Toxicology Letters 218 (2013) 159–165

Contents lists available at SciVerse ScienceDirect

Toxicology Letters

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

Interspecies extrapolation based on the RepDose database—A probabilistic approach

Sylvia E. Escher[a,*], Monika Batke[a], Simone Hoffmann-Doerr[b], Horst Messinger[b], Inge Mangelsdorf[a]

[a] Fraunhofer Institute for Toxicology and Experimental Medicine, Nikolai Fuchs Str. 1, 30625 Hannover, Germany
[b] BASF Personal Care and Nutrition GmbH, Henkelstr. 67, 40589 Duesseldorf, Germany

HIGHLIGHTS

- Interspecies distribution function based on the broad database RepDose.
- Distributions evaluated for different routes of exposure, e.g. oral, inhalation.
- Derived interspecies distributions are best represented by log normal distribution.
- Allometric scaling correlates well to the GM of the derived distributions.
- Proposal of a resulting distribution applicable to probabilistic risk assessment.

ARTICLE INFO

Article history:
Received 22 November 2012
Received in revised form 16 January 2013
Accepted 17 January 2013
Available online 5 February 2013

Keywords:
Assessment factor
Interspecies
Risk assessment
Probabilistic

ABSTRACT

Repeated dose toxicity studies from the RepDose database (DB) were used to determine interspecies differences for rats and mice. NOEL (no observed effect level) ratios based on systemic effects were investigated for three different types of exposure: inhalation, oral food/drinking water and oral gavage. Furthermore, NOEL ratios for local effects in inhalation studies were evaluated. On the basis of the NOEL ratio distributions, interspecies assessment factors (AF) are evaluated.

All data sets were best described by a lognormal distribution. No difference was seen between inhalation and oral exposure for systemic effects. Rats and mice were on average equally sensitive at equipotent doses with geometric mean (GM) values of 1 and geometric standard deviation (GSD) values ranging from 2.30 to 3.08. The local AF based on inhalation exposure resulted in a similar distribution with GM values of 1 and GSD values between 2.53 and 2.70.

Our analysis confirms former analyses on interspecies differences, including also dog and human data. Furthermore it supports the principle of allometric scaling according to caloric demand in the case that body doses are applied. In conclusion, an interspecies distribution animal/human with a GM equal to allometric scaling and a GSD of 2.5 was derived.

1. Introduction

Risk assessment of human exposure to chemicals often relies on animal experimental data because human data in general are scarce or unethical to generate. In order to translate animal data to the human situation, assessment factors (AF) are used as tools to adjust this extrapolation.

Traditionally, a default factor of 10 was used to account for interspecies differences from animals and humans. In 1993, Renwick suggested dividing this traditional factor into a factor of 4 and 2.5 to account for differences in toxicokinetics and toxicodynamics, respectively. The REACH guidance now suggests a factor of 4 for rats to account for differences in toxicokinetics, which comprises differences related to metabolic rate. An additional factor of 2.5 is suggested to account for remaining interspecies differences, basically differences in toxicodynamics (ECHA, 2010).

The scientific basis behind the suggested AF for interspecies differences in toxicokinetics has been established by analysing ratios of NOELs, benchmark doses or other endpoint values in different species. Results indicate that the median or geometric mean (GM) of the distribution for systemic effects correlates to species specific differences in caloric demand/metabolic rate (Schneider et al., 2004; Vermeire et al., 1999; Rennen et al., 2001; Bokkers and Slob, 2007). Thus animal data can be extrapolated to humans on the basis of allometry considering effect level, caloric demand and body weight; calculated with the following allometric equation: scaling factor = kg(humans)/kg(animal species)$^{(0.25)}$ (Kalberlah and Schneider, 1998). Assuming an average body weight of 300 g for

* Corresponding author. Tel.: +49 511 5350 330.
E-mail address: Sylvia.Escher@item.fraunhofer.de (S.E. Escher).

0378-4274/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.toxlet.2013.01.027

rats, 30 g for mouse and 70 kg for humans, this results in allometric scaling factors of 7 (mouse/human) and 4 (rat/human) (ECETOC, 2010; ECHA, 2010).

There is still a debate whether the additional assessment factor of 2.5 to account for differences in toxicodynamics is scientifically substantiated and thus indispensable for human risk assessment (Falk-Filipsson et al., 2007). ECETOC has issued guidance for applying AF proposing the sole application of allometric scaling factors for interspecies extrapolation (ECETOC, 2010).

Other authors evaluated distributions on interspecies differences to account for the remaining interspecies susceptibility (Schneider et al., 2005; Vermeire et al., 1999; Baird et al., 1996; Rennen et al., 2001; Bokkers and Slob, 2007). The group of Schneider et al. compared animal to human studies, whereas the other reports used animal to animal comparisons. In these studies, the data sets compared for deriving interspecies factors were very limited in number and inhomogeneous which limits the representativeness of the assessment factors/the distribution with regard to the chemical and mechanistic domain.

