Received: 11 November 2021 | Revised: 25 January 2022 | Accepted: 16 February 2022

DOI: 10.1002/jat.4305

RESEARCH ARTICLE

Journal of Applied Toxicology WILEY

Distributions for time, interspecies and intraspecies extrapolation for deriving occupational exposure limits

Marco Dilger[1] | Klaus Schneider[1] | Claudia Drossard[2] | Heidi Ott[2] | Eva Kaiser[1]

[1]Forschungs- und Beratungsinstitut Gefahrstoffe GmbH (FoBiG), Freiburg, Germany

[2]Federal Institute for Occupational Safety and Health, Dortmund, Germany

Correspondence
Dr. Klaus Schneider, Forschungs- und Beratungsinstitut Gefahrstoffe GmbH (FoBiG), Klarastraße 63, 79106 Freiburg, Germany.
Email: klaus.schneider@fobig.de

Funding information
Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, BAuA, Grant/Award Number: F2437; European Chemicals Agency, Helsinki

Abstract

This work aimed at improving the empirical database of time (i.e., exposure duration), interspecies and intraspecies extrapolation when deriving occupational exposure limits (OELs). For each extrapolation step, a distribution was derived, which can be used to model the associated uncertainties. For time and interspecies extrapolation, distributions of ratios of dose descriptors were derived from studies of different length or species. National Toxicology Program (NTP) study data were manually assessed, and data from REACH (Registration, Evaluation and Authorisation of Chemicals) registration dossiers were evaluated semi-automatically. Intraspecies extrapolation was investigated by compiling published studies on human toxicokinetic and toxicodynamic variability. A new database was established for toxicokinetic differences in interindividual susceptibility, including many inhalation studies. Using NTP data produced more reliable results than using REACH data. The geometric mean (GM) for time extrapolation subacute/chronic agreed with previous evaluations (GM = 4.11), whereas the GM for subchronic/chronic extrapolation was slightly higher (GM = 2.93) than the GMs found by others. No significant differences were observed between systemically and locally acting substances. Observed interspecies differences confirmed the suitability of allometric scaling, with the derived distribution describing remaining uncertainty. Distributions of intraspecies variability at the 1% and 5% incidence level had medians of 7.25 and 3.56, respectively. When compared with assessment factors (AFs) currently used in the EU, probabilities that these AFs are protective enough span a wide range from 10% to 95%, depending on the extrapolation step. These results help to select AFs in a transparent and informed way and, by allowing to compare protection levels achieved, to harmonise methods for deriving OELs.

KEYWORDS
assessment factors (AFs), distributions, interspecies extrapolation, intraspecies extrapolation, occupational exposure limits (OELs), time extrapolation, uncertainty

1 | INTRODUCTION

Health-based occupational exposure limits (OELs) are a key instrument to control airborne exposures to hazardous substances and to prevent adverse health effects at the workplace. From a methodological point of view, their derivation should ideally be based on sound human evidence, preferentially from long-term observations. Such information is lacking for newly introduced substances, but in many cases also for chemicals used for a long time. In such cases, experimental animal toxicity data are used to derive OELs. Assessment factors (AFs), which are partly based on empirical data, have been used for many years already to derive health-based exposure limits for the general population (Falk Filipsson et al., 2007; Kalberlah & Schneider, 1998; Vermeire et al., 1999; WHO, 2020). It is generally agreed that substance-specific AFs should be used whenever possible (Bhat et al., 2017). For deriving OELs, AFs for bridging data gaps gained wider acceptance only in recent years (Dankovic et al., 2015; Schenk & Johanson, 2018). They are now part of many frameworks such as the derivation of EU-wide OEL proposals by the European Chemicals Agency's (ECHA) Committee for Risk Assessment (RAC) (ECHA, 2019) or the German national OELs of the Committee on Hazardous Substances (Ausschuss für Gefahrstoffe, AGS) (AGS, 2010). Whenever possible, AFs should be based on empirical data derived from other known substances and should be accompanied with a characterisation of its uncertainty, which can best be accomplished by presenting it as probability distributions (Schneider et al., 2006; US Environmental Protection Agency [EPA], 2014; Vermeire et al., 1999).

There are differences between methodological approaches as well as numerical discrepancies between OELs derived for specific substances by different bodies. The latter is partly caused by numerical differences in AFs used and also by lacking guidance and transparency on how AFs should be used for deriving exposure limits (Deveau et al., 2015; Schenk & Johanson, 2010). This publication provides data evaluations aimed at improving the empirical basis for AFs. A major driver for differences in OELs is the uncertainty associated with interindividual differences in susceptibility, which is addressed here by creating a new empirical database on intraspecies extrapolation. Another objective of this work was to analyse (a) data provided to ECHA as part of REACH (Regulation [EC] No 1907/2006: Registration, Evaluation and Authorisation of Chemicals) registration dossiers and (b) studies from the US National Toxicology Program (NTP) to improve the data available for informing time and interspecies extrapolation.

These efforts were part of the research project F2437 ('Derivation of occupational exposure limits for airborne chemicals—Comparison of methods and protection levels') initiated by the German Federal Institute for Occupational Safety and Health (BAuA). This project analysed existing frameworks for deriving OELs and analogue values (such as Derived No Effect Levels for workers under REACH) at EU level in various regulatory sectors as well as those established at the national level in Germany. In a second publication, we describe the differences between the methods used in several frameworks and present a probabilistic analysis of the respective protection levels achieved using the distributions presented here (Schneider et al., 2022a).

2 | METHODS

2.1 | Selection and evaluation of NTP study data

Technical reports of the US NTP (available at https://ntp.niehs.nih.gov/data/tr/) that satisfied all of the following criteria were selected for in-depth analysis: (1) either inhalation or oral exposure; (2) at least two of the three exposure durations subacute (sa), subchronic (sc) or chronic (c) available; and (3) effects (as specified below) observed for at least one study type. Reports older than TR-184 (1979) were not screened as the investigation depth was considered insufficient for further analysis. In subacute NTP studies, endpoints are evaluated in less detail than in the longer studies. Consequently, the absence of reported effects does not imply the actual absence of toxicity, which could introduce bias in the subsequent analysis. Therefore, only effects on body weight were evaluated for the subacute studies. For subchronic and chronic studies, body weight and systemic effects were evaluated. In addition, local effects in the respiratory tract were evaluated for inhalation studies. For each evaluated endpoint type, the species, study type, No Observed Adverse Effect Level or Concentration (NOAEL or NOAEC) and Lowest Observed Adverse Effect Level or Concentration (LOAEL or LOAEC) were identified by consulting NTP study reports as described in detail in the project report (Schneider et al., 2022b; Report 6: Time extrapolation; section 2: Methods). This included documentation of the target organ(s) associated with the systemic LOAEL. The main goals of the evaluation criteria were to minimise subjectivity in identification of effect levels and to consider only those examinations that can be compared between two studies. Consequently, organ weights and reproductive parameters like sperm motility and oestrous cyclicity were not considered, as they are not evaluated in the chronic studies.

2.2 | Processing and evaluation of REACH study data

Study data on repeated dose toxicity for the oral and inhalation route were exported as structured data from the REACH database on 23 October 2018 by the ECHA and provided as Microsoft Excel® files. For reasons of confidentiality, it is not possible to provide these data, but we provide the ultimately calculated ratios in the depository at https://www.baua.de/EN/Tasks/Research/Research-projects/f2437.html. Studies were selected that fulfilled all of the following criteria: (1) Klimisch reliability of at least 2; (2) only experimental studies (no predictions from quantitative structure–activity relationships or read-across); (3) test material matches the registered substance; (4) appropriate study design (OECD test guideline or equivalent); (5) appropriate exposure duration*; (6) species needs to be identifiable*; (7) appropriate dose descriptor* (NO(A)EC/NO(A)EL, LO(A)EC/LO(A)EL, benchmark doses if comparable with no effect levels); and (8) appropriate dose units (daily doses normalised to body weight for oral studies, inhalation studies need to have a unit that can be converted to mg/m^{3*}). The marked (*) selection criteria are at least partially based on information in natural

language and involved data curation efforts by iteratively built string matching rules in order to assign a value that can be used for selection. Effect levels given as ranges or unbounded values were converted to the most appropriate discrete values, where possible. Details of this process and examples are given in Schneider et al. (2022a) (Report 6: Time extrapolation; section 2: Methods). Briefly, this step involved selecting the lower end of the reported dose range for a dose descriptor and equating values given as 'NO(A)EL > x' with 'NO(A)EL = x' and 'LO(A)EL < x' with 'LO(A)EL = x', while values given as 'NO(A)EL < x' and 'LO(A)EL > x' were considered inconclusive and discarded. Data curation further included extraction of the target organ from the available information in the results and discussion section, but this information was only used for data evaluation, not for selection.

Study selection and curation steps were performed entirely by scripts written in the R language (last checked with Version 4.0.3) (R Core Team, 2021), ensuring reproducibility and minimising subjectivity. The scripts can be downloaded from the depository at https://www.baua.de/EN/Tasks/Research/Research-projects/f2437.html.

2.3 | Deriving ratios of dose descriptors from study pairs

Due to the different characteristics of the two data pools, there are differences in the way ratios of dose descriptors are formed from a study pair. In case of the NTP data, identification of the two studies that form a pair was straightforward. In our evaluation, we distinguished between sex and up to three endpoint types (body weight, local and systemic effects). Each NTP report provided up to 28 NOAELs (specific for species, sex, exposure duration and endpoint type), which could be used for ratio building. However, a significant number of studies did not identify a true NOAEL, meaning that either the NOAEL was the highest tested dose or the LOAEL was at the lowest tested dose. If this was the case for one of two compared values, a ratio was calculated. The resulting ratio then represents either a minimum or a maximum and this uncertainty was accepted in return for the higher number of ratios. If both compared N(L)OAELs were facing this issue, no ratio was calculated.

