Toxicology and Applied Pharmacology 233 (2008) 64–70

Contents lists available at ScienceDirect

Toxicology and Applied Pharmacology

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

Characterizing interspecies uncertainty using data from studies of anti-neoplastic agents in animals and humans

Paul S. Price [a,*], Russell E. Keenan [b], Jeffrey C. Swartout [c]

[a] The Dow Chemical Company, Toxicology & Environmental Research & Consulting, 1803 Building, Midland MI 48674, USA
[b] AMEC Earth and Environmental, 15 Franklin Street, Portland, ME 04101, USA
[c] National Center for Environmental Assessment U.S. Environmental Protection Agency, 26 W. M. L. King Drive, Cincinnati, OH 45268, USA

ARTICLE INFO

Article history:
Available online 16 April 2008

Keywords:
Uncertainty factors
Interspecies
Reference dose
Risk characterization

ABSTRACT

For most chemicals, the Reference Dose (RfD) is based on data from animal testing. The uncertainty introduced by the use of animal models has been termed interspecies uncertainty. The magnitude of the differences between the toxicity of a chemical in humans and test animals and its uncertainty can be investigated by evaluating the inter-chemical variation in the ratios of the doses associated with similar toxicological endpoints in test animals and humans. This study performs such an evaluation on a data set of 64 anti-neoplastic drugs. The data set provides matched responses in humans and four species of test animals: mice, rats, monkeys, and dogs. While the data have a number of limitations, the data show that when the drugs are evaluated on a body weight basis: 1) toxicity generally increases with a species' body weight; however, humans are not always more sensitive than test animals; 2) the animal to human dose ratios were less than 10 for most, but not all, drugs; 3) the current practice of using data from multiple species when setting RfDs lowers the probability of having a large value for the ratio. These findings provide insight into inter-chemical variation in animal to human extrapolations and suggest the need for additional collection and analysis of matched toxicity data in humans and test animals.

Introduction

Regulatory agencies evaluate the safety of chemicals with respect to noncarcinogenic health effects by comparing either modeled or measured doses received as a result of exposure to doses believed to be "protective" of both the general population and sensitive individuals. These estimates of "protective" doses, most often expressed in units of milligram of chemical per kilogram body weight per day of exposure (mg/kg-day), include the US EPA's Reference Dose (RfD) (USEPA, 1988) and the Agency for Toxic Substances and Disease Registry's (ATSDR's) Minimum Risk Level (MRL; ATSDR, 1996). Other metrics include the European Union's Tolerable Daily Intake (TDI; EFSA, 2007), and Derived No Effect Levels (DNELs; European Union, 2007). While differing in certain details, these metrics have a common basis in their derivation, in that when human data are not available, they rely on test-animal toxicology studies. They are established by taking the dose associated with a specific toxicological endpoint and dividing the dose by a series of uncertainty factors.

One of these factors is the interspecies uncertainty factor which in the U.S. is typically assigned a value of 10. This factor addresses the fact that test animals may have higher tolerances to chemicals than humans and that the magnitude of the difference is likely to vary with the test species and the chemical. The value of 10 is based on the assumption that humans are unlikely to be 10 times more sensitive to a chemical than a test species when the doses are expressed on a mg of chemical per kg of body weight basis (Dourson and Stara, 1983).

In the last twelve years, a number of researchers have represented the uncertainty in noncancer risk assessment using probabilistic models (Baird et al., 1996; Swartout et al., 1998; Slob and Pieters, 1997, 1998; Vermeire et al., 1999, 2001; Kalberlah et al., 2003; Kodell and Chen, 2007), integrating this information into estimates of uncertainty and variation in risk findings (Price et al., 1997; Carlson-Lynch et al., 1999; Bosgra et al., 2005; Van der Voet and Slob, 2007). Such an approach requires a probability distribution that describes the uncertainty in the difference in toxicity between the test animal and humans for a given chemical. A number of researchers (Weil and McCollister, 1963; Weil, 1972; Dourson, 1994; Nair et al., 1995; Nauman and Weideman, 1995; Nessel et al., 1995) have suggested that this uncertainty might be characterized based on the inter-chemical variation in the ratio of similar toxicity endpoints in humans and animals.

Other researchers have proposed more detailed conceptual frameworks for defining this distribution. The approach in Price et al. (1997, 1999) and Swartout et al. (1998) defines the distribution in terms of the ratio of No Observed Adverse Effects Levels (NOAELs). Baird et al. (1996), Slob and Peiters (1997, 1998), and Evans et al. (2001) define the uncertainty distribution based on the ratios of ED_{50}s for the critical effects. Frameworks aslo differ on the issue of the existence or non existence of thresholds (Price et al., 1997, Evans et al., 2001).

* Corresponding author. Fax: +1 989 638 2425.
E-mail address: pprice@dow.com (P.S. Price).

0041-008X/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.taap.2008.03.026

A number of studies have focused on the empirical characterization of differences in noncancer responses across species of test animals (Rhomberg and Wolff, 1998; Kalberlah et al., 2002; Schneider et al., 2004; Bokkers and Slob, 2007). In addition, differences between test animals and humans have been defined based on physiological and toxicokinetic considerations (Dourson and Stara, 1983; Andersen et al., 1995) and in terms of consistency with historical regulatory policies (Slob and Pieters, 1997, 1998; Swartout et al., 1998; Evans et al., 2001).