Therefore in the present study a large data set of rat and mice studies was used, derived from the RepDose database (www.fraunhofer-repdose.de, Bitsch et al., 2006). The RepDose database was mined in order to identify chemicals being tested in both species under comparable conditions, e.g. exposure duration and route of application. Distributions of the NOEL ratios in both species were derived and compared to those recently described by Bokkers and Slob (2007), Schneider et al. (2005); Schneider et al. (2004), Vermeire et al. (1999) and Rennen et al. (2001). Based on this comparison we propose which distribution and consequently, which AF is likely to represent best chemical specific interspecies differences in toxicity when extrapolating to the human situation.

2. Materials and methods

Data from the RepDose database (DB) (www.fraunhofer-repdose.de) were used to evaluate interspecies differences. Currently, the RepDose DB contains about 2500 repeated dose toxicity studies in rodents obtained from approximately 700 organic chemicals. On average, the toxicity of each chemical is documented in 3–4 independent repeated dose toxicity studies, with a maximum of 14–19 studies for data rich chemicals. Each study derives a NOEC or NOEL value, which corresponds to the "no observed effect concentration (metric ppm)" or "no observed effect level (metric mmol/kg bw/day). NOEC(L) ratios (corresponding to assessment factors, AF) for one substance in the two species rat and mouse were calculated according to the following equation:

$$AF_{interspecies} = \frac{NOEC(L)_{mouse}}{NOEC(L)_{rat}}$$

The following routes of application were analysed separately for interspecies differences: inhalation; oral (food/drinking water); oral (gavage). For inhalation and oral treatment (food/drinking water) the metric ppm, for gavage the metric mmol/kg bw/day was evaluated. Using the metric ppm, a correction factor for caloric demand is not necessary as animals breathe and consume according to their caloric demand. The data set oral (gavage) was analysed with and without the application of allometric scaling factors. With an average rat weight of 250 g and an average mouse weight of 25 g, the scaling factor mouse versus rat is 1.78 according to the allometric equation.

LOEC(L) (lowest observed effect concentration(level)) replaced NOEC(L) ratios for pairs of studies, where in either study a NOEC(L) value was not available. NOEC(L) or LOEC(L) ratios were calculated for pairs of studies of the same study type; either subacute, subchronic or chronic studies. Narrow time frames were applied to define subacute (duration ranging from 21 to 32 days), subchronic (84–98 days) and chronic studies (greater than 699 days). Also, for study pairs the route of application had to be identical. For the oral route, feed to feed, drinking water to drinking water and gavage to gavage studies were compared. Furthermore, only studies of high to acceptable reliability, corresponding to Klimisch score 1 or 2 (Klimisch et al., 1997), were considered.

For inhalation, NOEC(L) values were based on systemic effects or on local effects observed in either species. Local effects in inhalation studies were defined as effects occurring in organs of first contact such as eye, nose, pharynx, larynx, trachea, lung, bronchi and respiratory system. For systemic effects from gavage application local effects in forestomach and stomach were excluded. The number of these effects was, however, too low for a separate analysis of local effects.

The distinction of route of exposure as well as systemic versus local N(L)OELs leads to the following data sets: inhalation (systemic); inhalation (local); oral food/drinking water; oral gavage (without allometric scaling); oral gavage (with allometric scaling) (Table 1).

In the case of data rich chemicals more than one pair of NOELs may be available and thus more than one interspecies assessment factor may be derived for an individual chemical. As this may thus trigger the resulting distribution curve, this "chemical" bias is analysed by differentiating for each data set:

- All AF – the data set contains all AF which meet the above described selection criteria. Multiple AF per chemical are possible.
- Median AF – the data set contains the median AF per chemical.

The characterisation of the empirical and fitted theoretical distributions was carried with the @risk software (Palisade Corporation, Ithaca, USA) (Appendix A). The full reference list of the data sets analysed in the present report is given in the supplemental material.

3. Results

NOEL ratios rats/mice and respective statistical analyses are shown in Table 1 for inhalation (systemic); inhalation (local); oral food/drinking water; oral gavage (without allometric scaling); oral gavage (with allometric scaling).