The REACH data for one substance often include several relevant studies. To resolve this issue, the arithmetic mean (AM) was calculated from the dose descriptors of only those studies that match the best Klimisch reliability among all studies under consideration for the specific ratio calculation. No distinction was made between the sex of the animals or between NOAEL/NOAEC and NOEL/NOEC or LOAEL/LOAEC and LOEL/LOEC. Ratios were always calculated from two NO(A)ELs (vast majority of cases). If only two LO(A)ELs were available, these were used instead.

Ratios were always calculated in such a way that for time comparisons the value of the shorter study was divided by that of the longer one (expecting ratios > 1). For interspecies comparisons, the value of the larger species was divided by that of the smaller species (with expected values < 1, as allometric principles predict lower NOAELs, expressed as dose per kg body weight, for larger species) (Kenyon, 2012; Schneider et al., 2004).

Procedures are explained in more detail by Schneider et al. (2022b) (Report 6: Time extrapolation; section 2: Methods) and are implemented in the R scripts provided in the depository at https://www.baua.de/EN/Tasks/Research/Research-projects/f2437.html.

2.4 | Evaluation of data to model toxicokinetic (TK) and toxicodynamic (TD) uncertainty

Literature searches for data on human interindividual variability were performed until October 2019 in the database 'PubMed' (at https://pubmed.ncbi.nlm.nih.gov/). Initial searches were restricted to publications from the last 10 years, but to avoid overrepresentation of oral data, further studies were searched without time restriction (for details on the search strategy, see Schneider et al., 2022b [Report 8: Intraspecies extrapolation; section 2: Evaluation of literature data]). For TK effects, studies in adults with oral or inhalation exposure to industrial chemicals or pharmaceuticals were screened for quantitative kinetic data (area under the curve [AUC] or C_{max}). For accepting a study for evaluation, the study group had to comprise at least four individuals (representing not a highly selective subgroup of individuals) and the reported data should allow the characterisation of the variability in the study group (either individual data, or AM plus standard deviation [SD] or variation coefficient [CV] or geometric mean [GM] plus 95th confidence interval [CI] or 25th and 75th percentile). All data were assumed to be lognormally distributed, as all data were restricted on the left side (no negative values possible for TK parameters). As the variability of these data cannot be described by their SD, a $\log_{10}$ GSD value (SD of the logarithmic data) (WHO, 2014) was derived for each evaluated study using the equations provided in Table S8. For a given $\log_{10}$ GSD (i.e., for a given variability), the factor required to cover susceptibilities higher than the median can be calculated according to the following equation (WHO, 2014): Factor covering $(1-i)$ of the population $= \mathrm{GSD}^{z1-i}$, where i is the incidence (or, in other words, the percentage of the target population not covered) and z_{1-i} is the z-score of the normal distribution corresponding to this incidence. For $i = 5\%$, the corresponding value for z_{1-i} is 1.6449; for $i = 1\%$, it is 2.3263.

Accordingly, the parametrised distribution (see below) of $\log_{10}$ $\mathrm{GSD_{TK}}$ values obtained from the literature evaluation is transformed by the relationship: $\mathrm{distribution_{incidence\ I}} = 10^{\left(\mathrm{lognorm_{log_{10}GSD_{TK}}(\mu,\sigma)*z_{1-i}}\right)}$ to get a distribution of factors corresponding to the intraspecies variability for a certain accepted incidence level. From this distribution, samples are drawn for the Monte Carlo simulation. In Schneider et al. (2022b) (Report 8: Intraspecies extrapolation; annex 2: The concept of $\log_{10}$ GSD), more explanation of $\log_{10}$ GSD and the necessary transformations is provided.

For investigating variability regarding TD effects, human studies (adults only) with oral, inhalation or parenteral exposure were considered. Studies were selected for evaluation if effect data for at least two different doses/concentrations were spread wide enough to observe the range of different susceptibilities. For the quantitative evaluation, ratios are calculated by dividing the highest dose or concentration without effects in some individuals by the

lowest dose or concentration with effects. Note that these ratios based on effect doses describe both the variability due to TK and TD reasons.

A further dataset describing TD variability was published by Abdo et al. (2015). These authors provide information on the variability in the in vitro cytotoxicity of 179 chemicals in immortalised human lymphoblastoid cell lines derived from 1086 individuals representing nine different populations from five different continents ('1000 Genomes project', Coriell Institute). EC_{10} values (effective concentration, 10th percentile) were determined by Abdo et al. (2015) from eight different concentrations covering 6 orders of magnitude. Variability for each substance was described by percentiles of the obtained empirical distributions of EC_{10} values. Further, factors were calculated for each dataset describing the difference between the 1st (or 5th) percentile and the median, reflecting the difference in response of the 1% (or 5%) with the lowest EC_{10} (highest susceptibility) and the median. These 'raw factors' were corrected for sampling variability (variation between replicate measurements), which reduced variability considerably.

2.5 | Stratifications into subdistributions

The ratios of the dose descriptors were further stratified by exposure route, species, sex, endpoint type, target organ or substance type, to check for differences between ratio distributions in the strata.

The stratification by substance type used two exemplary substance classes, which were deemed to be suitable for a meaningful comparison based on their occurrence in the NTP data. These were metal compounds (excluding metals in their elemental form) and alkylated aromatics (additionally containing benzene). A list of all members of these two categories is found in Table S4. Substance classes were not extracted from the REACH data.

2.6 | Statistics and estimation of distribution parameters

Assumption of lognormality of the empirically determined distributions was verified by quantile–quantile plots. Comparison of quantiles indicated that all distributions can be adequately described by lognormal distributions. Subsequently, lognormal distributions with the parameters μ (location parameter, corresponding to the expected value on the log scale) and σ (shape parameter, corresponding to the SD on the log scale) were fitted to the data using the function 'get.lnorm.par()' from the R package 'rriskDistributions' Version 2.1.2 (Belgorodski et al., 2017). The function was called with the sorted quantiles and the corresponding probabilities ($p_i = i - 0.5/n$, where n is the number of empirical data points and $i \in \mathbb{N}^+, i \le n$). Defaults were used for all other parameters.

The distribution for interspecies extrapolation was additionally corrected for variance introduced by using NOAEL ratios, because the NOAEL carries inherent uncertainty. The evaluation by Bokkers and Slob (2007) provided data for this correction as they derived ratios based on benchmark doses as well as on NOAEL for the same study selection (column 6 of table 2 in Bokkers & Slob, 2007). σ of the distribution was corrected by

$$\sigma_{corrected} = \sqrt{(\sigma_{NOAEL,uncorrected})^2 - (\sigma_{NOAEL-error})^2},$$

for explanation, see Schneider et al. (2022b) (Report 10: Synthesis report; section 2.4: Interspecies extrapolation).

The distribution for intraspecies variability is composed of the individual distributions for TK and TD differences, combined by multiplication:

$$\text{Factors}_{intra-combined,i} = \text{Factor}_{intra-TK,i} * \text{Factor}_{intra-TD,i}.$$

The individual distributions are dependent on the chosen target incidence i, as described above. The distribution of the combined intraspecies ratios cannot be described by a lognormal distribution (or another common probability distribution) anymore. Instead, the distribution is modelled in the Monte Carlo simulation by the multiplicative combination of samples drawn from the individual distributions for TK and TD differences:

$$\text{distribution}_{intra-combined,i} = 10^{\left(\text{lognorm}_{\log_{10}\text{GSD}_{TK}}(\boldsymbol{\mu},\sigma) * z_{1-i}\right)} * \text{lognorm}_{\text{factors}_{TD},i}(\boldsymbol{\mu},\sigma)$$

Statistical evaluation of differences between distributions was performed using the bootstrap method on the GM and 75th percentile. CIs were estimated using the 'boot.ci()' from the R package 'boot' with 10,000-fold resampling and type = 'perc'. A parameter was considered different between compared distributions, if the 95% CI did not overlap.

3 | RESULTS

3.1 | Compiled datasets from NTP and REACH to evaluate time and interspecies extrapolation factors

The compiled dataset from NTP technical reports used for time and interspecies extrapolation contained studies on 256 unique substances (some of them having data for both exposure routes). Split by exposure duration, species and exposure route, a total of 1366 studies are contained in the dataset. There are roughly equal numbers of subchronic and chronic studies, and about 25% less subacute studies, which is primarily due to less frequent inclusion of this study type in older reports. Nevertheless, for most substances, the NTP data provide a complete set of data regarding different exposure durations and species (i.e., a subacute, a subchronic and a chronic study both in rats and in mice). Therefore, these data are especially valuable for calculation of dose descriptor ratios for time and interspecies extrapolation.

The REACH data started with 150,000 study records for repeated dose toxicity. After applying the selection criteria and curation steps outlined in Section 2, this resulted in 8500 dose descriptors for oral studies and 1800 dose descriptors for inhalation studies. However, the number of ratios for time and interspecies extrapolation that can be derived from the REACH data is relatively low when compared with the NTP data. This results from the fact that the second study needed

for a comparison is frequently lacking, for example, because only a subacute study is available, or all studies were performed in rats.

Even though the IUCLID entries used for the REACH dataset closely follow harmonised templates and, in principle, store the study result data as structured and machine-readable data, there is still some information to be entered as natural language. This leaves a certain amount of leeway in reporting the data, such as the description of the exposure duration and guideline followed, or usage of 'exotic' dose descriptors. We iteratively built rulesets to convert natural language into categorical variables by integrating the available information from the study records. However, the overall quality of the study entries was lower than expected, and contradictory or equivocal information was often found.