In this paper, we present an analysis of a toxicity data set for anti-neoplastic agents in test animals and humans. The data consist of a series of toxicological endpoints that can be considered roughly equivalent to an acute maximum tolerated dose (MTD). These data are used to calculate ratios of the MTD in humans and test animals (toxicity ratios). Using these ratios we investigate the inter-chemical variation in the differences between humans and test animals, the effect of different test animals, the impact of sample size, and the impact of having test data in multiple species.

We then discuss the relevance of these distributions for the extrapolation of toxicity measurements from test animals to humans, required probabilistic models of noncancer risks. As part of this discussion, consideration is given to the limitations of the data set and concerns raised by Brand et al. (1999, 2001) on the use of ratios of empirically measured toxicological endpoints.

The anti-neoplastic agent data set

Anti-neoplastic agents are used in chemotherapy to preferentially destroy cancerous cells. Because differences in the relative toxicity to normal and cancerous cells for such agents may be small, chemotherapy protocols often call for administration of anti-neoplastic agents at levels that are close to doses that cause significant toxicity in humans. As part of the development of such agents, the National Cancer Institute (NCI) determines the maximum tolerated dose in humans (MTD_H) in short-term studies (typically 5 days). The MTD_H is defined as the dose level at which no more than one of six cancer patients experience dose limiting toxicity with the next higher dose group of six patients having two or more patients experiencing dose limiting toxicity (Storer, 1989).

During developmental trials for anti-neoplastic agents, the NCI estimates the MTD_H based on the results of a progressive series of acute, subacute, and subchronic toxicological studies in a battery of animal tests. This work results in toxicological data in a number of animal species. Toxicity in animals is determined differently for different species:

- LD_{10} Typically derived for mice and rats. The acute (single) dose resulting in the death of 10% of a population of test animals.
- TDL (Toxic Dose Low) Typically derived for dogs and monkeys. The lowest dose that produces pathological alterations in hematological, chemical, clinical, or morphological endpoints. Doubling the TDL produces no lethality.
- MTD_A The highest dose in test animals that suppresses body weight by no more than 10% in a 90-day subchronic study.

Source (Grieshaber and Marsoni, 1986).

The data for this study were taken from six articles, which were identified as containing data on the toxicity of agents in patients and in one or more test animals (Freireich et al., 1966; Goldsmith et al., 1975; Schein et al., 1979; Rozensweig et al., 1981; Grieshaber and Marsoni, 1986; Travis and White, 1988; Paxton et al., 1990). Only data on compounds administered by intravenous or intraperitoneal routes were included in the data set. For the majority of compounds the route of administration was consistent across test animals and humans. When necessary, data were normalized to a 5-day dosing regime (Freireich et al., 1966; Travis and White, 1988). This dose was estimated by summing the total dose administered to the animals over the course of the testing and dividing by five. Finally, where doses were reported in units of mass per surface area mg/m^2, the doses were converted to mg/kg based on standard estimates of the surface area and body weights of the relevant species (Freireich et al., 1966; Travis and White, 1988).

A total of 61 compounds animal to human was identified with data in humans and in one or more species of test animals. There were 161 animal to human ratios in total, with the number of ratios varying by species and ranging from 56 for the dog to 19 for the rat. The adjusted toxicity data from these six studies are presented in Table 1.

Relevance of the data set to probabilistic noncancer risk assessments

While the unique nature of these data (matched human and animal toxicity data) make this data set highly relevant to the investigation of noncancer risk assessments, there are several significant limitations in the use of the data on anti-neoplastic agents. The following is a description of some of these limitations and the ways that they could cause the distribution of MTD ratios to differ from measures of human and animal thresholds required by probabilistic models.

1. *The MTD differs from NOAELs (and benchmark doses) used in the derivation of RfDs.* The MTD is a dose that is likely to be associated with some level of effects and is not equivalent to a NOAEL. Indeed, it may be greater than the Lowest Observed Adverse Effect Level (LOAEL). If the dose–response curves in humans and animals diverge at lower doses, the ratios at the different effect levels will differ. As they are not in the lower dose range, distributions of MTDs may be more similar to those based on ED50s used in the approaches suggested by Baird et al. 1996), Slob and Peiters (1997, 1998), and Evans et al. (2001).
2. *The data are drawn from short-term studies, while RfDs generally address chronic effects.* It is generally considered that chronic effects occur at lower doses than acute effects. However, it is not clear that the ratios of acute doses across species will be higher or lower than ratios of chronic doses. Thus, there is no a priori reason to suspect that the ratios of MTDs and chronic thresholds will have different median values. Differences in the variation between the two distributions might be expected to occur because there are more ways that chronic doses could differ across species than acute doses (such as differences in repair functions or in excretion rates). These factors could result in greater total variance in the distributions of dose ratios for chronic doses than for acute doses.
3. *The toxicological endpoints in the various studies vary with species and can vary across compounds.* The MTD database is composed of the results of studies performed using different study designs, protocols, and endpoints. Thus, there is uncertainty as to whether the differences in the toxicity ratios are due to the species' responses or to differences in the endpoints examined. Specific concerns include the use of acute (single) doses, the comparability of doses that cause lethality to doses that cause low level toxicity or weight change, and the differences in the levels of lethality allowed (Rhomberg and Wolff, 1998; Rhomberg and Lewandowski, 2004c). However, it should be noted that these toxicity benchmarks in humans and test animals have historically been considered comparable to doses that cause frank effects, but with a limited potential for lethality (Goldsmith et al., 1975; Grieshaber and Marsoni, 1986; Travis and White, 1988).
4. *The compounds were administered via injection to both humans and test animals and therefore do not reflect certain aspects of interspecies variation.* Absorption from the GI tract and initial metabolism in the liver will not be reflected in the data, potentially biasing the toxic dose low, particularly for direct-acting agents that undergo detoxification in the liver. Because of the absence of these factors, the variance of MTDs is expected to be less than the variance in NOAELs from oral studies.

