CONTRASTING TOXICOLOGY AND EPIDEMIOLOGY ON "SAFE" LEVELS OF PERSISTENT ORGANIC POLLUTANTS - AN UPDATE

Tom Muir[1*] and Joel E Michalek[2]

[1]Environment Canada – Retired. 70 Townsend Ave, Burlington, Ontario, Canada; [2]Department of Epidemiology and Biostatistics, University of Texas Health Science Center at San Antonio, 314.9 Admin Building, 7703 Floyd Curl Drive, San Antonio TX.

Introduction

The motivations for this work lie in our observed use of experimental toxicology (both *in vitro and in vivo*) study results for chemicals being compared to human blood concentrations and being declared "safe", one chemical at a time, using simple or raw Margins Of Exposure. We observed that the experimental results were often at uM concentrations, with *in vitro* higher, on average. Furthering our concern was an apparent discordance on the average between these experimental results, and epidemiological studies of humans showing significant effect associations at low nM concentrations. These nM concentrations were often in the same range as the human blood concentrations used to declare "safety". Our view is that this apparent discordance, or discrepancy gap, often of several orders of magnitude, should be accounted for in risk assessment and regulation. Our hypothesis is that there is a general discordance between toxicological results and those of epidemiology. In an initial effort[1,2], we explored this hypothesis quantitatively by assessing reported significant internal dose effect concentrations variance within and between toxicological and epidemiological studies and found toxicology-epidemiology discordance, and *in vitro-in vivo* toxicology discordance. There is an apparent trend of increased significant effect concentrations means, with *in vitro* studies exhibiting the highest concentrations, *in vivo* exhibiting intermediate, and epidemiology the lowest. However, there were notable exceptions to the general findings in the thyroid effects category for toxicology, which warrant further explanation. There is also a notable inconsistency in the lipid weight contrasts with whole weight for in vivo compared to epidemiology. The aim of this paper is to present more detailed data and discussion on these inconsistencies, and to provide limited results of slight revisions to the underlying data.

Methods and Materials

We selected 68 relevant POPs *in vitro* (n= 40) and *in vivo* (n= 28) studies, and 53 epidemiological studies. We made the selection to include studies of BFRs, FRs and POPs with published internal dose potencies and specification of the effect. We expressed the internal dose in a common Molar metric expressed in log base 10. We assessed the statistical significance of variation in reported or minimum internal dose observed to be associated with an effect with study type (in vitro (toxicology), in vivo (toxicology), epidemiology), basis (wet, lipid), and effect category (non-thyroid endocrine (NTE), developmental neurotoxicity (DNT), thyroid). We contrasted toxicology with epidemiology and in vivo toxicology with in vitro toxicology with regard to the mean $\log_{10}$ (Molar) using analyses of variance and, for each contrast, a 95% confidence interval for the mean difference (toxicology mean dose minus the epidemiology mean body burden). We applied the Tukey method to correct multiple pairwise comparisons. All statistical testing was two-sided with a nominal experimentwise significance level of 5%. We used SAS Version 9.2 for Windows (SAS Institute, Cary, NC) throughout.

Results and Discussion

Table 1 shows the sample sizes by basis, study design, and effect category for the all toxin chemical category. We summarized 652 dose measurements in all studies (Lipid weight: Epidemiology 136, in vivo toxicology 29, in vitro toxicology 0, Wet weight: Epidemiology 141, in vivo toxicology 64, in vitro toxicology 282). Here, we used revised data for lipid weight in vivo toxicology (thyroid), adding n=8, and for wet weight in vivo toxicology for DNT, revising effect concentrations, and adding n=4. Epidemiology added n=4.

Table 1. Sample sizes by basis, effect, and study

	Lipid Weight			Wet Weight		
		Toxicology			Toxicology	
Effect Category	Epidemiological	in vivo	in vitro	Epidemiological	in vivo	in vitro
DNT	21	11	0	24	35	66
NTE	42	8	0	32	17	133
Thyroid	73	10	0	85	12	83
Total	136	29	0	141	64	282

The mean dose in epidemiology studies (Table 2) was significantly less than the mean in toxicology studies by basis [lipid weight: in vivo toxicology -6.05±1.01, epidemiology -6.95±0.97, CI (0.5, 1.3), wet weight: in vivo -6.42±0.97, in vitro -6.0±1.22, all toxicology -6.08±1.19, epidemiology -8.83±1.19, CI (2.07, 2.74; 2.58, 3.07; 2.51, 2.98, respectively), with $p<0.001$ for all contrasts. The relation between dose or body burden and study type changed significantly with lipid weight versus wet weight basis ($p<0.001$); in wet weight data, epidemiology exhibited a greater than 2 order of magnitude increase CI (2.07, 2.74), and in lipid weight an approximate 1 order of magnitude increase CI (0.5, 1.3) over in vivo toxicology potency. At the mean, the lipid weight contrast for epidemiology versus in vivo toxicology gained 1.51 orders of magnitude over the wet or whole weight contrast.