3.2 | Time extrapolation

3.2.1 | Empirical distributions obtained with NTP and REACH data

The ratios from NTP data resulted in distributions with a GM of 4.11, 1.60 and 2.93 for the extrapolations sa/c, sa/sc and sc/c, respectively (Table 1). Ratios from roughly 400 study pairs could be calculated for sa/c and sa/sc and about 1200 study pairs were available for calculation of ratios for sc/c. The difference of about a factor of 3 is because only the body weight was evaluated in the subacute studies. Therefore, no ratios could be calculated based on local or systemic effects for comparisons with this study type. The distributions obtained with the REACH data are generally based on a lower number of valid study pairs. In this case, this is not explained by the evaluated endpoints but rather an outcome of the different frequencies of subacute, subchronic and chronic studies in substance registrations (caused by the information requirements for REACH registrations). Interestingly, the GMs from the REACH data are all lower than those from NTP data: 2.49, 1.28 and 2.02 for sa/c, sa/sc and sc/c, respectively (Table 1). The corresponding GSDs however are all larger than those obtained with NTP data. In consequence, the higher percentiles are quite comparable between the two data sources. The larger GSDs for the REACH data correspond to our experience while processing the data: The quality of the REACH study entries varies considerably (see above and discussion for more details). The resulting distributions are shown in Figure 1.

TABLE 1 Summary statistics for the distributions of the ratios of dose descriptors for time extrapolation obtained from NTP and REACH study data (sa = subacute, sc = subchronic, c = chronic)

Source	Study pairs	GM	GSD	5th perc.	Median	75th perc.	95th perc.	*n*
NTP	sa/c	4.11	3.40	0.98	4.00	7.91	30.31	396
NTP	sa/sc	1.60	2.69	0.47	1.33	2.30	7.98	390
NTP	sc/c	2.93	3.04	0.50	2.67	5.00	18.94	1218
REACH	sa/c	2.49	5.22	0.21	2.95	5.44	33.50	68
REACH	sa/sc	1.28	3.48	0.20	1.03	2.40	10.00	478
REACH	sc/c	2.02	3.55	0.30	2.00	3.80	15.53	144

Abbreviations: GM, geometric mean; NTP, National Toxicology Program; REACH, Registration, Evaluation and Authorisation of Chemicals.

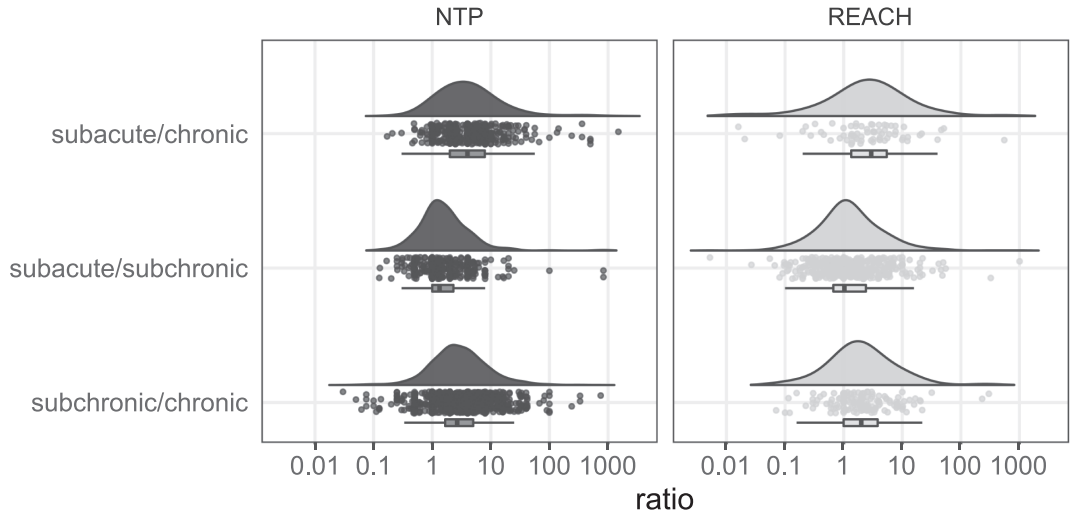

FIGURE 1 Distribution of all time comparison ratios derived from NTP and REACH data (presented as probability density function, scatter plot and box plot). NTP, National Toxicology Program; REACH, Registration, Evaluation and Authorisation of Chemicals

3.2.2 | Stratifications within the distributions for time extrapolation

The exposure route had little influence on the investigated ratio distributions for time extrapolation based on the GM and the 75th percentile (Table 2). As this may have been obscured by aggregation of systemic and local effects, the influence of the effect type was also investigated (Table 3). This stratification was only possible for the NTP data (as for the REACH data, no systematic information on the critical endpoint could be obtained) and only for the sc/c time extrapolation (as local and systemic effects were not evaluated in subacute studies).

No significant differences between ratio distributions from local, systemic or body weight effects are present in the sc/c extrapolation data from NTP inhalation studies. Similarly, no significant differences were observed between sc/c ratios derived from oral studies based on body weight or systemic effects (Table 3, Figure 2).

Other possible influencing factors on the ratio distributions are species, sex of the animals, target organ and substance type. None of these stratifications revealed relevant differences (Figure S1, Tables S1–S5), but the number of cases was small. More details on these stratifications can be found in Schneider et al. (2022b) (Report 6: Time extrapolation; section 3: Results).

3.3 | Interspecies extrapolation

3.3.1 | Empirical distributions obtained with NTP and REACH data

The vast majority of valid study pairs for ratio calculation are from the comparison of rat studies with studies on mice. Because NTP usually performed all studies on both rodent species, the number of ratios from this data source is particularly high (927 oral, 333 inhalation). In contrast, a second rodent species is generally not needed for REACH registrations, which results in the fairly low number of ratios (135 oral, 105 inhalation) (Figure 3). The REACH data also comprise some study pairs for less common species combinations, but only the comparison of oral studies in dogs versus rats gave enough ratios to draw conclusions (Table 4). The GMs for oral ratios were <1 regardless of the data source. The ratios from inhalation studies were close to 1 and the 95% CI of the GM bracketed 1 for both data sources. The oral effect

TABLE 2 Influence of the exposure route on the distributions of the ratios of dose descriptors for time extrapolation (sa = subacute, sc = subchronic, c = chronic)

Study pairs	Database	Exposure route	GM (95% CI)	GSD	75th perc. (95% CI)	*n*
sa/c	NTP	Oral	4.40 (3.85–5.06)	3.41	8.00 (6.27–8.33)	305
sa/c	NTP	Inhalation	3.25 (2.58–4.17)	3.31	6.83 (4.67–8.00)	91
sa/sc	NTP	Oral	1.65 (1.48–1.86)	2.78	2.50 (2.00–3.23)	303
sa/sc	NTP	Inhalation	1.44 (1.20–1.72)	2.38	2.00 (1.97–3.09)	87
sc/c	NTP	Oral	3.06 (2.86–3.27)	2.85	5.33 (4.76–6.00)	895
sc/c	NTP	Inhalation	2.60 (2.27–3.00)	3.54	4.01 (4.00–6.46)	323
sa/c	REACH	Oral	2.35 (1.42–3.96)	5.40	5.48 (3.63–7.67)	43
sa/c	REACH	Inhalation	2.76 (1.55–5.24)	5.08	5.02 (3.20–10.67)	25
sa/sc	REACH	Oral	1.24 (1.12–1.39)	3.10	2.12 (2.00–3.00)	379
sa/sc	REACH	Inhalation	1.45 (1.07–2.05)	5.01	2.49 (2.00–4.77)	99
sc/c	REACH	Oral	1.91 (1.52–2.38)	3.12	3.36 (2.54–5.47)	96
sc/c	REACH	Inhalation	2.28 (1.54–3.57)	4.47	4.25 (2.59–10.05)	48

Abbreviations: CI, confidence interval; GM, geometric mean; NTP, National Toxicology Program; REACH, Registration, Evaluation and Authorisation of Chemicals.

TABLE 3 Influence of the type of endpoint on the distributions of the ratios of dose descriptors for time extrapolation (subchronic/chronic) from National Toxicology Program studies

Exposure route	Endpoint type	GM (95% CI)	GSD	75th perc. (95% CI)	*n*
Oral	Body weight	2.97 (2.72–3.24)	2.52	4.82 (4.00–6.00)	447
Oral	Local effects	Not evaluated			
Oral	Systemic effects	3.17 (2.84–3.52)	3.19	6.00 (5.00–7.67)	448
Inhalation	Body weight	2.40 (1.83–3.14)	4.12	4.00 (3.00–7.98)	115
Inhalation	Local effects	2.73 (2.20–3.43)	2.96	6.25 (4.00–8.00)	101
Inhalation	Systemic effects	2.70 (2.17–3.48)	3.52	4.01 (4.00–7.50)	107

Abbreviations: CI, confidence interval; GM, geometric mean.