Table 1
Data on the toxicity[a] of select anti-neoplastic agents

Source and chemical	Toxic dose (mg/kg)				
	Human	Mouse	Rat	Monkey	Dog
Freireich et al., 1966[b]					
Amethopterin	0.41	3.2	0.58	3.0	0.12
6-Mercaptopurine	27	86	51	56	14
5-Flourouracil	15	42	25	18	10
5-Fluoro-2′-deoxyuridine	30	160	89	60	40
Nitrogen mustard	0.20	1.2	0.37	0.20	0.48
Nitromin	2.0	45	7	4.8	4.4
L-Phenylalanine mustard	0.20	5.1	2.3	0.55	0.63
Alaninemustard	0.90	6.3		1.5	1.5
Cytoxan	10	93	12	52	12
ThioTEPA	0.20	5.7	2.7	1.0	1
Myleran	0.64	15	3.7	6.0	5.8
Actinomcycin D	0.02	0.07	0.09		0.03
Mitomycin C	0.20	2.3	1.3	0.64	
Vinblastin	0.08	0.6			
Vinchristine	0.024	0.18			
Methyl GAG	11	59.00			
Schein et al., 1979; Travis and White, 1988[b]					
Mithramycin	0.025	0.16		0.12	0.12
9H-purine,6-(methylthio)-9-B-D-ribofuranosyl-,dehydrate	5.0	46	8.0	11	8.0
Imidazole mustard	10	230		170	14
Ammonium, trimethylpurin-6-yl-chloride	42	150		170	45
Pactamycin	0.45	3.1		0.09	0.11
Glycinen-(diazoacetyl)-,hydra-zide	160	400		96	48
Tylocrebrine	1.9	19		1.2	0.60
Acetophenone	550	700		1500	490
Cytosine, 1-B-d,arabinofurano-syl, monohydrochloride	7.0	130		36	18
Hydrazine, 1-acetyl-2-picolinoyl	30	87	61	120	58
Phosphorodiamidic acid, *N,N*,-bis (2-chloroethyl), with cyclohexylamine (1:1)	2.7	65	53	5.6	11
Urea, 1,3-bis (2-chloroethyl)-1-nitroso	1.5	9.6	42	3.8	3.8
Greishaber and Marsoni, 1986[c]					
Carboplatin	2.7	40			3.2
Teroxinone	11	27			8.3
Homoharringtonine	0.14	1.9			0.15
Fludarabine	1.1	410			110
Triciribinephosphate	1.5				8.5
N-Methylformamide	32	420			68
DHAC	68	320			12
Rozencweig et al., 1981[d]					
Anguidine	0.12	9.6			0.03
Piperazinedione	0.05	1.9			0.04
Deazuridine	32	190			64
Gallium nitrate	3.8	10			1.7
Bakers antifol	2.7	12			1.0
PALA	27	220			25
Thalicarpine	6.0	41			3.3
Maytansine	0.01	0.08			0.0031
Chlorozotocin	1.1	4.9			0.15
Paxton et al., 1990[e]					
Amsacrine	0.62	6.92			0.32
CI-921	2.8	8.4	6.00		0.32
Goldsmith et al., 1975[f]					
3-(2-Chloroethyl-2-(2-chloroethylamino)) tetrahydro-2H-1, 3, 2-oxazaphosphorine-2-oxide	27				3.2
3,3′-(Imino) di-1-propanol dimethane sulfonate (ester), hydrochloride	2.0	41		4.7	1.3
Cytosine arabinoside[g]	7.1			38	19
4-Amino-1-B-d-ribofuranosyl cytosine hydrochloride	6.8	12			0.52
Tubercidin[h]	0.50				6.4
2-Amino-9-(2-deoxy-(B-D-erythropentofuranosyl)-9H-purine-6-thiol, hydrate	8.1			0.51	0.77
2-Deoxy-2-(3-methyl-3-nitrosoureido) D-glucopyranose	14				3.3

Table 1 (*continued*)

Source and chemical	Toxic dose (mg/kg)				
	Human	Mouse	Rat	Monkey	Dog
Goldsmith et al., 1975[f]					
4-Amino-7-B-D-ribofuranosyl-7H-pyrrolo(2, 3-d) pyrimidine	0.95	8.7			0.54
Bleomycin[i]	0.16	12.8		0.11	0.27
3-Acetyl-1, 2, 3, 4, 6, 11-hexahydro-3, 5, 12-trihydroxy-7 (or 10)-methoxy-6,11-dioxo-1-naphthacenyl-3-amino-2, 3, 6-trideoxy-2-a-l-lyxo-hexopyranoside hydrochloride	1.2	1.4		0.73	1.2
5-(3,3 Dimethyl-1-triazeno)-imidazole-4-carboxamide	6.8	100		100	15
4,5-Dicarboxytetrahydro-6-methylene-pyran-2-succinic anhydride, dianhydride, 2,6B polymer	34			23	12
2′-(9,10-Anthrylenedimethylene) bis-2-pseudothiourea dihydrochloride dihydrate	4.1	90		16	14
3-Ethyl-1,3,4,6,7,11b-hexahydro-9, 10-dimethoxy-2-[(1,2,3,4-tetrahydro-6, 7-dimethoxy-1-isoquinoly) methyl] -2H-benzo[a] quinolizine, dihydrochloride	0.30	7.4		0.87	
2-Ethyl-9,11-dihydro-8-(hydroxymethyl)-9-oxo-indolizino [1,2-b] quinoline 7-glycolic acid, sodium salt	0.54	4.0		1.3	0.45

