Integrating Exposure, Toxicology And Epidemiology: A Prototype For Multi-Model, Translation, And Mixtures Study – An Update On Sample Size And Lipid-Wet Weight Contrasts

Tom Muir[1*] Joel E Michalek[2] and Ray Palmer[3]

[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. [3] Dept. of Family and Community Medicine, University of Texas Health Science Center at San Antonio, 314.9 Admin Building,7703 Floyd Curl Drive San Antonio

Introduction

Many chemicals have been identified as having similar effects on either thyroid and sex steroid homeostasis, or neurotoxicity, particularly showing a differential risk in fetuses and neonates, and may interfere at concentrations far below those traditionally used in regulatory toxicology and screening[1]. This life-stage dependent risk, using low doses relevant to human exposure, needs to be accounted for in experimental toxicology and risk assessment. There is also increasing concern about the effects on health of exposure of combined chemicals present in food, air or water, and calls for public risk assessment of combined human exposure to multiple chemical contaminants. The rationalization is based on emerging themes and topics at the Dioxin conferences, research programs, and in the literature. As well, our previous work has been developing data bases and methods for integrating published data on experimental exposure, mechanistic and animal toxicology, and epidemiology, and using ANOVA methods to estimate empirical uncertainty, and the possibility of predicting dosimetry across these categories statistically. Past conference sessions have been emerging on the similar theme of integrating toxicology and epidemiology, with different metrics and methods across the exposure and life stage. Emerging effects studies are integrating in silico, in vitro, in vivo, and epi. The idea is to reduce in vivo needs and develop predictive models with high throughput or other methods, and linkages. Other work includes using metabolomics, proteomics, systems biology, epi-and genetic markers of susceptibility, and integrative approaches of this data to POPs toxicology. Others are developing "omics chips to cluster putative chemicals, 'omics to look for mechanisms, pollutant effects and combine with epidemiological tools. This line of work is leading into quasi-translational approaches, which can be seen roughly if you look closely. Using the idea of multi-level models, for 3 or 4 levels, there is a kind of integrating and translating across, including to mixtures effect assessment. We assembled a data-base of comparable internal dose and response effect concentration data, from a number of toxicological (*in vitro* and *in vivo*) and epidemiological studies reporting on a range of POPs chemicals, multiple species, and for multiple toxicological responses or endpoints[2,] The aims of this paper are a) show how sample size, and the lipid weight and wet weight basis contrasts, present interpretation issues due to the new inferences contained in the integrated data, not contained in the individual studies alone, and b) to discuss how our integration efforts can fit the session theme of a multi-model, translational, and mixture-competent empirical data base,

Methods and Materials

In real time from 2000 to 2010, we have selected 71 relevant POPs *in vitro* (n= 40) and *in vivo* (n= 31) studies, and 53 epidemiological studies. We made the selection to include studies of BFRs, FRs and POPs with published applied dose amounts and protocols for *in vivo* studies, internal dose potencies and specification of the effect. Animal species included mouse, rat, monkey, sheep/lambs, kestrel, rainbow trout, flounder, and fathead minnow. Further, we added information on life stage of exposure, dose timing, and dose number. For the *in vivo* and epidemiology internal effect dose data we added information on sex and life stage of exposure. We stratified by basis (lipid weight, wet weight), study (*in vivo* toxicology, *in vitro* toxicology, epidemiology), chemical (in 22 categories), and effect (in multiple categories or markers (n=102), aggregated to DNT (n=22), thyroid (n=35), and NTE (n=45) due to sample size constraints). We expressed the internal dose and applied doses 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). In this paper, we added the applied doses, sex, life stage of exposure and dose timing and number

to the analysis. We contrasted with regard to the mean $\log_{10}$ (Molar) using analyses of variance and, for each contrast, a 95% confidence interval for the mean difference. 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 680 internal dose measurements in all studies (Lipid weight: Epidemiology 136, *in vivo* toxicology 41, *in vitro* toxicology 0, Wet weight: Epidemiology 152, *in vivo* toxicology 69, *in vitro* toxicology 282). Additionally, we summarized 234 applied dose measurements from all *in vivo* studies.