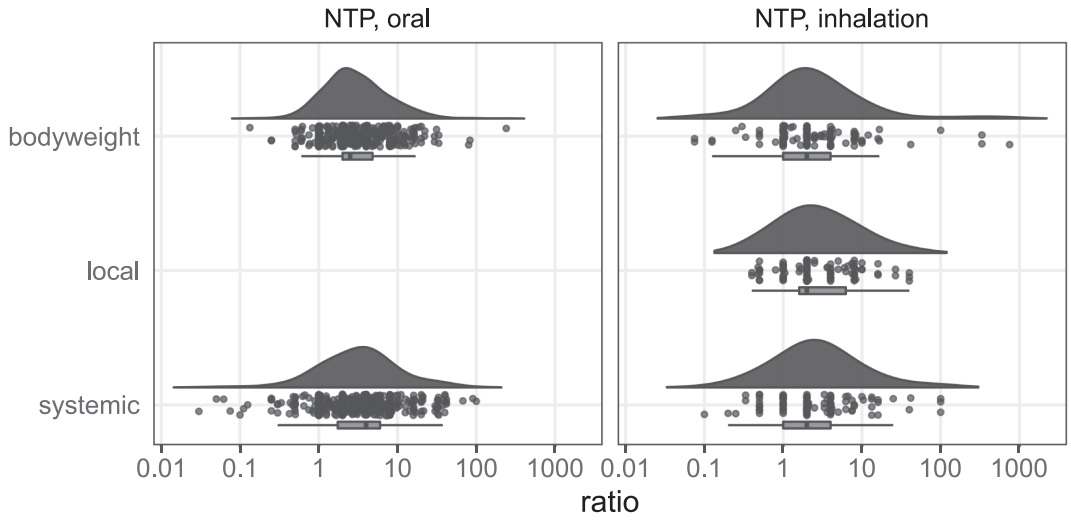

FIGURE 2 Distribution of ratios for the comparison of subchronic with chronic exposure, separated by exposure route and endpoint type. NTP, National Toxicology Program

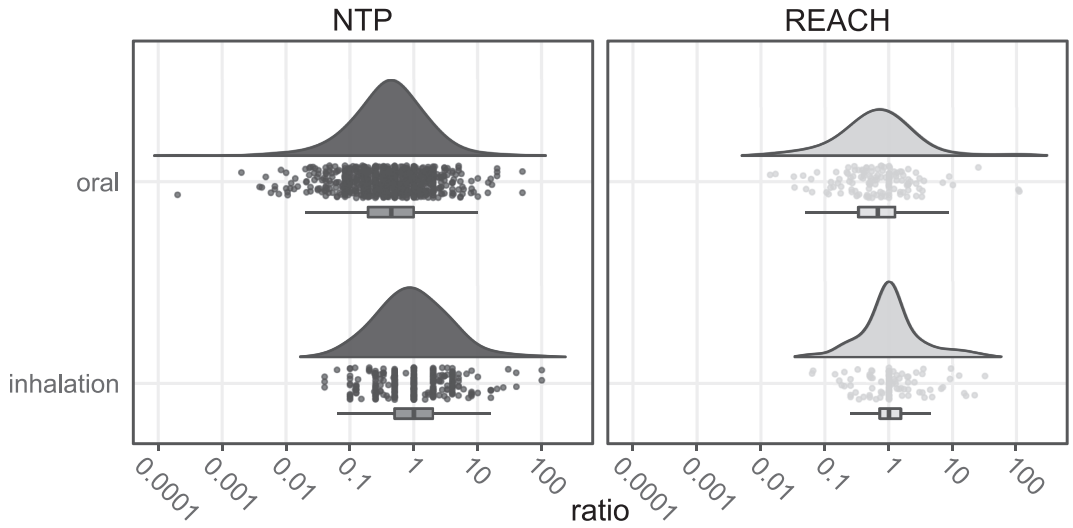

FIGURE 3 Distribution of ratios for the interspecies comparison of rats with mice obtained with NTP and REACH data, separated by exposure route. NTP, National Toxicology Program; REACH, Registration, Evaluation and Authorisation of Chemicals

TABLE 4 Summary statistics of distributions of the ratios of dose descriptors for interspecies extrapolation, separated by exposure route

Source	Study pairs	Route	GM (95% CI)	GSD	5th perc.	Median (50%)	75th perc.	95th perc.	*n*
NTP	Rat/mouse	Oral	0.40 (0.37–0.44)	3.78	0.04	0.44	1.00	2.97	927
NTP	Rat/mouse	Inhalation	0.96 (0.84–1.10)	3.61	0.12	1.00	2.00	8.00	333
REACH	Rat/mouse	Oral	0.66 (0.52–0.83)	3.85	0.06	0.67	1.27	3.39	135
REACH	Rat/mouse	Inhalation	1.09 (0.88–1.34)	2.98	0.20	1.00	1.56	9.28	105
REACH	Dog/rat	Oral	0.68 (0.48–0.98)	4.71	0.07	0.66	1.50	5.54	72
REACH	Dog/rat	Inhalation	0.82[a]	3.77	0.19	0.70	1.76	6.02	7

Abbreviations: CI, confidence interval; GM, geometric mean; NTP, National Toxicology Program; REACH, Registration, Evaluation and Authorisation of Chemicals.

[a]Not enough values to calculate meaningful confidence intervals.

levels are based on doses relative to body weight. Therefore, in agreement with allometric principles (Schneider et al., 2004), ratios < 1 for ratios from oral studies were expected. Expected values according to basal metabolic rate scaling (with an allometric exponent of 0.75) are 0.59 for rat/mouse comparisons and 0.34 in the case of dog/rat (for details, see Schneider et al., 2022b [Report 7: Interspecies extrapolation; section 2: Methods]). The GM for the rat/mouse ratios was slightly below the expected value in case of the NTP data and slightly above it for the REACH data (95% CI NTP: 0.37–0.44, REACH: 0.52–0.83). For dog/rat ratios derived from REACH data, the 95% CI for the GM ranges from 0.48 to 0.98, which is above the expected value (Table 4).

3.3.2 | Stratifications within the distributions for interspecies extrapolation

The data were further stratified to check whether exposure duration, endpoint type or target organ influenced the ratio distributions. All these stratifications were performed on the rat/mouse comparison only due to the limiting number of ratios for the other comparisons.

The exposure duration had little to no impact on interspecies differences. The interspecies differences in the oral NTP dataset from studies of varying exposure durations are very close, yet the 95% CI of the GM for chronic studies exceeds that of subacute and subchronic comparisons (Table S6). This difference is likely too small to be of practical relevance. The oral studies provided through REACH suggest that the ratios from subacute studies are higher than from longer studies, but this is due to the rather low sample number that is influenced by a few high ratios. This is also reflected in a high GSD and consequently a great overlap in the 95% CI of the GM (Table S6). The ratios from inhalation studies show no differences between exposure durations, and the NTP and REACH data are in good agreement.

The stratification according to types of endpoints showed only insignificant differences for ratios from oral studies. In the case of inhalation studies, the GM for ratios derived from effects on body weight is slightly shifted towards greater ratio values (1.26, i.e., a higher sensitivity of mice), whereas for local and systemic effects, the GM was <1.0 (Table S7). Because all three 95% CIs for inhalation studies comprise the ratio of 1 or are very close to 1 and the differences in the distribution are rather small, this is likely not of further relevance.

Target organ had no influence on the distributions, but the relatively low number of ratios made it difficult to detect any target organ-specific differences (Table S8). The stratification by substance type faced the same difficulty, yet here a difference was observed in the ratios from inhalation studies. It appears that alkylated aromatics have a higher GM (1.57, corresponding to mice being more sensitive) whereas metal compounds have a lower GM (0.56, rats more sensitive). The 95% CI for the GM suggests that rats may indeed react more sensitive towards metal compounds after inhalation exposure than mice (Table S9).

3.4 | Intraspecies extrapolation

3.4.1 | Differences in toxicokinetics

We identified 74 human studies in the data searches and evaluated them in detail (see Table S10). In about half of the studies, the route of exposure was inhalation, which is the most relevant pathway for occupational exposures. From 68 studies, we calculated $\log_{10}$ GSD values, characterising the interindividual variability in pharmacokinetic/TK parameters observed in each individual dataset. Six studies were dismissed, mainly because only ranges (maximum, minimum) were reported.

Inhalation studies showed a slightly lower variability than oral studies (GM for $\log_{10}$ GSD values from inhalation studies: 0.111 vs. GM for oral studies 0.168) (Table 5). Differences were significant between GMs (95% CI inhalation: 0.091–0.135, oral: 0.138–0.202), but not the 75th percentiles of the empirical distributions (95% CI inhalation: 0.127–0.223, oral: 0.175–0.351). Other stratifications (by substance class or health status of study participants) did not reveal significant differences (see Tables S13 and S14).

The data in Table 5 point to a smaller variability for the inhalation route. However, the opposite was observed when evaluating studies that involve both TK and TD differences (see below). Therefore, the full dataset (all routes combined) is used for describing intraspecies variability due to TK reasons.

3.4.2 | Differences in toxicodynamics

Variability observed in measured endpoints at a given concentration or dose cannot be linearly translated into variability in equipotent doses. Therefore, studies applying a large range of concentrations or doses allowing to identify interindividual differences in equipotent doses or concentrations were searched but, expectedly, proved to be scarce. We only identified 24 studies (12 inhalation studies, 4 with oral and 8 with other exposure routes; see Tables 6 and S11), in which similar effects were observed in individuals or groups at varying

TABLE 5 Summary statistics of the empirical distribution of the $\log_{10}$ GSD values from the evaluated pharmacokinetic/toxicokinetic studies

Exposure route	GM	GSD	5th perc.	Median	75th perc.	95th perc.	*n*
All	0.141	1.830	0.049	0.146	0.220	0.355	68
Oral	0.168	1.764	0.058	0.167	0.264	0.379	33
Inhalation	0.111	1.804	0.042	0.106	0.179	0.252	31
Other	0.218	1.305	0.174	0.205	0.236	0.300	4

Abbreviation: GM, geometric mean.

concentrations or doses. For these studies, we calculated the ratios of equipotent doses or concentrations between highly susceptible individuals and those with low susceptibility. These ratios varied over a broad range from 3 to 201, which include in most cases variability due to both TK and TD reasons, as external doses or concentrations were compared. Variability was especially high in inhalation studies, but differences due to exposure routes (or other parameters) are difficult to judge upon given the small number of studies.