3.1. Systemic AF

The data sets not requiring allometric scaling comprised 94 pairs of inhalation studies for 58 chemicals and 66 pairs of oral food/drinking water studies, representing 50 chemicals. The four empirical distributions for oral and inhalation exposure show similar characteristics with regard to the geometric mean (GM) and geometric standard deviation (GSD) values. The GM values are 1, whereas the GSD values range from 1.99 to 3.53 (Table 1). Because of the observed statistical similarity, the data sets for inhalation and oral food/drinking water were pooled resulting in a data set of in total 160 assessment factors for 108 chemicals (data set "both-routes", Table 1). The empirical distributions for the data set "both-routes" also show a GM of 1 and a GSD of 2.54–2.65. Many AF are equal to 1, e.g. 18 AF (36%) in the oral, 15 AF (26%) in the inhalation and 33 (31%) in the data set "both-routes" at the AF median data level (data not shown). For this reason the GM of the six empirical distributions is 1.00. The GM of 1 indicates that both species are on average affected at the same dose level and are thus equally sensitive to the exposure of the tested chemical. The GSD of the distribution characterises the spread of the AF in the evaluated data sets.

In Table 1 besides the empirical distribution also the theoretical distribution for the data is shown. While the empirical distribution describes the variation of assessment factors across the analysed substances and is thus sample specific, the whole population can be estimated based on the sample data by assuming a theoretical distribution function. Therefore it was evaluated which theoretical distribution function best represents the empirical data. The fitted distribution may then serve to propose a general assessment factor based on a certain percentile. Several theoretical distribution functions were fitted to the empirical data. Goodness-of-fit tests, such as Kolmogorow–Smirnow (KS), as well as graphical plots (P–P plot and Q–Q plot) were performed. In Appendix 1 the goodness of fit is shown for the largest data set "both-routes" on the data level "all AF" ($N = 160$ data points).

The fitting of distributions to the empirical data shows, that the numerous AF values of 1 cannot be optimally modelled with a continuous distribution. The lognormal distribution for example slightly underestimates the 35th–50th percentiles and slightly overestimates the 50th–70th percentiles (Appendix 1). The GM of the fitted lognormal distribution, which assumes a continuous data set, thus deviates slightly from 1 and ranges from 0.78 to

Table 1
Summary of the interspecies distributions for oral and inhalation application.

Exposure	Type	Data level	N	Empirical distribution					Fitted lognormal distribution				
				GM	GSD	10th	90th	95th	GM	GSD	10th	90th	95th
Inhalation[a]	Systemic	All AF	94	1.00	2.68	0.25	3.96	5.10	1.00	3.08	0.24	4.28	6.45
		Median AF	58	1.00	3.53	0.27	4.10	8.00	1.16	2.74	0.32	4.25	6.13
Oral (food/drinking water)[a]		All AF	66	1.00	1.99	0.25	2.00	3.13	0.78	2.51	0.24	2.56	3.59
		Median AF	50	1.00	1.99	0.20	2.00	3.13	0.83	2.30	0.28	2.41	3.27
Both routes[a]		All AF	160	1.00	2.65	0.25	3.33	5.00	0.91	2.87	0.23	3.51	5.12
		Median AF	108	1.00	2.54	0.25	3.13	4.67	0.98	2.50	0.32	3.02	4.43
Oral (gavage)[b,c]		All AF	78	1.67	2.39	0.67	4.00	7.03	1.59	2.50	0.49	5.16	7.21
		Median AF	56	1.97	1.57	0.75	4.00	4.13	1.77	2.23	0.63	4.97	6.67
Oral (gavage)[b,d]		All AF	78	0.95	2.39	0.38	2.29	4.02	0.91	2.50	0.28	2.95	4.12
		Median AF	56	1.13	1.57	0.43	2.29	2.36	1.01	2.33	0.36	2.84	3.81
Inhalation[a]	Local	All AF	72	1.00	2.68	0.25	4.06	5.10	1.09	2.70	0.30	3.92	5.63
		Median AF	40	1.00	2.51	0.33	5.00	5.50	1.2	2.53	0.36	3.97	5.56

[a] Metric ppm.
[b] Metric mmol/kg bw/day.
[c] Allometry not considered.
[d] Allometry considered.

1.16. The transformed ln(AF)-values, however, fit well to a normal distribution, with statistical characteristics close to the empirical values (Appendix 1). Taking the fitting of the transformed and non-transformed values into account, it can be concluded that the lognormal distribution is best to model the empirical data (Table 1, right column). The lognormal distribution fits well to the low and high ratios resulting in percentiles close to the empirical data. The corresponding GSD values are therefore good estimates to characterise the spread of the distribution curves and thus the variability of AF in the six data sets. The lognormal distribution is less suitable to model the 50th percentile, which is 1 according to the experimental data (summarised in Table 2).

In addition to the types of application following allometric principles, the data set oral gavage consisting of 78 pairs of studies, representing 56 chemicals was analysed (Table 1) with and without application of allometric scaling factors to the AF. The GM of the fitted distributions is 1.59/1.77 (median AF/all AF). After correction for allometry the GM decreases to 0.91/1.03 (Table 1, median AF/all AF). With allometric scaling, 17 out of 78 AF (22%) are equal to 1.The GSD values, ranging from 2.23 to 2.50, are similar to those observed for the data sets oral food/drinking water and inhalation.

