Environment International xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Review article

# Identification of novel uncertainty factors and thresholds of toxicological concern for health hazard and risk assessment: Application to cleaning product ingredients

Zhen Wang[a], W. Casan Scott[a], E. Spencer Williams[a], Michael Ciarlo[c], Paul C. DeLeo[d], Bryan W. Brooks[a,b,*]

[a] Environmental Health Science Program, Department of Environmental Science, Baylor University, Waco, TX, USA
[b] Institute of Biomedical Studies, Baylor University, Waco, TX, USA
[c] EA Engineering, Science & Technology, Inc., Baltimore, MD, USA
[d] American Cleaning Institute, Washington, DC, USA

ARTICLE INFO

Handling Editor: Robert Letcher

*Keywords:*
Uncertainty factors
Health hazard and risk assessment
Cleaning products
Chemical toxicity distributions
Thresholds of toxicological concern

ABSTRACT

Uncertainty factors (UFs) are commonly used during hazard and risk assessments to address uncertainties, including extrapolations among mammals and experimental durations. In risk assessment, default values are routinely used for interspecies extrapolation and interindividual variability. Whether default UFs are sufficient for various chemical uses or specific chemical classes remains understudied, particularly for ingredients in cleaning products. Therefore, we examined publicly available acute median lethal dose (LD50), and reproductive and developmental no-observed-adverse-effect level (NOAEL) and lowest-observed-adverse-effect level (LOAEL) values for the rat model (oral). We employed probabilistic chemical toxicity distributions to identify likelihoods of encountering acute, subacute, subchronic and chronic toxicity thresholds for specific chemical categories and ingredients in cleaning products. We subsequently identified thresholds of toxicological concern (TTC) and then various UFs for: 1) acute (LD50s)-to-chronic (reproductive/developmental NOAELs) ratios (ACRs), 2) exposure duration extrapolations (e.g., subchronic-to-chronic; reproductive/developmental), and 3) LOAEL-to-NOAEL ratios considering subacute/acute developmental responses. These ratios (95% CIs) were calculated from pairwise threshold levels using Monte Carlo simulations to identify UFs for all ingredients in cleaning products. Based on data availability, chemical category-specific UFs were also identified for aliphatic acids and salts, aliphatic alcohols, inorganic acids and salts, and alkyl sulfates. In a number of cases, derived UFs were smaller than default values (e.g., 10) employed by regulatory agencies; however, larger UFs were occasionally identified. Such UFs could be used by assessors instead of relying on default values. These approaches for identifying mammalian TTCs and diverse UFs represent robust alternatives to application of default values for ingredients in cleaning products and other chemical classes. Findings can also support chemical substitutions during alternatives assessment, and data dossier development (e.g., read across), identification of TTCs, and screening-level hazard and risk assessment when toxicity data is unavailable for specific chemicals.

## 1. Introduction

Human health risk assessment characterizes the likelihood of adverse health effects through a structured and expert review of hazard, dose-response, and exposure information (WHO, 1999). It is common for health organizations to utilize a “safe” dose concept (human limit values; HLVs) such as occupational exposure limits (OELs), acceptable daily intakes (ADIs), tolerable daily intake (TDI), reference dose (RfD), or reference concentrations (RfC) in the dose-response assessment of noncancer toxicity for chemicals. Human epidemiological data used for risk assessment are lacking for many chemicals, largely due to limitations of information on exposure in occupational and residential settings and the small number of available clinical studies that cover very few chemical substances. Therefore, human safety values are typically derived from existing mammalian information such as an animal threshold dose ($threshold_{animal}$), a no-effect level (e.g., no-observed-adverse-effect level (NOAEL), a benchmark dose level (BMDL)) or an effect level (e.g., lowest-observed-adverse-effect level (LOAEL)). The

* Corresponding author at: One Bear Place #97266, Waco, TX 76798, USA.
*E-mail address:* Bryan_Brooks@baylor.edu (B.W. Brooks).

https://doi.org/10.1016/j.envint.2018.02.011
Received 7 November 2017; Received in revised form 7 February 2018; Accepted 8 February 2018