In light of these uncertainties, we chose an alternative approach to characterise TD variability, based on the data from Abdo et al. (2015) as described in Section 2. With the data reported by Abdo et al. (2015), we calculated for each chemical the ratio median/5th percentile and the ratio median/1st percentile for the reported EC_{10} values. The statistical parameters of the distributions representing a protection goal of 95% and 99% of the target population, respectively, are given in Table 7.

The data in Table 7 can be interpreted as follows: To cover TD variability in 95% of the adult population with a probability of 50%, a factor of 1.95 is required. For covering 99% of the population with a probability of 95%, the extrapolation factor would be 10.32. These empirical distributions can be taken to derive parametric distributions for TD variability to cover 95% or 99% of the adult population.

3.5 | Distributions to model time, interspecies and intraspecies extrapolation

The stratifications of the distributions for time and interspecies extrapolation did not identify a study parameter that was influencing the empirical ratio distributions with sufficient magnitude or certainty to warrant the derivation of separate distribution parameters specific for that study parameter. The experiences made with the REACH data, both during preparation and curation as well as with the resulting ratio (high number of time extrapolation ratios < 1, generally wider distributions), led us to use only the NTP data for deriving parametrised distributions for time and interspecies extrapolation.

Our evaluation of interspecies differences confirmed the applicability of allometric principles for interspecies extrapolation. For describing the remaining variability, all ratios obtained from the evaluation of NTP data were normalised to a GM of 1, with the scatter around 1 describing the substance-to-substance variability.

The empirical distributions for time and interspecies extrapolation could be adequately described by lognormal distributions with μ (location parameter, corresponding to the expected value on the log scale) and σ (shape parameter, corresponding to the SD on the log scale) (Table 8). The parameters for the interspecies distribution in Table 8

TABLE 6 Summary statistics of the distribution of the ratios of dose descriptors from all evaluated studies for pharmacodynamic/toxicodynamic effects

Exposure route	GM	GSD	5th perc.	Median	75th perc.	95th perc.	*n*
All (inhalation [$n = 12$], oral [$n = 4$], other [$n = 8$])	7.39	2.48	3.00	6.00	8.00	33.00	24

Abbreviation: GM, geometric mean.

TABLE 7 Statistical characterisation of the ratios of dose descriptors describing toxicodynamic variability reported by Abdo et al. (2015)

Distribution	Median	5th perc.	95th perc.
Median/5th percentile ratios of EC_{10} values from substance-specific datasets[a] (5% incidence; 95% of population covered)	1.95	1.19	4.67
Median/1st percentile ratios of EC_{10} values from substance-specific datasets[a] (1% incidence; 99% of population covered)	3.04	1.44	10.32

[a]Own calculation from substance-specific data provided in supplemental material of Abdo et al. (2015).

TABLE 8 Parameters of derived distributions of assessment factors (μ: location parameter; σ: shape parameter)

Extrapolation	Data source	μ	σ	Median	75th perc.	95th perc.
Time: sa/c	NTP	1.31	1.05	3.71	7.52	20.85
Time: sc/c	NTP	1.04	0.99	2.83	5.53	14.49
Interspecies	NTP	0.02	0.75	1.02	1.69	3.49
Combined (TK and TD) intraspecies	TK and TD distributions for 1% incidence, see text for explanation	See text		7.25	12.53	34.26
	TK and TD distributions for 5% incidence	See text		3.56	5.15	10.37

Abbreviations: NTP, National Toxicology Program; sa/c, subacute/chronic; sc/c, subchronic/chronic; TD, toxicodynamic; TK, toxicokinetic.

are already corrected for the inflated variance as described in Section 2.

The combined (TK plus TD) intraspecies distributions were derived for the two protection objectives with coverage of 95% or 99% (or 5% and 1% remaining incidence, respectively) of the target population (adult workers) by combining the respective lognormal distributions for TK and TD as given by the formula in Section 2 using the Monte Carlo simulations. The parameters for the distribution of $\log_{10}$ GSD_{TK} were $\mu = 1.93$ and $\sigma = 0.61$. For the distribution of ratio-s_{TD}, lognormal distributions with $\mu = 1.11$ and $\sigma = 0.58$ (1% remaining incidence) and $\mu = 0.66$ and $\sigma = 0.37$ (5% incidence) were derived from the data by Abdo et al. (2015) as given in Table 7. The resulting distributions (TK and TD combined) are shown in Table 8.

3.6 | Comparison with currently used default values

Having concluded on appropriate distributions to model the uncertainties of the extrapolation steps involved in risk assessment, it is possible to compare the currently used default values for AFs with these distributions. The cumulative probability of the distribution at a certain value corresponds to the probability that this AF is adequately covering the uncertainty of this extrapolation step for a new substance or evaluation. In this project, we evaluated the following frameworks for deriving OELs or analogue values: EU REACH (ECHA, 2012), the EU Plant Protection Products Directive (EU PPPD) (European Commission [EC], 2006; European Food Safety Authority [EFSA], 2012), the EU Biocidal Products Directive (EU BPD) (ECHA, 2017), OELs set by ECHA's RAC (ECHA, 2019), the German Committees AGS (AGS, 2010, 2018) and MAK Commission (Deutsche Forschungsgemeinschaft [DFG], 2019), and European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC, 2010) (for more details, see Schneider et al., 2022a).

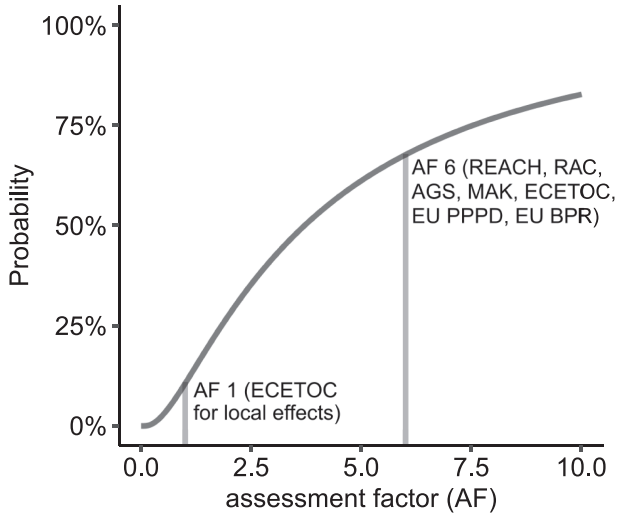

FIGURE 4 Cumulative probability distribution for subacute to chronic extrapolation and probabilities achieved by currently used assessment factors (vertical lines). AGS, Ausschuss für Gefahrstoffe; ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals; EU BPR, EU Biocidal Products Regulation; EU PPPD, EU Plant Protection Products Directive; RAC, Committee for Risk Assessment; REACH, Registration, Evaluation and Authorisation of Chemicals

Figure 4 shows an example for such a comparison. The common default value for the AF for subacute to chronic extrapolation is 6. This default covers about 68% of cases in our analysis, corresponding to a probability of 68% that the point of departure (POD) from the subacute study divided by 6 provides the same protection from adverse effects that the chronic POD would provide. Table 9 provides the achieved probabilities for the various extrapolation steps.

TABLE 9 Probabilities achieved by currently used default values for assessment factors

Extrapolation	Assessment factor (regulatory framework)	Probability
Time: sc/c	2 (REACH, RAC, AGS, MAK, ECETOC, EU PPPD, EU BPR)	36.3%
	1 (discussed by ECETOC for local effects)	14.6%
Time: sa/c	6 (REACH, RAC, AGS, MAK, ECETOC, EU PPPD, EU BPR)	67.7%
	1 (discussed by ECETOC for local effects)	10.6%
Interspecies	2.5 (REACH, RAC, EU PPPD[a], EU BPR[a])	88.4%
	1 (ECETOC)	48.6%
Intraspecies	10 (EU PPPD, EU BPR)	65.8% (1% incidence)
		94.6% (5% incidence)
	5 (REACH, RAC)	30.7% (1% incidence)
		73.4% (5% incidence)
	3 (ECETOC)	10.8% (1% incidence)
		36.6% (5% incidence)
Interspecies and intraspecies combined[b]	5 (AGS)	34.5% (1% incidence)
		61.9% (5% incidence)
	2 (MAK)	10.1% (1% incidence)
		24.3% (5% incidence)

AGS, Ausschuss für Gefahrstoffe; ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals; EU BPR, EU Biocidal Products Regulation; EU PPPD, EU Plant Protection Products Directive; RAC, Committee for Risk Assessment; REACH, Registration, Evaluation and Authorisation of Chemicals; sa/c, subacute/chronic; sc/c, subchronic/chronic.

[a]EU PPPD and BPR recommend a default of 10 for interspecies extrapolation (without a factor for allometric scaling). Hence, it is identical to the default of 2.5 used by REACH and RAC in case of extrapolation from rat data, for which the allometric scaling factor is 4.

[b]Combined factors of interspecies and intraspecies extrapolation are used in the German frameworks; probabilities were calculated by combining the two distributions by the Monte Carlo simulation.

4 | DISCUSSION

This study aimed at improving the empirical database of time, interspecies and intraspecies extrapolation when deriving OELs. For each extrapolation step, a distribution was derived. In a further step, we used the established distributions to evaluate the differences between various OEL frameworks and their levels of protection achieved.