One source of the overall lipid weight – wet weight inconsistency may be the different lipid nature of the matrices used to report the internal doses. Not shown here, the *in vivo* matrices lipid weight data comprised 17 out of 29 results in either liver (4-5% lipid), or adipose (80-90% lipid), the two compartments where POPs tend to partition. There were also 9 results using Rainbow Trout muscle lipid of about 2%. The epidemiology lipid weight data comprised 130 out of 136 results in serum/plasma/blood (0.5-0.8% lipid), cord blood (0.2-0.26% lipid), or placenta (0.68% lipid). The higher lipid matrices accumulate more POPs in the whole weight, but the leaner epidemiology matrices have higher lipid normalization factors. As well, lipid extraction and recovery at analysis, and normalization, at low levels of lipid can compound measurement error and confound epidemiology. Furthering this situation, the wet weight and lipid weight samples are not matched, so we are not comparing like with like, differing only in the basis, although there is some overlap. We suggest that the lipid based contrasts have unresolved uncertainty and reliability questions, and recommend that any use in RA utilize the wet or whole weight factors.

Table 2 Contrasts between Toxicological and Epidemiological Studies on mean Log_{10}(Dose or Body Burden in Molar units) [N. mean (SD)]

	Toxicology					
Basis	In vivo	In vitro	All	Epidemiology	p-value	95% CI
Lipid Weight	29	0	29	136		
	-6.05 ±1.01		-6.05±1.01	-6.95±0.97)	<0.001	(0.5, 1.3)
Wet Weight	64	282	346	141		
	-6.42±0.97	-6±1.22	-6.08±1.19	-8.83±1.19)		
	•			•	<0.001	(2.07, 2.74)
		•		•	<0.001	(2.58, 3.07)
			•	•	<0.001	(2.51, 2.98)

Among studies reporting DNT, NTE or Thyroid effects in wet weight (Table 3a), the mean dose was significantly decreased ($p<0.001$ for all contrasts) in epidemiology relative to wet weight *in vivo* and wet

weight *in vitro* toxicology [DNT: in vivo toxicology -6.61±0.71, in vitro -5.26±0.79, epidemiology -8.87±1.15, CI (1.78, 2.74; 3.18, 4.03, respectively), NTE: in vivo toxicology -6.88±0.74, in vitro -5.89±1.06, epidemiology -8.24±1.08, CI (0.78, 1.95; 1.94, 2.76, respectively), Thyroid: in vivo toxicology -5.23±1.05, in vitro -6.76±1.33, epidemiology -9.03±1.18, CI (3.09, 4.52; 1.89, 2.65, respectively)].

An exception to our general trend hypothesis is the in vitro result for thyroid showing a lower effect concentration than the wet weight in vivo (-6.76 versus -5.23), which, at least, reflects the chemical, species, and model type. The in vitro are mostly half maximal results (e.g., IC50). The overall in vitro results for thyroid have sample size n=83, of which n=49 are transthyretin or thyroxine-binding prealbumin (TTR) and n=11 are thyroxine-binding globulin (TBG) competitive binding assays. These assay results represent altered plasma transport of thyroid hormone, and as such are highly upstream biological perturbations. For TTR, hydroxylated, and halogenated phenolic compounds (including TBBPA, TCBPA and triclosan) dominated the sample (OH-PCB, n=7; OH-PBDE, n=17; OH-PFB, n=1; H-phenolics, n=14; PCBs, PBDEs, and PFCs, n=10). For TBG, hydroxylated compounds (OH-PBDE, n=6; OH-PCB, n=1; OH-PFB, n=1; and triclosan, n=1) dominated, with PBDE, n=2, the remainder. Summary statistics for TTR are: mean, -7.23; range -8.59 to -4.44. And for TBG: mean, -6.09; range -7.0 to -4.18.