3.2. Local AF

Interspecies differences for local effects in inhalation studies were evaluated for 72 pairs of studies representing 40 substances in rats and mice. Out of the 40AF, 15 AF (38%) are equal to 1. As already seen for the systemic assessment factors, none of the continuous theoretical distribution functions fits perfectly to the experimental data. As demonstrated for the systemic data sets, the lognormal distribution fits the high and low percentiles of the empirical interspecies distributions (Table 1). Therefore, the lognormal distribution has been considered to be a good estimate. The resulting distributions are similar to the distribution based on systemic effects, with GM values of 1 and the GSD values between 2.53 and 2.70.

3.3. General

In Table 1 also the 90th and 95th percentiles for the distribution of NOEL ratios are shown, both for the empirical and the fitted data. As already indicated by the low value of the GSD, the factors obtained are rather homogeneous. For the fitted distributions,

Table 2
Comparison of distributions for interspecies differences (RepDose results shown for "all AF").

Source	Allometry applied	Reference value	Species	Route	Number	GM	GSD	95th
Vermeire et al. (1999)	Yes	NOAEL	Mouse/rat	Oral	67 substances	2.4	5.7	42.2
			Mouse/rat	Inhalation	21 substances	3.1	7.8	91.8
			Rat/dog	Oral	63 substances	0.5	5.1	6.6
			Mouse/dog	Oral	40 substances	1.3	6.1	24.9
Rennen et al. (2001)	Yes	NOAEL	Mouse/rat	Oral	78 substances	1.9	4.4	21
			Mouse/rat	Inhalation	19 substances	1.5	3.5	11
			Rat/dog	Oral	71 substances	0.8	4.6	9
			Mouse/dog	Oral	20 substances	1.2	3.7	10
Vermeire et al. (2001)	Yes	NOAEL	Mouse, rat, dog	Oral/Inhalation	184 ratios	1.0	4.5	11.9
Schneider et al. (2002) and BAuA (2005)	Yes	(MTD, LD10, TDL)/MTD	(Mouse, hamster, rat, monkey)/human	Parenteral	183 ratios (63 substances)	1.0	3.2	6.7
Bokkers and Slob (2007)	No	NOAEL	Mouse/rat	Oral	135 ratios	1.9	3.1	12.0
		CED		Oral	135 ratios	1.9	2.0	5.7
RepDose	No	NOEC	Mouse/rat	Oral[a]/inhalation	160 ratios (108 substances)	1.0	2.9	5.1
	No	NOEL	Mouse/rat	Gavage	78 ratios (56 substances)	1.6	2.5	7.2
	Yes	NOEL	Mouse/rat	Gavage		1.0	2.5	4.1

[a] Food/drinking water.

Table 3
Strengths and limitations of the discussed interspecies distributions, the derived experimental GSD (Exp. GSD) as well as the expected trend for the true GSD value are indicated.

Dataset	Strength	Exp. GSD	Limitation	Trend for GSD
RepDose (present report)	Broad chemical domain	2.9	Interrodent comparison	Increase
	Comparable studies (exposure duration, route of application)		NOEL approach	Decrease
Vermeire/Rennen et al. (2001)	Broad chemical domain	4.5	Less comparable studies (broad categories for exposure duration, global oral exposure)	Decrease
	Interrodent and rodent to non-rodent comparisons		NOAEL approach	Decrease
Schneider et al. (2002, 2005)	Animal to human comparisons	3.2	Small chemical and mechanistic domain	No estimation
			Heterogeneity of evaluated endpoints	Decrease
			Parenteral dosing	No estimation
Bokkers and Slob (2007)	BMD approach	2.0 (2.4)	Interrodent comparisons	Increase
	Comparable studies (exposure duration, route of application)		Small chemical domain	No estimation
	Comparable studies (exposure duration, route of application)	3.1	Interrodent comparisons	Increase
			Small chemical domain	No estimation
			NOAEL approach	Decrease

the 90th percentile ranges from 2.4 to 5.2 and the 95th percentile ranges from 3.3 to 7.2.

4. Discussion

The distinction of "all AF" versus "median AF" has no major impact on the resulting distribution functions. This implicates that a "chemical" bias is not introduced when all AF per chemical were considered for the analyses, compared to the median AF. Therefore, the discussion focuses on the analyses of the larger data sets containing "all AF" in the following. First, the presented distributions curves are discussed with regard to toxicokinetic and toxicodynamic differences. Second, other sources of data variability are discussed and the presented distributions are compared to recently published reports (Bokkers and Slob, 2007; Schneider et al., 2004, 2005; Vermeire et al., 1999; Rennen et al., 2001). Finally, a distribution for the extrapolation animal/human is suggested.