[a] MTD and other doses, normalized to a 5-day dosing schedule. When necessary, doses were converted to mg/kg from mg/m^2 as follows: (dose in mg/kg)=(dose in mg/m^2)/*k*; where *k* is a species-specific conversion factor (human)=37; (mouse)=3.0, (rat)=5.2, monkey=11.5, and; dog=19.4.
[b] Human dose=MTD, Mouse=$LI0_{10}$, Rat=LD_{10}, Monkey=MTD, Dog=MTD.
[c] Human dose=MTD, Mouse=LD_{10}, Dog=LD_{10}.
[d] Human dose=MTD, Mouse=LD_{10}, Dog=TDL. Dosage converted from a single administration.
[e] Human, rat, and mouse doses=MTD. Dosage converted from a single administration.
[f] Human dose=MTD, Mouse=LD10, Dog=TDL.
[g] Dosage converted from a 10-day schedule.
[h] Dosage converted from a 21-day schedule.
[i] Dosage converted from a bi-weekly schedule.

5. *The majority of the compounds are direct-acting compounds known to have significant biological activity in rapidly dividing cells and may not be fully relevant to the universe of general chemicals.* RfDs are established for a variety of substances including essential nutrients, industrial chemicals, pesticides, metals, inorganics and organics. The substances with existing RfDs, or potentially requiring RfDs, include compounds that are direct acting, compounds that require metabolic activation, compounds with receptor-mediated effects, and compounds that act through more general mechanisms. The impact of this greater heterogeneity may result in greater variance than the MTD ratios.
6. *Human MTD data are drawn from individuals with advanced stages of cancer, and as such may overestimate the toxicity of the compounds for typical humans.* The stress from cancer may reduce the ability of the patients to tolerate the effects of the agents. This would tend to lower the values of the human MTD and raise the value of the ratios. In addition, the condition of the patients could vary from one study to another and may contribute to the variance in the data.
7. *Ratios of MTDs (and other dose metrics) reflect both variation in toxicity between the species and uncertainty in the measurement of the MTD. Thus the distribution of MTD ratios will tend to have greater variance than the true animal to human variability.* The uncertainty in the toxicological measurements used to define the ratios has been discussed by Brand et al. (1999, 2001) and Slob and Peiters

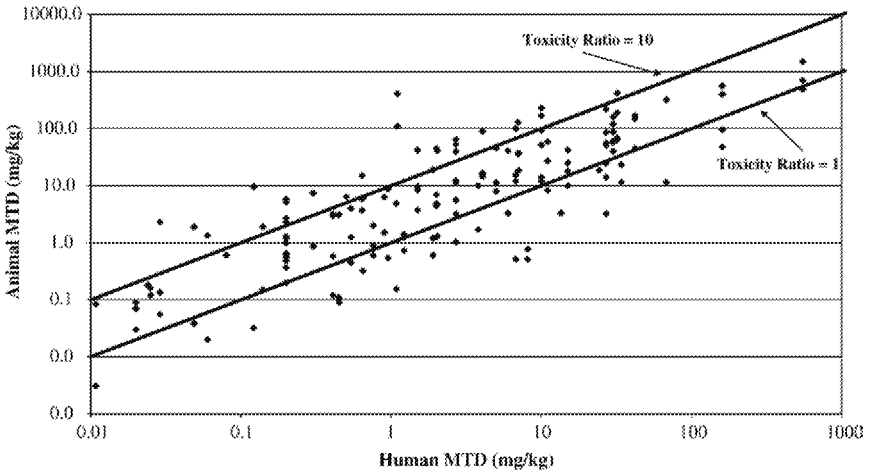

Fig. 1. Scatter plot of animal and human MTDs.

(1998). This uncertainty comes from the impact of dose spacing, background rates of adverse effects, and the finite number of animals used in the studies. The effect of this measurement uncertainty is to inflate the variation in the empirical ratios. While the published work has focused on NOAELs and benchmark doses, similar inflation would be expected in the measures of the MTD. In fact, the inflation of the interspecies differences may be more pronounced for the MTD ratios since MTD studies are conducted on smaller numbers of test animals (or humans) than are studies that have generated NOAELs.

Because of these differences, it is not clear that the empirical distributions can be applied directly as measures of uncertainty in the differences between animal and human NOAELs. This issue is addressed in more detail in the discussion section of this paper.

Analyses of data

Fig. 1 presents a scatter plot of the 161 data points. The plot also includes two lines, one corresponding to equivalent human and animal toxicity and the second to a 10-fold higher sensitivity in humans.