Table 1. Toxicological sample sizes by basis, study type, and effect category

	Wet Weight	Lipid Weight				Wet Weight			
Effect Category	Applied Dose	Epidemiology	*in vivo*	*in vitro*	All	Epidemiology	*in vivo*	*in vitro*	All
DNT	75	21	11	0	11	24	35	66	101
NTE	77	42	12	0	12	32	17	133	150
Thyroid	82	73	18	0	18	96	17	83	100
Total	234	136	41	0	41	152	69	282	351

Not shown here, the wet weight contrasts across the dose metrics and effect category contrasts show a consistent increase in the mean dose from epidemiology relative to *in vivo* and *in vitro* (not shown: $p<0.001$). The addition of the AD adds to this increasing trend. Among studies reporting any effect and with regard to wet weight applied and internal doses, the wet weight mean dose was significantly decreased in epidemiology relative to both *in vivo* and *in vitro* toxicology for all three effect categories. In wet weight, the *in vivo* ADs were not significantly different from the *in vitro* internal dose for DNT and NTE, but not Thyroid (DNT p=0.26; 95% CI -0.12 to 0.71: NTE p=0.8; 95% CI -0.87 to 0.03: Thyroid p=0.001; 95% CI -2.29 to -1.0). Table 2 shows lipid weight contrasts across the dose metrics and effect categories. Corresponding contrasts between toxicology and epidemiology in lipid weight internal doses were in the same direction as the wet weight contrasts, were generally smaller, and did not reach significance for studies expressing any effect category (DNT p=0.34; 95% CI -1.55 to 0.4: NTE p=0.14; 95%CI -1.93 to 0.21; Thyroid p=0.14; 95% CI -1.87 to 0.2). In all effect categories, lipid weight *in vivo* internal dose is not significantly different from the AD, which is administered in whole weight (DNT p=0.62; 95% CI -0.51 to 1.17: NTE p=0.54; 95% CI -0.56 to 1.47; thyroid p=0.014 95%CI 0.2 to 2.25). Integrating all lipid wt. results (i.e. Total) without regard to effect category, the epidemiology dose was significantly less than the *in vivo* dose (95% CI -1.45 to -0.24; p=0.003). Previous interpretation of this effect category result suggested that the ADs include doses that are environmentally relevant to humans, and that accumulation in lipids presents this. The result for Total, suggested that pooling data may average down significant effects otherwise seen in relevant stratifications. Reinterpretation suggested the small number of selected lab animal studies with lipid-adjusted internal dose is the reason for the failure of the contrast with the epi internal dose to achieve statistical significance for each of the health effect categories separated. An inference available from integration allows the observation that the contrast between epi and applied dose of in vivo studies shows a statistically significant difference, as does the contrast between epi and lab animal internal dose for the combined health effects. These comparisons are inconsistencies showing the sample size bias to the null in the lipid weight contrasts of individual effect categories.

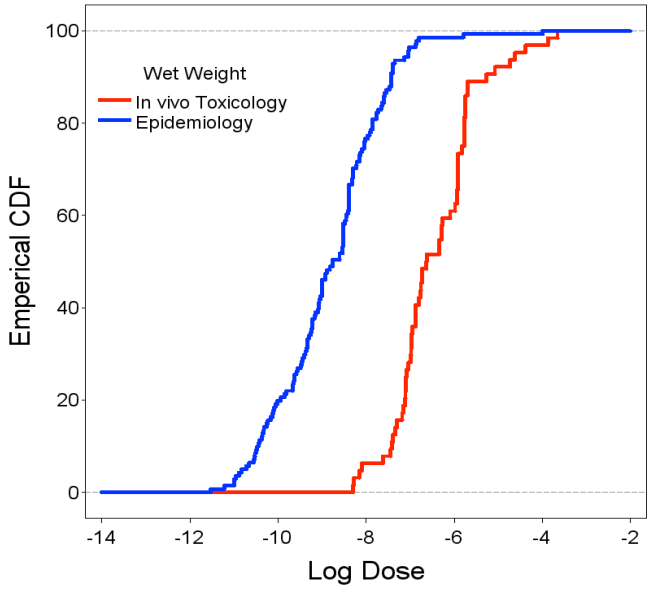
Table 2. In vivo applied dose and internal concentration, in vitro and epidemiology internal dose lipid weight.