NOAEL (or LOAEL) is the common point of departure (PoD) in establishing HLVs, while BMDL, defined as the lower 95% confidence limit (95% CI) of the critical effect dose, is an alternative considering the dose-response relationship as a whole compared to NOAEL determinations (Crump, 1984; US EPA, 2002). Because actual relationships among $threshold_{animal}$ and $threshold_{human}$ are largely unknown, uncertainty is inherent during interspecies extrapolation. Uncertainty factors (UFs; also known as assessment or safety factors) are then applied to the PoD values from toxicology studies in an attempt to account for specific types of uncertainties (e.g., intra- ($UF_{H-H}$) and interspecies variability from animal to human ($UF_{A-H}$), subchronic-to-chronic ($UF_{S-C}$), LOAEL-to-NOAEL ($UF_{L-N}$), adequacy of the total dataset ($UF_D$), and route-to-route extrapolations (e.g., oral-to-dermal/inhalation)).

Initial publications on uncertainty factors for health risk assessment practice were presented by Lehman and Fitzhugh (1954) of the United States Food and Drug Administration (FDA). They suggested that a safety level for food additives or contaminants can be derived by dividing a chronic NOAEL from animal studies by a 100-fold UF (10 for $UF_{H-H}$ × 10 for $UF_{A-H}$). This approach was then adopted by the JECFA (1961) and by the JMPR (1962) in the 1960s for ADIs derivations. Rationale for the 10-fold UF was also examined by Dourson and Stara (1983) for specific areas of uncertainty in risk calculations; such values were considered sufficient to protect the majority of human populations from adverse health effects (Bigwood, 1973; Lu, 1979; Vettorazzi, 1980; Calabrese, 1985). Subsequently, this default UF concept has been adopted in regulatory guidelines and hazard and risk assessment practice around the globe (Table 1). For example, US EPA (1988) promulgated UFs as 10-fold to account for uncertainties during estimations of RfDs. Further, a modifying factor (MF) ranging from < 1 to up to 10 was applied when a database includes a very large number of animals per dose level (MF ≤ 1). Calabrese and Gilbert (1993) suggested modifying uncertainty by the lack of total independence of these factors. For example, a factor of 5 was suggested to account for sufficient protection for most humans (e.g., less-than-lifetime animal study). To account for differences in size among animals and human, the Health Council of the Netherlands (1985) presented an allometric scaling approach for the establishment of HLVs based on body weight or caloric demand. For instance, if extrapolation on basis of caloric demands is chosen, the Health Council of the Netherlands (1985) proposed a factor of 30 over 10 for $UF_{H-H}$ and $UF_{A-H}$, respectively, and an MF of 3 for observation errors. Allometric approaches have also been adopted by TNO (Stevenson et al., 1995a; Stevenson et al., 1995b; EC, 1996) and ECETOC (ECETOC, 1995, 2003) for $UF_{A-H}$ in establishment of HLVs on the basis on caloric demands (equivalent to $(body\ weight)^{0.75}$; see Table 1 for more details).

The use of default UFs inherently imparts precaution to human health risk assessment when there are insufficient hazard data. However, scientific justification for the size of the UFs is often lacking, and the selection of default values can reflect policy decisions rather than scientific determinations. Fortunately, an increasing understanding of intra- and interspecies sensitivity, knowledge of comparative mechanisms of action (MOA), and compilations of existing data for computation toxicology efforts has led to improvements that allow for incorporation of more robust scientific data during dose-response assessments of non-cancer toxicity. Such advances can afford opportunities to use data-derived UFs (also known as chemical-specific assessment factors; CSAFs) rather than standard default values. For example, Lewis et al. (1990) developed an alternative methodology (LLN model) for establishing guidelines for determining acceptable atmospheric emissions. The LLN model intended: 1) to separate scientific judgments from policy/value judgments (i.e., UFs 1–10); 2) to provide a plausible estimate of actual risk from a defined exposure; and 3) to value UFs over independent assessments. This approach made a distinction between factor adjustments, for which the appropriate magnitude of UFs was considered within the purview of a risk assessor.