Development of species-specific distributions

Toxicity ratios are likely to differ as a function of the species of the test animal (Travis and White, 1988). For this reason it is important to characterize inter-chemical variation in the ratios as a function of the species of the test animal. This can be done by developing species-specific distributions or by seeking to minimize species differences by using different dose metrics such as body weight$^{3/4}$ or surface area (Travis and White, 1988; Watanabe et al., 1992; Baird et al., 1996; Schneider et al., 2004; Bokkers and Slob, 2007). In this paper, the first approach of developing species-specific distributions is used, as this approach is the one most closely related to application of uncertainty factors in the U.S. EPA RfD methodology (Barnes and Dourson, 1988; USEPA, 2007). The toxicity ratios were sorted by test animal and ranked in order of increasing value. The resulting cumulative distributions are presented in Fig. 2. Summary statistics for the distributions are given in Table 2.

The species-specific distributions were tested to determine if they differed statistically from one another. This was determined using 90% confidence limits on the medians and 90th percentiles of the distributions. The confidence limits were determined using the non-parametric bootstrap approach described below.

Evaluation of the impact of data in multiple species

Where data are available for multiple species the current practice is to use the most sensitive species (on a body weight basis) to establish an RfD (Barnes and Dourson, 1988; USEPA, 1988). Such a practice should have the effect of minimizing the chance of underestimating the toxicity of a compound in humans, as species with greater sensitivity to a compound

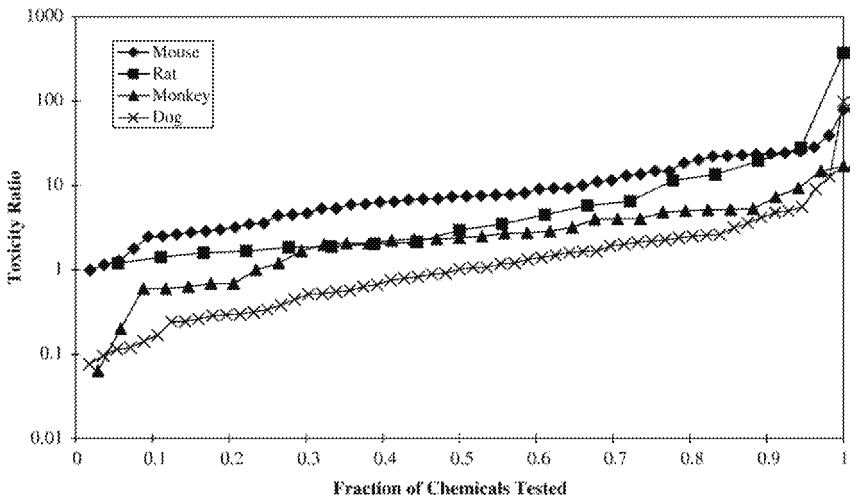

Fig. 2. Distribution of toxicity ratios in test animals and humans for four test species.

Table 2
Summary statistics for species-specific distributions of toxicity ratios

	Mouse/human	Rat/human	Monkey/human	Dog/human
Number of agents	54	17	34	56
Median (CI)	7.7 (6.8–9.3)	3.0 (1.9–5.8)	2.5 (2.1–3.3)	1.0 (0.7–1.5)
Mean	20	6.5	3.6	3.5
90th percentile (CI)	25 (23–63)	16 (6.5–28)	6.7 (5.0–14)	4.4 (2.6–7.7)
Standard deviation	51	7.6	3.7	13

would be used to establish the RfD. The impact of this practice is examined by plotting the ratio of the lowest test-animal dose to the human MTD in the subset of agents (14 compounds) that have toxicity data for all four species. Fig. 3 presents the distributions of ratios of human MTDs to the lowest MTD for specific combinations of test animals. The combinations consisted of: the mouse dose; the lower of the doses for mouse or rat; the lowest of the doses for mouse, rat, and monkey; and, the lowest of the doses for mouse, rat, monkey, and dog. These combinations were selected based on increasing body weight of the test animals.

Evaluating the effect of the limited number of compounds on the distribution

As discussed above, the database for anti-neoplastic agents is relatively small and varies with the test species. An empirical distribution of a relatively small number of chemicals, as in the database for anti-neoplastic agents, is likely to underestimate the true range of the distribution, resulting in potentially significant uncertainty in the estimate of the values of specific percentiles of the distribution. As recommended by Brand et al. (2001), a bootstrap analysis was performed to determine the confidence limits for the means and 90th percentiles (Efron and Tibshirani, 1993). Using unbounded parametric distributions (such as normal or lognormal distributions) has the advantage in that they allow the sampling of plausible values that occur outside of the range of the observed data. Sampling the raw data will limit the values that can be selected and will result in an underestimation of the uncertainty at the extreme tails of the distribution. However, several of the species-specific data sets showed significant deviation from parametric fits (e.g., lognormal). Therefore, the use of unbounded parametric sampling is not appropriate. Sampling in this analysis is performed using an empirical cumulative distribution that allows sampling of values between the raw data points but is still bounded by the range of reported data. Because of this limitation, no estimates were made for the confidence limits of the highest portions of the distributions (>90th percentile).

For this analysis, 10,000 bootstrap samples were generated. In this study, we found that taking 10,000 samples generated stable estimates (within 2%) of the uncertainty in the 90% confidence limits for the medians and the 90th percentiles.