Table 3 Contrasts between Toxicological and Epidemiological Studies on mean Log_{10}(Dose or Body Burden in Molar units)

a) Wet weight [N, mean±SD)]

Effect	Toxicology in vivo	in vitro	All	Epidemiology	p-value	95% CI
DNT	35	66	101	24		
	-6.61±0.71	-5.26±0.79	-5.73±1	-8.87±1.15)		
	•			•	<0.001	(1.78, 2.74)
		•		•	<0.001	(3.18, 4.03)
			•	•	<0.001	(2.68, 3.6)
NTE	17	133	150	32		
	-6.88±0.74	-5.89±1.06	-6±1.07	-8.24±1.08)		
	•			•	<0.001	(0.78, 1.95)
		•		•	<0.001	(1.94, 2.76)
			•	•	<0.001	(1.83, 2.65)
Thyroid	12	83	95	85		
	-5.23±1.05	-6.76±1.33	-6.57±1.39	-9.03±1.18)		
	•			•	<0.001	(3.09, 4.52)
		•		•	<0.001	(1.89, 2.65)
			•	•	<0.001	(2.08, 2.84)

For the non-TTR/TBG markers, (n=23), the OH-PCB (n=9), OH-PBDE (n=2), and TBBPA/TCBPA (n=4) dominate, with PCB (n=1), PBDE (n=4), and HBCD (n=3) sharing the remainder. Overall, these markers were predominantly associated with significant TH pathway down regulation, or inhibition, or antagonism, through several mechanisms (n=15). The balance show T3/TR inhibition combined with TH induction and agonism (TBBPA/TCBPA, n=2; 4-OH-PCBs, n=4) , or T3 potentiation (TBBPA/PBDE/HBCD, n=2). Summary statistics for these data are: mean -6.09; range -11.3 to -4.0. In contrast, wet weight in vivo results are for two chemicals: TBBPA (H-phenolic) in rat liver (n=7); PBDE-47 in minnow muscle (n=4);

and PBDE-47 in flounder carcass (n=1). There were no identified hydroxylated PBDE metabolites, which appear more potent in the in vitro thyroid, and would be expected to be metabolically active in the in vivo.

Corresponding contrasts in lipid weight (Table 3b) were in the same direction, but were generally smaller, and did not reach significance for studies expressing DNT effects. [DNT: in vivo -5.89±1.12, epidemiology -6.46±1.13, p=0.18, CI (-0.28, 1.43), NTE: in vivo -5.56±0.45, epidemiology -6.79±1.13, p=0.004, CI (0.41, 2.04), Thyroid: in vivo -6.61±1.02, epidemiology -7.18±0.75, p=0.04, CI (0.04, 1.09)]. These decreases in the magnitude and significance of the contrasts between lipid weight and the wet or whole weight, by effect category, mirror the analogous contrasts shown in Table 2 for all effects combined.

Also noteworthy for thyroid, the in vivo lipid weight result showed a lower mean concentration than the wet weight. This too may reflect the species model, and biomarker matrix, and is an example of the variability displayed by the toxicology. As above, the wet weight results (n=12) correspond to TBBPA (liver) in two one-generation rat studies (van der Ven et al, n=5; and Lilienthal et al, n=2), and PBDE-47 in one fathead minnow (muscle), and one flounder (carcass) study (Lema et al, n=4 and Kuipers et al, n=1 respectively). This compares to the lipid weight in vivo results from two studies - one 28-day rat OECD model (liver; n=1) testing HBCD as a technical mixture (van der Ven et al]), and one rainbow trout model (muscle; n=9) testing the three main HBCD diastereoisomers (alpha, beta, and gamma) administered individually, with measurements of each of the three diastereoisomers for each test that showed significant effects (Palace et al, Law et al,). The rat liver effect concentration is -4.17, and the rainbow trout muscle mean is -6.88. The reported effect in each case was thyroid epithelial cell and gland hypertrophy.

Table 3 Contrasts between Toxicological and Epidemiological Studies on mean Log10(Dose or Body Burden in Molar units)

b) Lipid weight [N, mean±SD)]

Effect	in vivo Toxicology	Epidemiology	p-value	95% CI
DNT	11	21		
	-5.89±1.12	-6.46±1.13	0.18	(-0.28, 1.43)
NTE	8	42		
	-5.56±0.45	-6.79±1.13	0.004	(0.41, 2.04)
Thyroid	10	73		
	-6.61±1.02	-7.18±0.75	0.04	(0.04, 1.09)

Overall, in effects category results, *in vitro* data contrasts with epidemiology for DNT by 3 to 4 orders of magnitude and *in vivo* data contrasts with epidemiology for DNT by two to three orders of magnitude. For NTE, the in vivo contrast is about 1 to 2 orders, whereas the in vitro is about 2 to 3 orders. In these cases, the in vivo is a closer comparison. An exception is for thyroid effects, where in vivo contrasts with epidemiology by 3 to 4.5 orders of magnitude, and in vitro contrasts by almost 2 to almost 3 orders of magnitude. The possible variation seen across species and basis become more evident when the three effect categories are stratified and thyroid appears separately for both toxicology model types, and in specific contrast with epidemiology.

References

1. Muir T and Michalek JE. 2011. Proceedings, 31st International Symposium on Halogenated Persistent Organic Pollutants.
2. Muir T and Michalek JE. 2012. Manuscript in Preparation.