Renwick (1991) and Renwick (1993) examined the nature of the UF (100-fold) for ADIs derivations and attempted to provide a scientific basis to the default UF of 10 for intra- and interspecies differences. Specifically, Renwick proposed to partition each UF to sub-factors to allow for evaluation of the differences of toxicokinetics (TK) and toxicodynamics (TD) separately. Relative magnitudes of TK and TD variations between and within species were examined in detail, based on data availability. It was found that TK differences are generally greater than TD differences, and the overall 10-fold default UF for intra- and interspecies was subdivided into sub-factors of 4 for TK and 2.5 for TD. This approach offered a possibility to incorporate mechanistic human pharmacokinetic and pharmacodynamic information during establishment of factors for pharmaceuticals in risk assessment (Naumann et al., 1997; Silverman et al., 1999; Beck and Clewell III, 2001). The International Programme on Chemical Safety (IPCS, 1994, 2001, 2005, 2014) has subsequently adopted this TK-TD principle to derive guidance values for health-based exposure limits, but the sub-factors were refined as 4-fold (TK) and 2.5-fold (TD) for interspecies extrapolation and 3.16-fold for both TK and TD to account for intraspecies extrapolations, respectively. US EPA (2014) also described a data-derived approach on the basis of the TK-TD concept that assigns values for TK and TD differences as components within an established 10 × 10 framework for intra- and interspecies extrapolations.

Baird et al. (1996) proposed a probabilistic alternative to the traditional default UFs approach for characterizing the uncertainty in estimates of RfDs or ADIs. It relied on probabilistic characterization of uncertainties in each step of extrapolations from animal NOAELs ($UF_{H-H}$, $UF_{A-H}$, $UF_{S-C}$, $UF_{L-N}$). Individual distributions, constructed using available acute toxicity data, were assumed to be log-normally distributed, and the UFs were defined from particular percentiles (e.g., 50th and 95th) of these probability distributions. They recommended: 1) UFs of 10 and 3 (i.e., $10^{0.5}$) representing the 95th and 50th percentiles; 2) $UF_{S-C}$ and $UF_{L-N}$ for interspecies extrapolations were bounded to 1 and 50, respectively; and 3) $UF_{A-H}$ was bounded by values of 0.2 and 50. Other distributions have also been proposed (Price et al., 1997; Slob and Pieters, 1998; Swartout et al., 1998), from which these distributions were considered consistent with current use of a default UF of 10.

Cleaning products (e.g., soaps, detergents, personal care products), like other consumer products, are routinely used in large quantities across the globe. Consequently, the main chemical ingredients in these products are often considered high volume chemicals (ACI, 2010). Exposure to ingredients in cleaning products occurs either directly (e.g., dish soap) or indirectly (e.g., contact with residues when wearing laundered clothes). Thus, there remains a need to understand hazards and risks of cleaning products to public health and the environment, and to select or design alternatives when risks are unacceptable. We subsequently developed the Cleaning Product Ingredient Safety Initiative (CPISI; http://www.cleaninginstitute.org/CPISI/), through which a database consisting of > 7000 hazard data for 588 ingredients was compiled to provide hazard data for each ingredient. Such a unique mammalian toxicology database for ingredients in cleaning products provided an exceptionally unique opportunity to examine alternative approaches to identify various UFs. Similar to most other chemical classes, whether default UFs (e.g., 10) are sufficient for various chemical uses or specific chemical classes remains understudied for ingredients in cleaning products.

In the present study, we examined three primary objectives. First, we systematically examined all available acute median lethal dose (LD50), and reproductive and developmental NOAEL and LOAEL values for the rat model following oral exposures. Chemical toxicity distributions (CTDs) for these ingredients and various endpoints, and chemical category-specific CTDs were then constructed. Likelihoods of encountering acute, subacute, subchronic and chronic toxicity threshold concentrations (TCs) for specific chemical categories and all available ingredients were subsequently computed from the corresponding CTDs. Third, we identified alternatives to the traditional default UFs approach