Results

The data set spans a wide range of toxicity in both test animals and humans (Fig. 1). In general, the toxicity ratios are greater than 1.0 suggesting that the healthy test animals are, on the whole, less sensitive than the human cancer patients when dose is expressed on a body weight basis. In addition, a sizable fraction of the data points (29 of 161) had ratios greater than 10. The size of the ratios varies by species and is larger in species with smaller body weights (Fig. 2). As is displayed in Table 2, the median values for the ratios vary from 1 to 8, or approximately one order of magnitude. The results of the bootstrap analysis provide confidence limits to the estimate of the median and 90th percentile (Table 2). These confidence limits indicate that the medians and 90th percentiles of the distributions of ratios for several of the species overlap and thus some of the species differences in this data set are not statistically significant.

All species had ratios that exceeded 10. The fractions of the compounds with values above 10 were approximately 3%, 5%, 19%, and 37% for the dog, monkey, rat, and mouse. Having data on more than one species tended to reduce the sizes of ratios in the set of compounds with data on all four test species. In the subset of 14 drugs that had data on all four species, use of the lowest MTD from both rats and mice versus the lowest mouse MTD results a reduction of the fraction of compounds with values greater than 10 from 25% to 15%. Data from rats or mice when combined with either the dog or monkey data eliminated values of ratios above 10. The median values of the distributions were reduced by up to a factor of five (Fig. 3).

The results from the bootstrap analysis in Table 2 indicate that there is considerable uncertainty in the estimate of the median and the 90th percentiles of the species-specific distributions. Estimates of the 90% confidence interval of the 90th percentile of each species' ratios show that the number could vary by a factor of 2.7 to 4.3. The uncertainties in the median values are slightly smaller, with the differences in the 90% confidence intervals varying by a factor of 1.4 to 3.1.

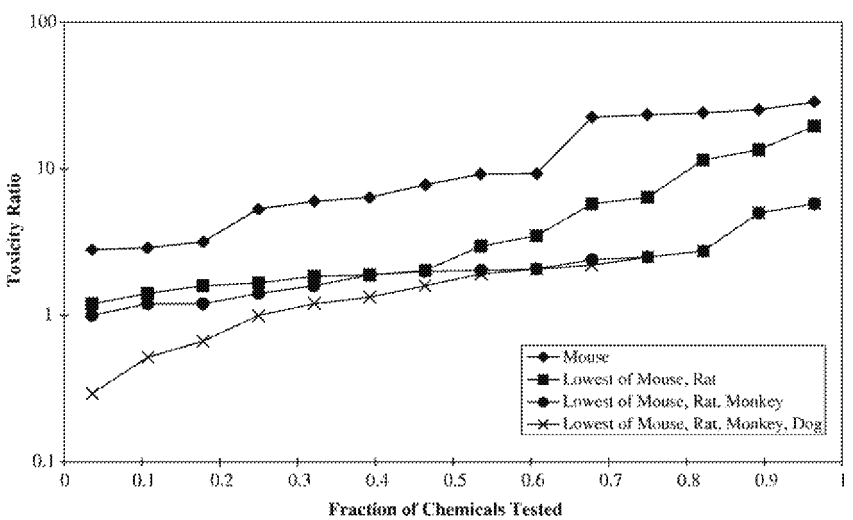

Fig. 3. Distribution of toxicity ratios of humans and the most sensitive species in a battery of test species.

Discussion

The anti-neoplastic data, matched in humans and test animals, provide a unique opportunity to investigate the relationship between the toxicity of this set of chemicals in test animals and human cancer patients. However, as discussed above, there are a number of limitations to the data set.

The issues raised by the limitations can be organized in terms of the impact on likely relationships between the empirical distributions and the uncertainty in the differences between test animals and human toxicity for a chemical. The data set addresses an endpoint that is associated with frank effects rather than NOAELs or LOAELs. Thus, these data are not necessarily representative of the animal to human differences in equivalent-effect doses in the region of interest at the lower end of the dose–response continuum. Similar to ED_{50}-based ratios, MTD-based ratios will tend to underestimate the desired ratios (at low dose) to the extent that human response variance exceeds that for test species. On the other hand, to the extent that the human cancer patients in these studies represent a sensitive subpopulation (i.e., older and in poor health), the ratios will tend to be overestimated. Indeed, this limitation of our study has important implications for the selection of an uncertainty factor distribution for interindividual variation as the human subjects were almost certainly compromised. The differences in the severity of the endpoints in the different species may also affect the means of the distributions. None of these potential biases can be quantified at this time.

Three characteristics of the data set will tend to minimize the variation in ratios. These are 1) the use of acute rather than chronic toxicity data, 2) parenteral route of administration (intravenous or intraperitoneal) and 3) the use of agents that in general do not require metabolic activation. Future work should seek to determine the magnitude of additional uncertainty that these factors cause in interspecies extrapolation.

Finally, as discussed by Brand et al. (1999, 2001), measures of ratios of empirical toxicity reflect both the true differences between the test animal and humans, and the experimental uncertainty in the measurements of toxicity. Brand correctly notes that NOAELs and benchmark doses are inherently "messy" measurements, subjected to a number of significant and poorly characterized sources of uncertainty. In addition, limited sample sizes and the large range of values result in wide confidence limits. While Brand did not investigate uncertainty in the MTDs, the same issues that affect NOAELs and benchmark doses affect the MTD. Brand found that this measurement of uncertainty inflated the observed ratios. If this is the case, then the above distributions may overestimate the upper bound (90th percentile) values of the distributions. Additional work in this area is needed.

Despite these limitations in the data, a number of findings are supported by these analyses. First, the ratios of the MTD are generally consistent with traditional assumptions concerning interspecies variation in toxicity for direct-acting compounds. All of the test animals evaluated tend to have toxicity ratios of 1 or more, which supports the use of values greater than 1 for the interspecies uncertainty factor.