4.1. Toxicodynamic and toxicokinetic interspecies differences

In the present report different data sets were evaluated to assess interspecies differences taking the rat/mouse comparison as model. Without allometric scaling, the GM of the distribution curve gavage oral is 1.7 (Table 1). Applying allometric scaling the distribution curve shifts left, resulting in a GM of about 1.0 supporting the principle of allometric scaling according to caloric demand. The other evaluated data sets in the present report (inhalation local/systemic; oral food/drinking water; both-routes) analyse equipotent doses (metric ppm), which already includes allometric scaling as animals breathe and consume according to their caloric demand. The GM of all resulting distribution curves is 1, indicating that both species are on average equally sensitive to equipotent doses. The type of oral application (bolus versus continuous dosing) has no major impact on interspecies differences. For inhalation there are no differences between local and systemic effects with regard to interspecies differences. Taking these finding as model for interspecies differences in general it can be assumed that the existing allometric factors are suitable for interspecies extrapolation. If the GM forms the basis for the assessment factor, as this is the case also for the assessment factors for study duration (ECHA, 2010; Batke et al., 2011), no further assessment factor on remaining uncertainties would be necessary.

There is, however, concern for the remaining interspecies differences (ID), which do not only reflect toxicodynamic differences, but also toxicokinetic/metabolic differences not implied in allometric scaling and, as discussed in more detail in the following, experimental variation. These may be covered by taking a higher percentile, i.e. the 90th or 95th percentile of our distribution. The 90th percentile would be a factor of maximum 5, the 95th a maximum of 7. The spread of the distribution is also characterised by the GSD which ranges from 2.2 to 3.1 in our analysis. If we consider that rats and mice are closely related species, the remaining interspecies differences may be underestimated in our data set. On the other hand, one has to bear in mind that the spread of the distribution does not purely depict species specific differences in toxicity but also involves data variability caused by differences in, e.g. study design, homogeneity of evaluated endpoints, and so on. In the following, the strength and limitations of our data sets are compared to recently published distribution functions investigated by Schneider et al. (2005), Vermeire et al. (1999) and Rennen et al. (2001) and Bokkers and Slob (2007) (summarised in Table 3). Except for Bokkers and Slob (2007), the *in vivo* values in the cited analyses were corrected with the respective species specific allometric scaling factor: 4 for rat, 7 for mouse and 1.6 for dog. Bokkers and Slob (2007) applied both a NOAEL (no observed adverse effect level) and a BMD (benchmark dose) approach to derive interspecies distribution functions.

Vermeire et al. (1999) compared 184 NOAEL ratios of substances tested in three species, namely mouse, rat and dog. The data set included oral and to a lesser extent inhalation studies on pesticides and other existing chemicals from IPCS Environmental Health Criteria documents and JMPR evaluations. NOAEL ratios were analysed for pairs of studies of one exposure duration category (subacute or semi-chronic exposure). Subacute studies were defined to range from 21 to 50 days for mice and rats (28–90 days for dogs), semi-chronic studies from 90 to 730 days for mice and rats (365–730 days for dogs). A resulting distribution with a GM of 1 and a GSD of 6 was proposed. The authors however discussed that the relatively broad spread of the distribution (GSD of 6) may have been caused by differences in experimental conditions, which are not related to differences in species susceptibility. For example, NOAEL ratios were not based on the same critical effect and there were differences in strains and in substance purity. In a further report, the same data set was re-evaluated and extended to 188 values (Rennen et al., 2001). In this analysis, exposure duration categories were more strictly defined. The semi-chronic category was split into a subchronic (90–365 days – rat/mice, >90 days – dogs) and a chronic category (1–2 years – rat/mouse). The resulting distributions show smaller GSD values and GM values closer to

1 (Table 2). As seen in the RepDose data set (present study), the distributions do not differ for the oral and inhalation route. This leads to the proposal of one overall distribution with a GM of 1 and a GSD of 4.5 (Vermeire et al., 2001). The GSD value of the RepDose data sets ranging from 2.5 to 3.1 are considerably smaller than the GSD of 4.5 described by Rennen et al. (2001), although both analyses are based on a comparable approach with regard to the amount of data points and the investigated endpoint.

The RepDose data set compares interrodent differences, whereas Vermeire/Rennen et al. also included rodent to non-rodent (dog) ratios, for which usually a higher interspecies difference is assumed. The distribution functions of Rennen et al. do, however, not show higher differences for the rodent to non-rodent comparison (Table 2). The mouse–dog and the rat–dog distributions have GSDs of 3.7 and 4.6, respectively, being in the same range as the GSD values of the inter-rodent distributions, 4.4 (oral)/3.5 (inhalation). Ideally, a NOAEL ratio is derived from duplicate studies from one laboratory, using identical dose selection, dose spacing, same critical effect and study types. The scope of examination differs for different study types, e.g. studies of short- and long term duration. The re-evaluation of Vermeire's data set by Rennen et al. demonstrates that data variability is reduced if studies of better comparable study duration are used. In our analyses, the definition of comparable study duration was even stricter. Furthermore, in the RepDose data set generalised oral to oral comparisons were replaced by more specific feed to feed and drinking water to drinking water and gavage to gavage ratios. Both restrictions increase the comparability of study pairs and thus reduce the uncertainty of the NOEC/NOEL ratios and the GSD of the resulting distribution (Table 3). As a result, about one third of the AF in the RepDose distributions is equal to 1 (26–38%) if equipotent doses are compared.