We used two different approaches for time and interspecies extrapolation according to the nature of the data. A completely manual assessment in case of the unstructured, but high-quality, NTP data and a largely automated data analysis was based on the mostly computer-readable REACH data. The resulting distributions agree in the basic characteristics, but the distributions from REACH data, in general, have a lower GM but a larger spread. In the data analysis, we experienced obstacles in getting quality results from the REACH data, which generally could be overcome by iteratively customising the analysis pipeline. Yet, limitations due to the inconsistent reporting quality of the REACH data persist, as evidenced, for example, by the high frequency of implausible ratios for time extrapolation, which are less than 1 (i.e., effects in the shorter study are reported at lower concentrations). This finding agrees with the ones observed by others (Lampe et al., 2018; Luechtefeld et al., 2016; Zarn et al., 2011). Many of these cases could have been resolved by a manual assessment, but this was not possible due to the large amount of data. Researchers, who use REACH data directly or indirectly via aggregators like eChemPortal, EPAs CompTox Dashboard or the OECD QSAR Toolbox, should be wary of such limitations and consider appropriate quality controls. Considering all aspects that contribute to the quality of the resulting distributions, we consider the ratios derived from NTP data a better representation of the reality and therefore use the NTP data to derive parametrised distributions for comparison with default AFs. Nevertheless, the REACH data are of high value to train predictive models, although Luechtefeld et al. (2016) also acknowledged limitations in the data for time extrapolation of repeated dose toxicity (only sa/sc was investigated).

The GMs for subacute to chronic ratios, which we obtained with the NTP data, are close to the results published by others (Figure 5). The variability (expressed as dimensionless GSD) is slightly lower compared with other evaluations, which might be due to the strict criteria for comparing related endpoint types.

For substances acting locally in the respiratory tract, ECETOC (2010) argued that these effects are concentration driven and not depending on exposure time and, hence, proposed an AF of 1 for differences in exposure duration. This assumption is not confirmed by our analysis. Also, Mangelsdorf et al. (2021) found even higher ratios for local effects (which included effects on the eye and the respiratory tract in their evaluation) compared with systemic toxicity.

For subchronic to chronic extrapolation, our GMs from the NTP studies (3.1 for oral, 2.6 for inhalation data) are at the upper end of values reported by others (Figure 5). This might partly be due to our rigorous exclusion of endpoints in the subchronic NTP studies, which are not reported in the chronic studies (e.g., organ weights). Their inclusion would have tended to lower the NOAELs and LOAELs in the subchronic studies and would have reduced the ratios when compared with the respective chronic values. As the NTP studies form part of many of the existing evaluations (due to the scarcity of chronic studies), this methodological difference might have a relevant impact. It is noteworthy that the multiplication product of our GMs for sa/sc and sc/c (see Table 1) shows a high accordance with the GM for sa/c, indicating consistency in the whole dataset.

The comparison between species performed with the NTP data revealed a distinctive difference in the ratios for inhalation and oral data, which agrees with the principles of allometric scaling (Table 4). The accordance with the predictions by basal metabolic rate scaling is high, considering that only data from rats and mice were used, two rather small species. Such an agreement has been shown by many authors before (Bokkers & Slob, 2007; Escher et al., 2013; Kratchman et al., 2018;Price et al., 2008; Schneider et al., 2004). Although the information obtained from the comparison of two smaller species is limited and does not provide immediate conclusions on animal–human interspecies differences, it supports the principle of allometric scaling and therefore is in agreement with analyses comparing effect levels between animal species and humans (Price et al., 2008; Schneider et al., 2004).

The distribution shown in Figure 3 reflects the remaining variation around a GM of 1 after correction for the additional variability introduced by the measurement error inherent to NOAEL/LOAEL values. This variability is modest in size and, as can be seen in Table 9, the coverage achieved by the often-used AF of 2.5 is high. Application of this distribution, which is based on the evaluation of NTP reports for 256 substances, in combination with basal metabolic rate scaling provides a robust, data-derived procedure in case substance- and effect-specific information on species differences is not available. The range of variability is narrower in this large dataset compared with the ones observed earlier in smaller datasets with short-term exposure conditions, which compared TK parameters and subacute toxicity between experimental animals and humans (Schneider et al., 2004). Consequently, the use of distributions derived from these evaluations would result in more conservative OEL estimates.

Compared with mice, rats appeared to be less sensitive to alkylated aromatic compounds in NTP inhalation studies. On the other side, rats seem to be more sensitive towards metal compounds. It is known that inflammatory responses in the lungs after particle exposure can be more pronounced in rats compared with mice (Carter et al., 2006; Elder et al., 2005). Although this might be an explanation for the difference observed with metal compounds, it is unlikely that it had a large impact on the total dataset (due to the limited number of substances concerned), which can also be deduced from the GM close to 1 (i.e., 0.96) for all NTP inhalation studies combined. However, it is an important aspect to consider when deriving OELs for particulate substances based on rat inhalation studies. For adjusting for species differences regarding particle dosimetry and retention in the respiratory tract, models such as the 'Multi-Path Particle Dosimetry' (MPPD) model can be used to calculate a 'human equivalent concentration'. Advantages and limitations of this approach are discussed in

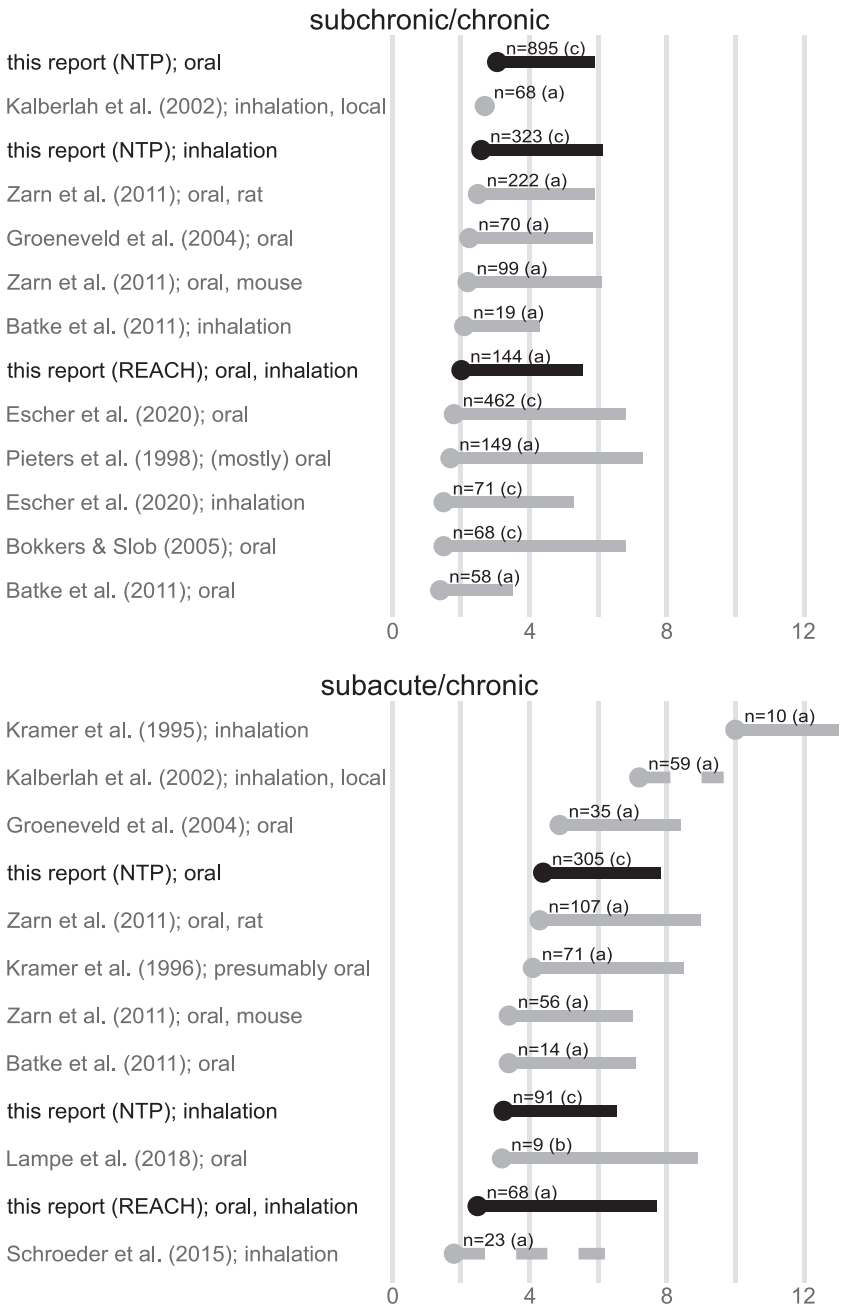

FIGURE 5 Geometric means (dot) and GSD (horizontal lines; dashed, if our estimates) for subchronic/chronic (top panel) and subacute/chronic (bottom) time extrapolation; comparison of results from this analysis (in black) with published data (for references see Schneider et al., 2022b; Report 6: Time extrapolation). (a): one ratio per substance, (b): ratios derived from multiple studies per substance, (c): multiple ratios derived from two studies per substance. NTP, National Toxicology Program; REACH, Registration, Evaluation and Authorisation of Chemicals

detail by Schneider et al. (2022b) (Report 9: Human Equivalent Concentration and Kinetic Modelling of Aerosols in the Lower Respiratory Tract).

Interindividual variability caused by differences in TKs was subject to various investigations. In an attempt to develop AFs specific for certain metabolism pathways, Renwick, Dorne and colleagues evaluated human studies published between 1966 and 2003 (Dorne et al., 2001a, 2001b, 2004, 2005; Dorne et al., 2002; Renwick et al., 2001). These authors noted that for pathways involving polymorphically expressed xenobiotic metabolising enzymes (e.g., CYP2C19 and CYP2D6), TK variability often exceeded a factor of 3. Hattis and colleagues developed a database documenting interindividual variability in TKs and TDs (Hattis, 1996a, 1996b; Hattis et al., 2002; Hattis, Banati, & Goble, 1999; Hattis, Banati, Goble, & Burmaster, 1999; Hattis & Lynch, 2007; Hattis & Silver, 1994). The data were used to describe intraspecies variability in a probabilistic assessment framework developed by WHO (2014). The data evaluated by both groups mostly consist of studies on pharmaceuticals with oral administration. Other data for differences due to TDs are extremely scarce.