Second, ratios generally are larger for smaller animals. Some of this difference may be a reflection of differences in toxicity quantification, that is, impacts in small animals were measured as lethality (LD10s) and large animal toxicity was determined using less-than-lethal endpoints (TDLs). Nonetheless, it appears from this analysis that body weights are an important factor in accounting for interspecies adjustments and that the differences in sensitivity drop coincidentally with converging weights.

Third, the effect of species is not large in comparison to inter-chemical variation. The difference in the median and 90th percentile across species is less than 8-fold. In contrast, the inter-chemical variations in the toxicity ratios of each species' range of data exceed two orders of magnitude. Fourth, up to 37% of the agents tested in a given species had ratio values greater than 10. This finding may be influenced by the contribution of measurement uncertainty in the MTDs. Fifth, where data are available in multiple species, the probability of large values for a ratio is decreased. For example, the 90th percentile of the distribution for the mouse is 5 times higher than the 90th percentile of the distribution of ratios for the lowest endpoint of the mouse, rat, dog, and monkey. This suggests that a lower interspecies extrapolation factor is appropriate for compounds with relevant toxicity data in multiple species than would be indicated by the species-specific MTD ratio distribution alone. Finally, the bootstrapping analysis shows the impact of limited sample size on the uncertainty in the estimates of the median and 90th percentile of the distribution. The difference between the 90% confidence limits ranges up to a factor of 4.3. Increasing the number of agents would provide a significant improvement in the characterization of interspecies variation.

Conclusion

The distributions developed in this paper are a relevant source of information on the uncertainty in the differences in test animal models and humans and provide a starting point for future research in the relationship between animal and human toxicity for probabilistic noncancer models. Although limited by a number of factors, the empirical data set on toxicity ratios in anti-neoplastic agents is one of the largest in the published literature for the characterization of interspecies uncertainty. The work in this paper establishes the importance of considering the total number of species tested when evaluating the uncertainty in extrapolating findings from animal models to humans. Finally, this work demonstrates the value of additional efforts to collect and evaluate data on anti-neoplastic agents for improving probabilistic non-cancer risk assessments.

Conflict of interest disclosure statement

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper was initially developed as part of the cooperative research and development agreement (CRADA) between McLaren/Hart Environmental Engineering Corporation and the U.S. Environmental Protection Agency under the U.S. Federal Technology Transfer Act of 1996. The paper was revised with additional support from General Electric and The Dow Chemical Company. The views expressed in this paper are of the authors and do not necessarily reflect the views and policies of the U.S. Environmental Protection Agency.

References

Andersen, M.E., Clewell, H., Krishnan, K., 1995. Tissue dosimetry, pharmacokinetic modeling, and interspecies scaling factors. Risk Anal. 15 (4), 533–553.

ATSDR, 1996. Minimal risk levels for priority substances and guidance for derivation 61 Fed. Regist. 101, 25873–25882.

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. Hum. Ecol. Risk Assess. 2, 79–102.

Barnes, D.G., Dourson, M.L., 1988. Reference dose (RfD): description and use in health risk assessment. Regul. Toxicol. Pharmacol. 8, 471–486.

Bokkers, B.G.H., Slob, W., 2007. Deriving a data-based interspecies assessment factor using the NOAEL and the benchmark dose approach. Crit. Rev. Toxicol. 37, 355–373.

Bosgra, S., Bos, P.M.J., Vermeire, T.G., Luit, R.J., Slob, W., 2005. Probabilistic risk characterization: an example with di(2-ethylhexyl) phthalate. Reg. Toxicol. Pharmacol. 43, 104–113.

Brand, K.P., Catalano, P.J., Hammitt, J.K., Rhomberg, L., Evans, J.S., 2001. Limitations in empirical extrapolation studies the case of the BMD ratios. Risk Anal. 21 (4) pages.

Brand, K.P., Rhomberg, L., Evans, J.S., 1999. Estimating non cancer uncertainty factors: are ratios of NOAELs informative. Risk Anal. 19 (2) pages.

Carlson-Lynch, H., Price, P.S., Swartout, J.C., Dourson, M.L., Keenan, R.E., 1999. Application of quantitative information on the uncertainty in the RfD to noncarcinogenic risk assessments. HERA 5 (3), 527–547.