Schneider et al. (2002, 2004; BAuA, 2005) investigated the remaining susceptibility differences between animals and humans by analysing toxicity data from 63 anti-neoplastic agents tested in monkeys, rabbits, mice, rats, dogs and humans. 187 dose ratios ($\text{Critical dose}_{\text{Animal, substance } x}/\text{Critical dose}_{\text{Human, substance } x}$) were pooled and resulted in a lognormal distribution with a GM of 1 and GSD of 3.2 (Table 2). The direct comparison of animal and human data is the main advantage of this analysis, as the protection of humans is the aim of regulatory risk assessment. However, some limitations in the data set restrict the general applicability of the resulting distribution to risk assessment. The data set consists of subacute studies in humans and animals compiled from six different publications (Freireich et al., 1966; Goldsmith et al., 1975; Schein et al., 1979; Rozencweig et al., 1981; Grieshaber and Marsoni, 1986; Paxton et al., 1990). Differences in study duration were circumvented by normalising the cumulative doses to 5 days of exposure. Maximal tolerated doses (MTD) were evaluated in humans being exposed for 5 days. These values were compared to LD10 doses (rodents) and TDL (Toxic Dose Low)/MTD values (dogs or monkeys). The authors noted however, that these evaluated human and animal reference values were of limited comparability. In addition, the human data were obtained from cancer patients, for whom the susceptibility to anti-neoplastic agents may differ, compared to the normal population. In the chosen study, humans and animals were exposed using parenteral dosing, e.g. intraperitoneal and intravenous application. As discussed by the authors (Schneider et al., 2002, 2004; BAuA, 2005), it is likely that parenteral dosing will result in a different bioavailability compared to oral or inhalation exposure. Human exposure to chemicals for consumers is, however, normally via the oral route, where species specific metabolism in the liver plays an important role. At the workplace, inhalation or dermal exposure are usually the prominent routes of exposure. Another difference is bolus versus continuous dosing. The impact of the parenteral dosing on the variability of the data points and its difference compared to inhalation and oral dosing cannot be estimated. Schneider et al. (2002) note that the anti-neoplastic chemicals in their data set are mainly directly alkylating substances with closely related chemical structures and similar mode of action in vivo. The range of chemical studies as well as the mechanistic domain of the distribution are therefore relatively small and can thus not be regarded as representative for all types of chemicals such as industrial chemicals currently assessed under REACH. Despite the difference of evaluated endpoints and data compilation the RepDose and Schneider distributions show nearly identical characteristics with a GM of approximately 1 and GSD of 2.9 and 3.2, respectively.

Bokkers and Slob (2007) compared the NOAEL and CED (Critical effect doses) – based on benchmark dose approach – using oral subchronic to chronic NTP repeated dose toxicity studies being tested in rats and mice. NOAELs and alternatively CEDs were established for changes in: body weight, liver and kidney weight (absolute and relative weight), and erythrocyte counts. A critical effect size of 5% was chosen for the CED analysis. Only studies with similar exposure duration and similar oral application, e.g. feed–feed, gavage–gavage or drinking water–drinking water, were compared. Allometric scaling was not applied. The final data set consists of 985 ratios from 91 pairs of studies based on 58 compounds. From these 985 ratios, 135 ratios could be derived for both the NOAEL and benchmark dose approach. The CED and the NOAEL distributions of these 135 ratios (the number of chemicals is not given) showed an identical GM of 1.9. The benchmark dose approach resulted in a GSD value of 2.0 whereas the NOAEL based distribution showed a GSD of about 3.1. The organ specific evaluation of NOAEL ratios differs from our approach, in which the type of target organ/effect at NOAEL was not considered. In regulatory risk assessment, it is, however, also not common practice to consider only those dose levels/NOAELs, where similar effects in both species are observed. The GM 1.6/1.8 of our data sets (all AF/median AF) RepDose gavage is in the same range as those of the data set Bokkers and Slob with 1.7–2.1 (BMD) and 1.6–2.3 (NOAEL). Both analyses encourage the principle of allometric scaling according to caloric demand, which has been earlier proposed by several groups, e.g. Kalberlah and Schneider (1998), Vermeire et al. (1999), and Schneider et al. (2002). The authors discussed that the GSD of 2.0 might underestimate the true variation in CED ratios, as ratios without similar dose–response in both species were not included in the analysis. Including these data, the GSD increased to 2.4. The authors concluded that the CED ratios of the benchmark dose approach results in less uncertainty than the NOAEL ratios and thus are a better indicator of substance induced interspecies difference. The chemical domain of this analysis is however limited, consisting at maximum of 58 compounds.