Our new data compilation fills important data gaps with relevance for the derivation of OELs. Half of the TK studies we used are inhalation studies with industrial chemicals. For the TD part, we used the in vitro data presented by Abdo et al. (2015), which avoids several limitations of in vivo studies: The differences observed can be attributed to TD variability alone, and by repeating the experiments, the data

could be corrected for additional variability caused by measurement uncertainty. With cells from more than 1000 individuals, they represent a much broader range of the population than the available in vivo studies. However, some uncertainty arises from the unknown representativity of these immortalised human cell lines for cells in living human individuals. Most importantly, it is unclear whether the inter-individual variability measured for the endpoint cytotoxicity is adequate for effects on highly specialised cells such as immune cells. The observed variability should therefore be considered to represent the lower end of TD variabilities in humans. Although it is somewhat lower than that described by Hattis and colleagues and as used in WHO (2014) based on human in vivo data, Abdo et al. (2015) found the two datasets being largely consistent. A detailed comparison and discussion of the two datasets can be found in Schneider et al. (2022b) (Report 8: Intraspecies extrapolation; section 4.2: Variability in toxicodynamics).

With a median of 3.6 and a 95th percentile of 10.4 at the 5% incidence level, our distribution, derived from adult data for the target population of workers, describes a slightly lower variability compared with the distribution obtained by WHO (2014) (median 5.0, 95th perc. 14.0). Although the WHO distribution is meant to be applied for health-based exposure limits for the general population, inclusion of data on vulnerable groups (e.g., children) does not seem to substantially increase the variability. Rather, considering the results obtained by Renwick, Dorne and colleagues as discussed above (Dorne et al., 2001a, 2001b, 2004, 2005; Dorne et al., 2002; Renwick et al., 2001), polymorphisms in xenobiotic metabolising enzymes are likely drivers for large differences in susceptibility for some substances. These polymorphisms are relevant for both workers and the general population. In light of our distribution for intraspecies variability, currently used default values for AFs are showing a low probability of providing sufficient protection (Table 9).

In conclusion, this study compiled new data from various sources and now provides an improved empirical database to set data-derived values for AFs in the context of OEL derivation from toxicological studies. Using these data, our analysis shows that probabilities that default values of AFs currently used in the different OEL frameworks in the EU are protective enough span a wide range from 10% to 95%, depending on the extrapolation step. Combining the distributions with probabilistic methods in full assessments confirmed that considerable differences exist between existing methodologies for deriving OELs (Schneider et al., 2022a). Such evaluations increase the transparency of OELs and thus contribute to harmonising the approaches.

ACKNOWLEDGEMENTS

We are grateful to the European Chemicals Agency, Helsinki, for compiling and providing the REACH data for our analysis.

We would like to thank Dr. Werner Wosniok, University of Bremen, for his support in statistics and Prof. Dr. Thomas Gebel, BAuA, for the valuable comments provided throughout the project.

The project 'Derivation of occupational exposure limits for airborne chemicals—Comparison of methods and protection levels' was funded by the German Federal Institute for Occupational Safety and Health (Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, BAuA) (project number F2437).

CONFLICT OF INTEREST

The authors did not report any conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available at https://www.baua.de/EN/Tasks/Research/Research-projects/f2437.html.

ORCID

Klaus Schneider https://orcid.org/0000-0002-7752-7578

REFERENCES

Abdo, N., Xia, M., Brown, C. C., Kosyk, O., Huang, R., Sakamuru, S., Zhou, Y. H., Jack, J. R., Gallins, P., Xia, K., Li, Y., Chiu, W. A., Motsinger-Reif, A. A., Austin, C. P., Tice, R. R., Rusyn, I., & Wright, F. A. (2015). Population-based in vitro hazard and concentration-response assessment of chemicals: The 1000 genomes high-throughput screening study. *Environmental Health Perspectives*, *123*, 458–466. https://doi.org/10.1289/ehp.1408775

AGS, Ausschuss für Gefahrstoffe. (2010). Bekanntmachung zu Gefahrstoffen. Kriterien zur Ableitung von Arbeitsplatzgrenzwerten. BekGS 901. *Das Gemeinsame Ministerialblatt*, *61*(32), 691–696. https://www.baua.de/DE/Angebote/Rechtstexte-und-Technische-Regeln/Regelwerk/TRGS/Bekanntmachung-901.html

AGS, Ausschuss für Gefahrstoffe. (2018). Technische Regeln für Gefahrstoffe – Arbeitsplatzgrenzwerte (TRGS 900). Ausgabe: Januar 2006. BArBl Heft 1/2006 S. 41–55. Geändert und ergänzt: GMBl 2018 S.542–545[Nr.28] (v.07.06.2018). Bundesanstalt für Arbeitsschutz und Arbeitsmedizin. https://www.baua.de/DE/Angebote/Rechtstexte-und-Technische-Regeln/Regelwerk/TRGS/pdf/TRGS-900.pdf?__blob=publicationFile&v=11

Belgorodski, N., Greiner, M., Tolksdorf, K., Schueller, K., Flor, M. & Göhring, L. (2017). Package 'rriskDistributions'. Online: https://cran.r-project.org/web/packages/rriskDistributions/rriskDistributions.pdf (08.11. 2018).

Bhat, V. S., Meek, M. E., Valcke, M., English, C., Boobis, A., & Brown, R. (2017). Evolution of chemical-specific adjustment factors (CSAF) based on recent international experience; increasing utility and facilitating regulatory acceptance. *Critical Reviews in Toxicology*, *47*, 729–749. https://doi.org/10.1080/10408444.2017.1303818

Bokkers, B. G. H., & Slob, W. (2007). Deriving a data-based interspecies assessment factor using the NOAEL and the benchmark dose approach. *Critical Reviews in Toxicology*, *37*(5), 355–373. https://doi.org/10.1080/10408440701249224

Carter, J. M., Corson, N., Driscoll, K. E., Elder, A., Finkelstein, J. N., Harkema, J. N., Gelein, R., Wade-Mercer, P., Nguyen, K., & Oberdorster, G. (2006). A comparative dose-related response of several key pro- and antiinflammatory mediators in the lungs of rats, mice, and hamsters after subchronic inhalation of carbon black. *Journal of Occupational and Environmental Medicine*, *48*, 1265–1278. https://doi.org/10.1097/01.jom.0000230489.06025.14

Dankovic, D. A., Naumann, B. D., Maier, A., Dourson, M. L., & Levy, L. S. (2015). The scientific basis of uncertainty factors used in setting occupational exposure limits. *Journal of Occupational and Environmental Hygiene*, *12*(Suppl. 1), S55–S68. https://doi.org/10.1080/15459624.2015.1060325

Deveau, M., Chen, C. P., Johanson, G., Krewski, D., Maier, A., Niven, K. J., Ripple, S., Schulte, P. A., Silk, J., Urbanus, J. H., Zalk, D. M., &

Niemeier, R. W. (2015). The global landscape of occupational exposure limits—Implementation of harmonization principles to guide limit selection. *Journal of Occupational and Environmental Hygiene*, *12*(Suppl. 1), S127–S144. https://doi.org/10.1080/15459624.2015.1060327

DFG, Deutsche Forschungsgemeinschaft. (2019). *MAK- und BAT-Werte-Liste 2019*. *Senatskommission zur Prüfung gesundheitsschädlicher Arbeitsstoffe*. *Mitteilung 55*. WILEY-VCH Verlag GmbH. https://doi.org/10.1002/9783527826155

Dorne, J. L. C. M., Walton, K., & Renwick, A. G. (2001a). Human variability in glucuronidation in relation to uncertainty factors for risk assessment. *Food and Chemical Toxicology*, *39*, 1153–1173. https://doi.org/10.1016/S0278-6915(01)00087-4

Dorne, J. L. C. M., Walton, K., & Renwick, A. G. (2001b). Uncertainty factors for chemical risk assessment: Human variability in the in vivo pharmacokinetics of CYP1A2 substrates. *Food and Chemical Toxicology*, *39*, 681–696. https://doi.org/10.1016/S0278-6915(01)00005-9

Dorne, J. L. C. M., Walton, K., & Renwick, A. G. (2004). Human variability in the renal elimination of foreign compounds and renal excretion-related uncertainty factors for risk assessment. *Food and Chemical Toxicology*, *42*(2), 275–298. https://doi.org/10.1016/j.fct.2003.09.002

Dorne, J. L. C. M., Walton, K., & Renwick, A. G. (2005). Human variability in xenobiotic metabolism and pathway-related uncertainty factors for chemical risk assessment: A review. *Food and Chemical Toxicology*, *43*(2), 203–216. https://doi.org/10.1016/j.fct.2004.05.011

Dorne, J. L. C. M., Walton, K., Slob, W., & Renwick, A. G. (2002). Human variability in polymorphic CYP2D6 metabolism: Is the kinetic default uncertainty factor adequate? *Food and Chemical Toxicology*, *40*, 1633–1656. https://doi.org/10.1016/S0278-6915(02)00117-5

EC, European Commission. (2006). Guidance for the setting and application of acceptable operator exposure levels (AOELs). Draft. SANCO 7531—rev.10. Directorate E—Safety of the food chain. E3—Chemicals, contaminants, pesticides.

ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals. (2010). Technical Report No. 110. Guidance on assessment factors to derive a DNEL. Brussels, Belgium.

