Company, New York, NY, USA.
- Solomon, K.R., Takacs, P., 2002. Probabilistic risk assessment using species sensitivity distributions. In: Posthuma, L., Suter IIG.W., Traas, T.P. (Eds.), *Species Sensitivity Distributions in Ecotoxicology*. CRC Press, Boca Raton, FL, USA, pp. 285–313.
- Solomon, K., Giesy, J., Jones, P., 2000. Probabilistic risk assessment of agrochemicals in the environment. *Crop Prot.* 19, 649–655.
- State of Michigan, 2017. *Guidesheet V: guidance for providing toxicity information to determine a rule 2225(5) standard*. http://www.michigan.gov/documents/deq/wrd-groundwater-p22-toxicity-v_564946_7.pdf.
- Stephan, C.E., Mount, D.I., Hansen, D.J., Gentile, J.H., Chapman, G.A., Brungs, W.A., 1985. Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses. PB85-227049. National Technical Information Service, Springfield, VA, USA.
- Suter II, G.W., 1995. Introduction to ecological risk assessment for aquatic toxic effects. In: Rand, G.M. (Ed.), *Fundamentals of Aquatic Toxicology: Effects, Environmental Fate and Risk Assessment*. Chapter 28, 2nd ed. Taylor and Francis, Washington DC, USA, pp. 803–816.
- Swartout, J., 1997. Exposure-duration uncertainty factors for the RfD. *Fundam. Appl. Toxicol.* 209.
- Tietjen, G.L., Moore, R.H., 1972. Some Grubbs-type statistics for the detection of several outliers. *Technometrics* 14, 583–597.
- US EPA, 1988. *Recommendations for and Documentation of Biological Values for Use in Risk Assessment*. EPA/600/6-87/008 (NTIS PB88179874). US Environmental Protection Agency, Washington, DC, USA.
- US EPA, 1998. *Guidelines for Ecological Risk Assessment*. EPA/630/R-95/002F. Risk Assessment Forum. US Environmental Protection Agency, Washington, DC, USA.
- US EPA, 2002. *A Review of the Reference Dose and Reference Concentration Processes*. EPA/630/P-02/002F. US Environmental Protection Agency, Washington, DC, USA.
- US FD&C Act, 1958. *Federal Food, Drug, and Cosmetic Act, as Amended Title 21 U.S. Code*. U.S. Governmental Printing Office, Washington, DC, USA.
- US FDA, 1995. *Food additives: threshold of regulation of substances used in food contact articles: final rule*. *Fed. Regist.* 60, 36582–36596.
- US FDA, 2008. *Guidance for Industry Genotoxic and Carcinogenic Impurities in Drug Substances and Products: Recommended Approaches*. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER).
- Veenstra, G., Kroese, D., 2005. *Threshold of Toxicological Concern A Concept in Toxicological Risk Assessment*. (Discussion document, SEG1. March 14-15, 2005. RIP3.3 SEG1-19).
- Vettorazzi, G., 1980. Safety factors and their application in the toxicological evaluation. In: *The Evaluation of Toxicological Data for the Protection of Public Health*. Pergamon Press, Oxford, UK, pp. 207–223.
- Wagner, C., Lokke, H., 1991. Estimation of ecotoxicological protection levels from NOEC toxicity data. *Water Res.* 25, 1237–1242.
- Wang, Z., Leung, K.M.Y., 2015. Effects of unionised ammonia on tropical freshwater organisms: implications on temperate-to-tropic extrapolation and water quality guidelines. *Environ. Pollut.* 205, 240–249.
- Wang, Z., Kwok, K.W.H., Lui, G.C.S., Zhou, G.-J., Lee, J.-S., Lam, M.H.W., Leung, K.M.Y., 2014. The difference between temperate and tropical saltwater species' acute sensitivity to chemicals is relatively small. *Chemosphere* 105, 31–43.
- Wang, Z., Scott, W.C., Williams, E.S., Ciarlo, M., DeLeo, P., Brooks, B.W., 2018. Identification of novel uncertainty factors and thresholds of toxicological concern for health hazard and risk assessment: application to cleaning product ingredients. *Environ. Int.* 113, 357–376.
- Weil, C.S., McCollister, D.D., 1963. Relationship between short- and long-term feeding studies in designing an effective toxicity test. *J. Agric. Food Chem.* 11, 486–491.
- Weil, C.S., Wright, G.J., 1967. Intra- and interlaboratory comparative evaluation of single oral test. *Toxicol. Appl. Pharmacol.* 11, 378–388.
- Williams, E.S., Berninger, J.P., Brooks, B.W., 2011. Application of chemical toxicity distributions to ecotoxicology data requirements under REACH. *Environ. Toxicol. Chem.* 30, 1943–1954.