For local effects, our analysis of the distribution of AF based on local effects after inhalation exposure showed the same characteristics as the distributions based on systemic effects. The geometric mean of the empirical data is 1, indicating again that on average both species are equally sensitive to equipotent doses. The GSD value (2.7) is similar to those derived for the data set "inhalation systemic" and "both-routes". Local toxicity includes irritation and inflammation reactions in tissues of first contact, e.g. the respiratory tract. Although other aspects following allometric principles are not involved in local toxicity, the rate of metabolisms depends on allometric principles. Thus generation of reactive metabolites seems to be relevant for toxicity in the respiratory tract.

The data set "inhalation local" is, however, relatively small with only 40 different chemicals so that its representativeness with regard to its chemical and mechanistic domain is limited. Kalberlah et al. (2002) investigated interspecies differences for locally acting substances using ATSDR toxicological profile reports. Rodent and human data were compared. The authors noted that the data set was very small and the evaluated endpoints differed significantly

between species. AF and their distributions were not reported. Rodents breathe by nose only, so it could be assumed that the respiratory tract of rodents is more sensitive to local effects than in humans. Nonetheless, Kalberlah et al. concluded that rodents were not more sensitive to irritants compared to humans.

Table 3 summarises the strength and limitations of our data and the three discussed analyses and indicates whether the limitations are expected to increase/decrease the spread of the respective distribution and thus its GSD value. None of the data sets is able to depict only toxicity induced interspecies differences ("true" GSD). This is in line with previous analyses (Batke et al., 2011). These demonstrated that differences in study design such as dose selection and dose spacing contribute to the variability of the NOAEL ratios and thus increase GSD values of the distribution functions. Critical effect doses (CED) based on a critical effect size of selected endpoints are less dependent on experimental conditions and, as shown by Bokkers and Slob (2007), are a better basis for deriving an interspecies distribution. Other sources of data variability are heterogeneity of evaluated endpoints, limited chemical domain, interrodent versus non-rodent ratios (Table 3). Our data sets show, that the spread of the distribution (GSD) is not dependent on the application of allometric scaling (Table 2). Taking our analyses together, the spread of the distribution depends strongly on the quality of the study pairs compared. Furthermore, the evaluated interspecies differences and thus the derived GSD values are not strictly independent from intraspecies differences because different mice and rat strains are compared.

5. Conclusion

Our data as well as other analyses based on the NOAEL comparisons (Vermeire/Rennen et al., 2001; Schneider et al., 2002, 2005) do all derive a GM of 1 for interspecies differences. This supports the biological hypothesis that species are on average equally sensitive to equipotent doses, if doses are related to energy turnover, such as uptake via inhalation, drinking water or food.

Remaining uncertainties, represented by 90th percentiles of maximum 5 or a GSD of 2.5, are rather low, if one bears in mind that they might comprise not only toxicodynamic differences between the two species, but also differences in study design such as dose spacing, number of animals and evaluated endpoints, and that they may include also interindividual differences in susceptibility between animals of the same species.

We therefore suggest a default interspecies extrapolation distribution animal/human with a GM identical to the allometric scaling factor, e.g. 4 rat/human, 7 mouse/human (ECHA, 2010) for body doses (e.g. gavage studies), based on allometric principles and a factor of 1 for doses in ppm or mg/m^3 (e.g. uptake via food, drinking water or inhalation). This reflects the geometric mean of the analysed distributions for interspecies differences in the present report.

Distributions such as the present one for interspecies differences can be used in regulatory risk assessment to substantiate and/or replace the currently used deterministic assessment factors. Furthermore, the distributions of single assessment factors could be combined by using mathematical models such as the Monte Carlo simulation (Schneider et al., 2005). The resulting distribution of this probabilistic risk assessment could serve to recommend an "overall" assessment factor. In future analyses of the database RepDose, certain chemical domains/mode of actions could be identified for which higher/lower AF than the standard factor would be appropriate.

Conflict of interest

The authors have no conflict of interest.

Acknowledgements

The authors like to thank CEFIC LRI and the EU FP-6 project OSIRIS for the sponsoring of parts of the FhG database RepDose. The work presented in this publication was funded by ERASM.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.toxlet.2013.01.027.

References

Baird, S.J.S., Cohen, J.T., Graham, J.D., Shlyakhter, A.I., Evans, J.S., 1996. Noncancer risk assessment: a probabilistic alternative to current practice. Human and Ecological Risk Assessment 2, 79–102.