ECHA, European Chemicals Agency. (2012). Guidance on information requirements and chemical safety assessment. Chapter R.8: Characterisation of dose [concentration]-response for human health. Version 2.1. November 2012. Helsinki, Finland. https://echa.europa.eu/guidance-documents/guidance-on-information-requirements-and-chemical-safety-assessment

ECHA, European Chemicals Agency. (2017). Guidance on the biocidal products regulation. Volume III Human health—Assessment & evaluation. (Parts B+C). Version 4.0. December 2017. https://echa.europa.eu/documents/10162/23036412/biocides_guidance_human_health_ra_iii_part_bc_en.pdf/30d53d7d-9723-7db4-357a-ca68739f5094

ECHA, European Chemicals Agency. (2019). Guidance on information requirements and chemical safety assessment. Appendix to Chapter R.8: Guidance for preparing a scientific report for health-based exposure limits at the workplace. Version 1.0. August 2019. Helsinki, Finland. https://echa.europa.eu/guidance-documents/guidance-on-information-requirements-and-chemical-safety-assessment

EFSA, European Food Safety Authority. (2012). Guidance on selected default values to be used by the EFSA Scientific Committee, Scientific Panels and Units in the absence of actual measured data. *EFSA Journal*, *10*(3). https://doi.org/10.2903/j.efsa.2012.2579

Elder, A., Gelein, R., Finkelstein, J. N., Driscoll, K. E., Harkema, J., & Oberdorster, G. (2005). Effects of subchronically inhaled carbon black in three species. I. Retention kinetics, lung inflammation, and histopathology. *Toxicological Sciences*, *88*, 614–629. https://doi.org/10.1093/toxsci/kfi327

Escher, S. E., Batke, M., Hoffmann-Doerr, S., Messinger, H., & Mangelsdorf, I. (2013). Interspecies extrapolation based on the RepDose database—A probabilistic approach. *Toxicology Letters*, *218*, 159–165. https://doi.org/10.1016/j.toxlet.2013.01.027

Falk Filipsson, A., Hanberg, A., Victorin, K., Warholm, M., & Wallén, M. (2007). Assessment factors—Applications in health risk assessment of chemicals. *Environmental Research*, *104*, 108–127. https://doi.org/10.1016/j.envres.2006.10.004

Hattis, D. (1996a). Human interindividual variability in susceptibility to toxic effects: From annoying detail to a central determinant of risk. *Toxicology*, *111*, 5–14. https://doi.org/10.1016/0300-483X(96)03388-4

Hattis, D. (1996b). Variability in susceptibility—How big, how often, for what responses to what agents? *Environmental Toxicology and Pharmacology*, *2*, 135–145. https://doi.org/10.1016/S1382-6689(96)00044-0

Hattis, D., Baird, S., & Goble, R. (2002). A straw man proposal for a quantitative definition of the RfD. In Final Technical Report, U.S. Environmental Protection Agency STAR grant # R825360, Human Variability in Parameters Potentially Related to Susceptibility for Non-cancer Risks. Full version available at http://www2.clarku.edu/faculty/dhattis; shortened version in Drug and Chemical Toxicology 2002; 25: 403–436.

Hattis, D., Banati, P., & Goble, R. (1999). Distributions of individual susceptibility among humans for toxic effects. *Annals of the New York Academy of Sciences*, *895*, 286–316. https://doi.org/10.1111/j.1749-6632.1999.tb08092.x

Hattis, D., Banati, P., Goble, R., & Burmaster, D. E. (1999). Human interindividual variability in parameters related to health risks. *Risk Analysis*, *19*, 711–726. https://doi.org/10.1111/j.1539-6924.1999.tb00441.x

Hattis, D., & Lynch, M. K. (2007). Empirically observed distributions of pharmacokinetic and pharmacodynamic variability in humans implications for the derivation of single point component uncertainty factors providing equivalent protection as existing RfDs. In J. C. Lipscomb & E. V. Ohanian (Eds.), *Toxicokinetics and risk assessment* (1st ed.). CRC Press.

Hattis, D., & Silver, K. (1994). Human interindividual variability-A major source of uncertainty in assessing risks for noncancer health effects. *Risk Analysis*, *14*(4), 421–431. https://doi.org/10.1111/j.1539-6924.1994.tb00260.x

Kalberlah, F. & Schneider, K. (1998). Quantification of extrapolation factors. Final Report of the Research Project No. 116 06 113 of the Federal Environmental Agency. Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin Dortmund Fb 797 Wirtschaftsverlag NW Bremerhaven.

Kenyon, E. M. (2012). Interspecies extrapolation. In B. Reisfeld & A. Mayeno (Eds.), *Computational toxicology. Methods in Molecular Biology (Methods and Protocols)* (Vol. 929, pp. 501–520). Humana Press.

Kratchman, J., Wang, B., & Gray, G. (2018). Which is most sensitive? Assessing responses of mice and rats in toxicity bioassays. *Journal of Toxicology and Environmental Health. Part A*, *81*, 173–183. https://doi.org/10.1080/15287394.2018.1423799

Lampe, B. J., Fuller, E., & Kuppusamy, S. P. (2018). A quantitative comparison of points of departure between 28-day and 90-day repeated dose studies with a proposed extrapolation factor. *Regulatory Toxicology and Pharmacology*, *92*, 189–200. https://doi.org/10.1016/j.yrtph.2017.12.007

Luechtefeld, T., Maertens, A., Russo, D. P., Rovida, C., Zhu, H., & Hartung, T. (2016). Analysis of public oral toxicity data from REACH registrations 2008–2014. *ALTEX - Alternativen zu Tierexperimenten*, *33*, 111–122. https://doi.org/10.14573/altex.1510054

Mangelsdorf, I., Schröder, K., Escher, S. E., Kolossa-Gehring, M., & Debiak, M. (2021). Risk assessment for irritating chemicals – Derivation of extrapolation factors. *International Journal of Hygiene and Environmental Health*, *232*, 113668. https://doi.org/10.1016/j.ijheh.2020.113668

Price, P. S., Keenan, R. E., & Swartout, J. C. (2008). Characterizing interspecies uncertainty using data from studies of anti-neoplastic agents in

animals and humans. *Toxicology and Applied Pharmacology*, *233*, 64–70. https://doi.org/10.1016/j.taap.2008.03.026

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Renwick, A. G., Dorne, J-L. C. M., & Walton, K. (2001). Pathway-related factors: The potential for human data to improve the scientific basis of risk assessment. *Human and Ecological Risk Assessment: An International Journal*, *7*(1), 165–180. https://doi.org/10.1080/20018091094286

Schenk, L., & Johanson, G. (2010). Use of uncertainty factors by the SCOEL in their derivation of health-based occupational exposure limits. *Critical Reviews in Toxicology*, *40*, 791–798. https://doi.org/10.3109/10408444.2010.507628

Schenk, L., & Johanson, G. (2018). Use on uncertainty factors by the European Commission Scientific Committee on Occupational Exposure Limits: A follow-up. *Critical Reviews in Toxicology*, *48*, 513–521. https://doi.org/10.1080/10408444.2018.1483891

Schneider, K., Dilger, M., Drossard, C., Ott, H., & Kaiser, E. (2022a). Derivation of occupational exposure limits: Differences in methods and protection levels. *Journal of Applied Toxicology*. https://doi.org/10.1002/jat.4307

Schneider, K., Dilger, M., Kaiser, E., Kalberlah, F., & Wosniok, W. (2022b). Derivation of occupational exposure limits for airborne chemicals - Comparison of methods and protection levels. 1. Auflage. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin 2022. Seiten 876, Projektnummer: F 2437, PDF-Datei. https://doi.org/10.21934/baua:bericht20220112

Schneider, K., Oltmanns, J., & Hassauer, M. (2004). Allometric principles for interspecies extrapolation in toxicological risk assessment—Empirical investigations. *Regulatory Toxicology and Pharmacology*, *39*, 334–347. https://doi.org/10.1016/j.yrtph.2004.03.001

Schneider, K., Schuhmacher-Wolz, U., Hassauer, M., Darschnik, S., Elmshäuser, E., & Mosbach-Schulz, O. (2006). A probabilistic effect assessment model for hazardous substances at the workplace. *Regulatory Toxicology and Pharmacology*, *44*, 172–181. https://doi.org/10.1016/j.yrtph.2005.11.001

US EPA, Environmental Protection Agency. (2014). Guidance for applying quantitative data to develop data-derived extrapolation factors for interspecies and intraspecies extrapolation. EPA/100/R-14/002F. U.S. Environmental Protection Agency, Risk Assessment Forum, Office of the Science Advisor, Washington, DC. https://www.epa.gov/sites/production/files/2015-01/documents/ddef-final.pdf

Vermeire, T., Stevenson, H., Pieters, M. N., Rennen, M., Slob, W., & Hakkert, B. C. (1999). Assessment factors for human health risk assessment: A discussion paper. *Critical Reviews in Toxicology*, *29*, 439–490. https://doi.org/10.1080/10408449991349249

WHO, World Health Organization. (2014). Guidance document on evaluating and expressing uncertainty in hazard characterization. IPCS Harmonization Project Document 11. WHO, Geneva, Switzerland. http://www.inchem.org/documents/harmproj/harmproj/harmproj11.pdf

WHO, World Health Organization. (2020). Environmental Health Criteria 240, Principles and methods for the risk assessment of chemicals in food. IPCS International Programme on Chemical Safety; World Health Organization, Geneva, Switzerland. https://www.who.int/publications/i/item/9789241572408

Zarn, J. A., Engeli, B. E., & Schlatter, J. R. (2011). Study parameters influencing NOAEL and LOAEL in toxicity feeding studies for pesticides: Exposure duration versus dose decrement, dose spacing, group size and chemical class. *Regulatory Toxicology and Pharmacology*, *61*, 243–250. https://doi.org/10.1016/j.yrtph.2011.08.004

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Dilger, M., Schneider, K., Drossard, C., Ott, H., & Kaiser, E. (2022). Distributions for time, interspecies and intraspecies extrapolation for deriving occupational exposure limits. *Journal of Applied Toxicology*, *42*(5), 898–912. https://doi.org/10.1002/jat.4305