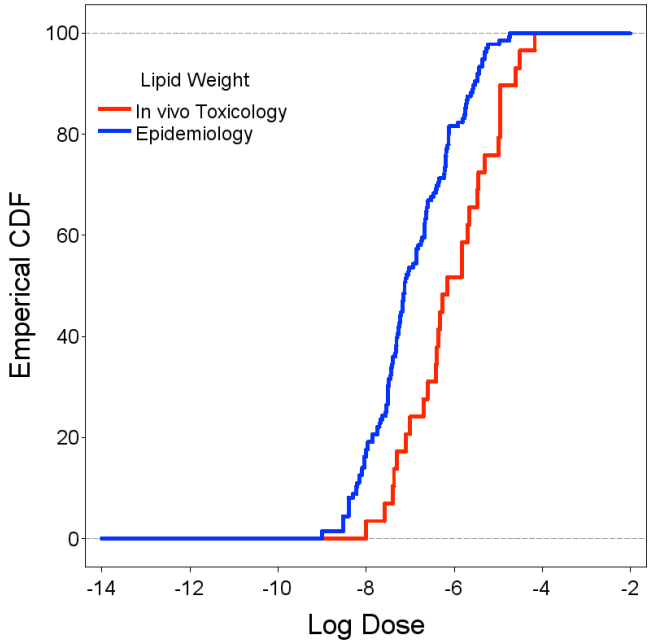
There are other concerns identified in the lipid weight data compared to the whole weight by the integration and ANOVA contrasts. The relation between internal dose concentration and study type changed

Effect		*In vivo applied dose*	*In vivo internal concentration*	*In vitro internal concentration*	*Epidemiology internal conc.*	*p-value*[2]
DNT	N	75	11		21	
	Mean (SD)	-5.56 (1.09)	-5.89 (1.12)	.	-6.46 (1.13)	
		•			•	0.003
			•		•	0.34
		•	•			0.62
NTE	N	77	12		42	
	Mean (SD)	-5.47 (1.56)	-5.93 (0.79)	.	-6.79 (1.13)	
		•			•	<0.001
			•		•	0.14
		•	•			0.54
Thyroid	N	82	18		73	
	Mean (SD)	-5.12 (2.2)	-6.35 (1.42)		-7.18 (0.75)	
		•			•	<0.001
			•		•	0.14
		•	•			0.014
Total	N	234	41		136	
	Mean (SD)	-5.38 (1.7)	-6.1 (1.18)		-6.95 (0.97)	
		•			•	<0.001
			•		•	0.003
		•	•			0.009

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 (1.9, 2.92), and in lipid weight an approximate 1 order of magnitude increase CI (0.24, 1.45) over in vivo toxicology potency. At the mean, the lipid weight contrast for epidemiology versus in vivo toxicology gained 1.56 orders of magnitude over the wet or whole weight contrast. This is another indicator of the movement of the epi mean toward the in vivo mean in lipid weight and may also contribute to the null bias. 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 27 out of 41 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%, and the balance were mostly plasma. 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. We recommend that any use in RA consider what is the more meaningful and biologically accurate measure to study POPs, and extent to which lipid adjustment may bias towards the null [4]. **Figure 1** shows plots of distribution functions of the whole weight versus lipid weight data used in the contrasts of in vivo compared to epidemiology.





1. Diamanti-Kandarakis E, Bourguignon J-P, Giudice L, Hauser R, Prins G, Soto A et al. 2009. *Endocrine Reviews* 30(4): 293-342.
2. Muir T and Michalek J. 2014. *Organohalogen Compounds - Proceedings of Dioxin 2014*
3. Michele La Merrill, Claude Emond*3* Min Ji Kim, Jean-Philippe Antignac, Bruno Le Bizec, Karine Clément, Linda S. Birnbaum, and Robert Barouki. (2013) Environmental Health Perspectives Vol 121, Number 2.