Batke, M., Escher, S., Hoffmann-Doerr, S., Melber, C., Messinger, H., Mangelsdorf, I., 2011. Evaluation of time extrapolation factors based on the database RepDose. Toxicology Letters 205, 122–129.

BAuA, 2005. Uncertaintyanalysis in Workplace Effect Assessment. Research Report F1824, F1825, F1826. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Dortmund/Berlin/Dresden, 249 pp.

Bitsch, A., Jacobi, S., Melber, C., Wahnschaffe, U., Simetska, N., Mangelsdorf, I., 2006. REPDOSE: A database on repeated dose toxicity studies of commercial chemicals – a multifunctional tool. Regulatory Toxicology and Pharmacology 46, 202–210.

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, 355–373.

ECETOC, 2010. Guidance on Assessment Factors to Derive a DNEL. Brussels (Technical Report Nr. 110), 202 pp.

ECHA, 2010. Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.8: Characterisation of Dose [Concentration]–Response for Human Health. Version 2. European chemicals Agency, Helsinki, Finnland.

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.

Freireich, E.J., Gehan, E.A., Rall, D.P., Schmidt, L.H., Skipper, H.E., 1966. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey and man. Cancer Chemotherapy Reports 50, 219–244.

Goldsmith, M.A., Slavik, M., Carter, S.K., 1975. Quantitative prediction of drug toxicity in humans from toxicology in small and large animals. Cancer Research 35, 1354–1364.

Grieshaber, C.K., Marsoni, S., 1986. Relation of preclinical toxicology to findings in early clinical trials. Cancer Treatment Reports 70, 65–72.

Kalberlah, F.S., Schneider, K. (Eds.), 1998. Quantifizierung von Extrapolationsfaktoren: Endbericht des Forschungsvorhabens Nr. 116 06 113 des Umweltbundesamtes. Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin: Forschung 796. Wirschaftsverlag NM, Bremerhaven, p. 200 pp.

Kalberlah, F., Föst, U., Schneider, K., 2002. Time extrapolation and interspecies extrapolation for locally acting substances in case of limited toxicological data. Annals of Occupational Hygiene 46, 175–185.

Klimisch, H.J., Andreae, M., Tillmann, U., 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regulatory Toxicology and Pharmacology 25, 1–5.

Paxton, J.W., Kim, S.N., Whitfield, L.R., 1990. Pharmacokinetic and toxicity scaling of the antitumor agents amsacrine and CI-921, a new analogue, in mice, rats, rabbits, dogs, and humans. Cancer Research 50, 2692–2697.

Rennen, M.A.J., Hakkert, B.C., Stevenson, H., Bos, P.M.J., 2001. Data-base derived values for the interspecies extrapolation: a quantitative analysis of historical toxicity data. Comments on Toxicology 7, 423–436.

Renwick, A.G., 1993. Data-derived safety factors for evaluation of food additives and environmental contaminants. Food Additives and Contaminants 10, 275–305.

Rozencweig, M., Von Hoff, D.D., Staquet, M.J., Schein, P.S., Penta, J.S., Goldin, A., Muggia, F.M., Freireich, E.J., DeVita, V.T., 1981. Animal toxicology for early clinical trials with anticancer agents. Cancer Clinical Trials 4, 21–28.

Schein, P., Davis, R.D., Carter, S., Newman, J., Schein, D.R., Rall, D.P., 1979. The evaluation of anticancer drugs in dogs and monkeys for the prediction of qualitative toxicities in man. Clinical Pharmacology and Therapeutics 11, 3–40.

Schneider, K., Hassauer, M., Oltmanns, J. (Eds.), 2002. Überprüfung der maßgerechten Übertragung (Scaling) von Schadstoffen aus Tierversuchen auf den Menschen (Interspeziesextrapolation). (UFOPLAN) 201 65 202. Umweltforschungsplan des Bundesministeriums für Umwelt, Naturschutz und Reaktorsicherheit: Aktionsprogramm "Umwelt und Gesundheit", Freiburg. Forschungs- u. Beratungsinstitut Gefahrstoffe GmbH (FoBiG), 257 pp.

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.

Schneider, K., Hassauer, M., Oltmanns, J., Schuhmacher-Wolz, U., Elmshäuser, E., Mosbach-Schulz, O., 2005. Uncertainty Analysis on Workplace Effect Assessment. Research Report F 1824, F 1825, F 1826. Publication Series of the Federal Institute for Occupational Safety and Health, Dortmund, Germany, 249 pp.

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. CRC Critical Reviews in Toxicology 29, 439–490.

Vermeire, T., Pieters, M.N., Rennen, M., Bos, P., 2001. Probabilistic Assessment Factors for Human Health Risk Assessment. RIVM Report 601516005; TNO Report V3489. RIVM, Bilthoven, The Netherlands.