Dourson, M.L., 1994. Methodology for establishing oral reference doses (RfDs). In: Mertz, W., Abernathy, C.O., Olin, S.S. (Eds.), Risk Assessment of Essential Elements. ILSI Press, Washington, D.C.
Dourson, M.L., Stara, J.F., 1983. Regulatory history and experimental support of uncertainty (safety) factors. Regul. Toxicol. Pharmacol. 3, 224–238.
Efron, B., Tibshirani, R.J., 1993. An introduction to the bootstrap. Monographs on Statistics and Applied Probability, 57. Chapman & Hall, New York.
EFSA, 2007. Scientific Colloquium Summary Report 21–22 March 2006 – Parma, Italy, Development of Food-Based Dietary Guidelines European Food Safety Authority.
European Union, 2007. Registration Evaluation Authorization of Chemicals (REACh).
Evans, J.S., Rhomberg, L., Williams, P.L., Wilson, A., Baird, S.J., 2001. Reproductive and developmental risks from ethylene oxide: a probabilistic characterization of possible regulatory thresholds. Risk Anal. 21, 697–717.
Freireich, E.J., Gehan, E.A., Rall, D.P., Schmidt, L.H., Skipper, H.E., 1966. Quantitative comparison of toxicity of anti-cancer agents in mouse, rat, hamster, dog, monkey, and man. Cancer Chemotherapy Rep. 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 Res. 35, 1354–1364.
Grieshaber, C.K., Marsoni, S., 1986. Relation of pre-clinical toxicology to findings in early clinical trials. Cancer Treatment Rep. 70, 65–72.
Kalberlah, F., Föst, U., Schneider, K., 2002. Time extrapolation and interspecies extrapolation for locally acting substances in case of limited toxicological data. Ann. Occup. Hyg. 46 (2), 175–185.
Kalberlah, F., Schneider, K., Schuhmacher-Wolz, U., 2003. Uncertainty in toxicological risk assessment for non-carcinogenic health effects. Reg. Toxicol. Pharmacol. 37, 92–104.
Kodell, R.L., Chen, J.J., 2007. On the use of hierarchical probabilistic models for characterizing and managing uncertainty in risk/safety assessment. Risk Analysis 27 (2).
Nair, R.S., Sherman, J.H., Stevens, M.W., Johannsen, F.R., 1995. Selecting a more realistic uncertainty factor: reducing compounding effects of multiple uncertainties. Hum. Ecol. Risk Assess. 1, 576–589.
Nauman, B.D., Weideman, P.A., 1995. Scientific basis for uncertainty factors used to establish occupational exposure limits for pharmaceutical active ingredients. Hum. Ecol. Risk Assess. 1, 590–613.
Nessel, C.S., Lewis, S.C., Stauber, K.L., Adgate, J.L., 1995. Subchronic to chronic exposure extrapolation: toxicologic evidence for a reduced uncertainty factor. Hum. Ecol. Risk Assess. 1, 516–526.
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 Res. 50, 2692–2697.
Price, P.S., Keenan, R.E., Swartout, J.C., Gillis, C.A., Carlson-Lynch, H., Dourson, M.L., 1997. An approach for modeling noncancer dose responses with an emphasis on uncertainty. Risk Anal. 17, 427–437.
Price, P.S., Keenan, R.E., Schwab, B.W., 1999. Defining the interindividual (intraspecies) uncertainty factor. Hum. Ecol. Risk Assess. 5, 1023–1033.
Rhomberg, L.R., Wolff, S.K., 1998. Empirical scaling of single oral lethal doses across mammalian species based on a large database. Risk Anal. 18 (6), 741–753.
Rhomberg, L.R., Lewandowski, T.A., 2004. Methods for identifying a default cross-species scaling factor. Gradient Corporation, Prepared under Contract 3W-0477-NASX for: Risk Assessment Forum U.S. Environmental Protection Agency 1200 Pennsylvania Avenue, N.W. Washington, D.C. 20460 April 2.
Rozensweig, 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 Clin. 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. Clin. Pharmacol. Therap. 11, 3–40.
Schneider, K., Oltmanns, J., Hassauer, M., 2004. Allometric principles for interspecies extrapolation in toxicological risk assessment – empirical investigations. Reg. Toxicol. Pharmacol. 39, 334–347.
Slob, W., Pieters, M.N., 1997. A Probabilistic Approach for Deriving Acceptable Human Intake Limits and Human Health Risks from Toxicological Studies: General Framework. National Institute of Public Health and the Environment, The Netherlands. Report no. 620110 005.
Slob, W., Pieters, M.N., 1998. A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: general framework. Risk Anal. 18 (6) pages.
Storer, B.E., 1989. Design and analysis of phase I clinical trials. Biometrics 45 (3), 25–37.
Swartout, J.C., Price, P.S., Dourson, M.L., Keenan, R.E., 1998. A probabilistic framework for the reference dose. Risk Anal. 18 (3), 271–282.
Travis, C.C., White, R.K., 1988. Interspecies scaling of toxicity data. Risk Anal. 8, 119–125.
U.S. Environmental Protection Agency (USEPA), 1988. Integrated Risk Information System (IRIS); Background Document 1. National Center for Environmental Assessment, Cincinnati Office, Cincinnati, Ohio.
U.S. Environmental Protection Agency (USEPA), 2007. Integrated Risk Information System (IRIS). Office of Research and Development, National Center for Environmental Assessment, Washington, DC. Online http://www.epa.gov/iris.
Van der Voet, H., Slob, W., 2007. Integration of probabilistic exposure assessment and probabilistic hazard characterization. Risk Anal. 27 (2) pages.
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(5), pp. 439–490.
Vermeire, T., Stevenson, H., Pieters, M.N., Rennen, M., Slob, W., Hakkert, B.C., 2001. Probabilistic assessment factors for human health risk assessment: a practical guide. National Institute of Public Health and the Environment (RIVM), RIVM Report 601516 005, TNO report V3489.
Watanabe, K., Bois, F.Y., Zeise, L., 1992. Interspecies extrapolation: a reexamination of acute toxicity data. Risk Anal. 12, 301–310.
Weil, C.S., 1972. Guidelines for experiments to predict the degree of safety of a material for man. Toxicol. Appl. Pharmacol. 21, 194–199.
Weil, C.S., McCollister, D.D., 1963. Relationship between short- and long-term feeding studies in designing an effective toxicity test. Agric. Fd. Chem. 11, 486–491.