April 29, 2012

CASE STUDY:

Development of Screening Tools for the Interpretation of Chemical Biomonitoring Data[1]

Richard A. Becker[a], Sean M. Hays[b] , Steven Robisonc[c] , Lesa L. Aylward[d] Christopher R. Kirman[e]
[a]American Chemistry Council, Washington, D.C.; [b]Summit Toxicology, LLP, Lyons, CO; [c]Procter & Gamble, Cincinnati, OH; and [d]Summit Toxicology, LLP, Falls Church, VA, [e]Summit Toxicology, LLP, Orange Village, OH

1. Overview

In this case study we describe how the approach for using the Biomonitoring Equivalents (BEs) -- forward dosimetry to interpret human biomonitoring data in a risk context -- has been extended by the development of a framework with decision nodes to enable the BE approach to be applied to substances with varying degrees of toxicity and pharmacokinetic data.

Human biomonitoring – the measurement of chemicals or their metabolites in biological media such as blood or urine -- has become a powerful tool for characterizing chemical exposures in people. Human biomonitoring data provide a reflection of integrated exposure from multiple pathways and routes in terms of an internal, biologically relevant , or absorbed dose. When exposures occur through multiple or ill-defined exposure routes or pathways, well-designed and conducted human biomonitoring studies can provide robust and reliable exposure data that can complement and refine or replace external exposure dos estimates based on more indirect approaches and generic assumptions.

However, biomonitoring results are not clinical diagnostic metrics, and interpretation is challenging. It is only for a handful of environmental chemicals, such as lead and mercury, where there are robust datasets and objective medical findings that can relate biomarker concentrations in human populations to potential adverse health effects. The data to support such assessments does not exist for the majority of environmental chemicals because this approach requires establishment of causality in clinical and epidemiological studies and a robust understanding of human dose response. Thus, an alternative approach, the concept of Biomonitoring Equivalents (BEs) has been developed, and guidelines for the derivation and communication of these values have been published [Hays et al., 2008; LaKind et al., 2008].

The concept of Biomonitoring Equivalents (BEs) was developed to permit interpretation of human biomonitoring results into a health risk context. Because biomarker data are typically expressed in units of biomarker concentration (e.g., ug/L urine), and risk-based benchmarks (such as Reference Doses [RfDs] or Acceptable Daily Doses [ADIs] or Tolerable Daily Doses [TDIs]) are typically expressed in units of applied dose (mg/kg-day), a direct comparison between the two cannot be made. The Biomonitoring Equivalent approach was thus developed.

[1] Adapted from the paper of Becker et al., 2011.

The Biomonitoring Equivalent approach translates the external dose health-based guidance value into the estimated corresponding steady state biomarker concentrations in blood or urine. A Biomonitoring Equivalent (BE) is defined as the concentration or range of concentrations of a chemical or its metabolites in a biological medium (blood, urine, or other medium) that is consistent with an existing health-based exposure guidance value such as a RfD or TDI or ADI [Hays et al., 2008]. BEs are intended to be used as screening tools to provide an assessment of which chemical exposures (biomarkers) are present at levels well below, near, or at or above concentrations that are consistent with existing risk assessments and exposure guidance values, and thus can provide an evaluation of relative priority for risk assessment follow-up.

1.1 Problem Formulation Discussion

In the broadest sense, the problem that needs to be addressed is how to interpret human biomonitoring results to determine whether or not the exposures detected are of a magnitude that raises concerns for health. In absence of interpretation tools to understand human biomonitoring results some may equate the mere detection of a substance to impending illness or injury..

Now, with advanced analytical methods, it is now possible to quantitatively measure 10's to 100's of substances in reasonable sample volumes at the individual level. However, interpreting the results of human biomonitoring is challenging, even for well-designed and executed studies. Without tools to interpret biomonitoring results in a risk context, risk assessors and risk managers cannot distinguish the significance of the exposures, whether a result is of potential health concern or not, or whether the levels detected are so low as to be of no health concern whatsoever or whether the results signify that additional risk management or product stewardship actions may be warranted.

BEs were designed to make maximal use of existing, authoritative (government) health guidance values (such as RfDs, ADIs, TDIs, etc.) that have been established through a transparent, scientifically rigorous process which included independent peer review. BEs derived from such health guidance values, once peer reviewed, can be applied to interpret potential human health risks with virtually the same certainty as the underlying health guidance value.

However, it is now apparent that approaches are needed to aid in the interpretation of human biomonitoring data for substances which lack such extensive and well vetted health guidance values. In addition, there are also cases where although toxicity data may be available, robust toxicokinetic models are lacking and also cases where available data concerning both toxicity and toxicokinetics are scarce.

Thus, the problem statement we sought to address was:

> *How can existing data and knowledge of toxicity and toxicokinetics be integrated to enable human biomonitoring results to be interpreted in a health risk context where available information can range from very complete to very sparse?*
>
> *What would be a consistent and scientifically justified framework and associated decision criteria that could be applied to guide the development of such interpretation tools?*

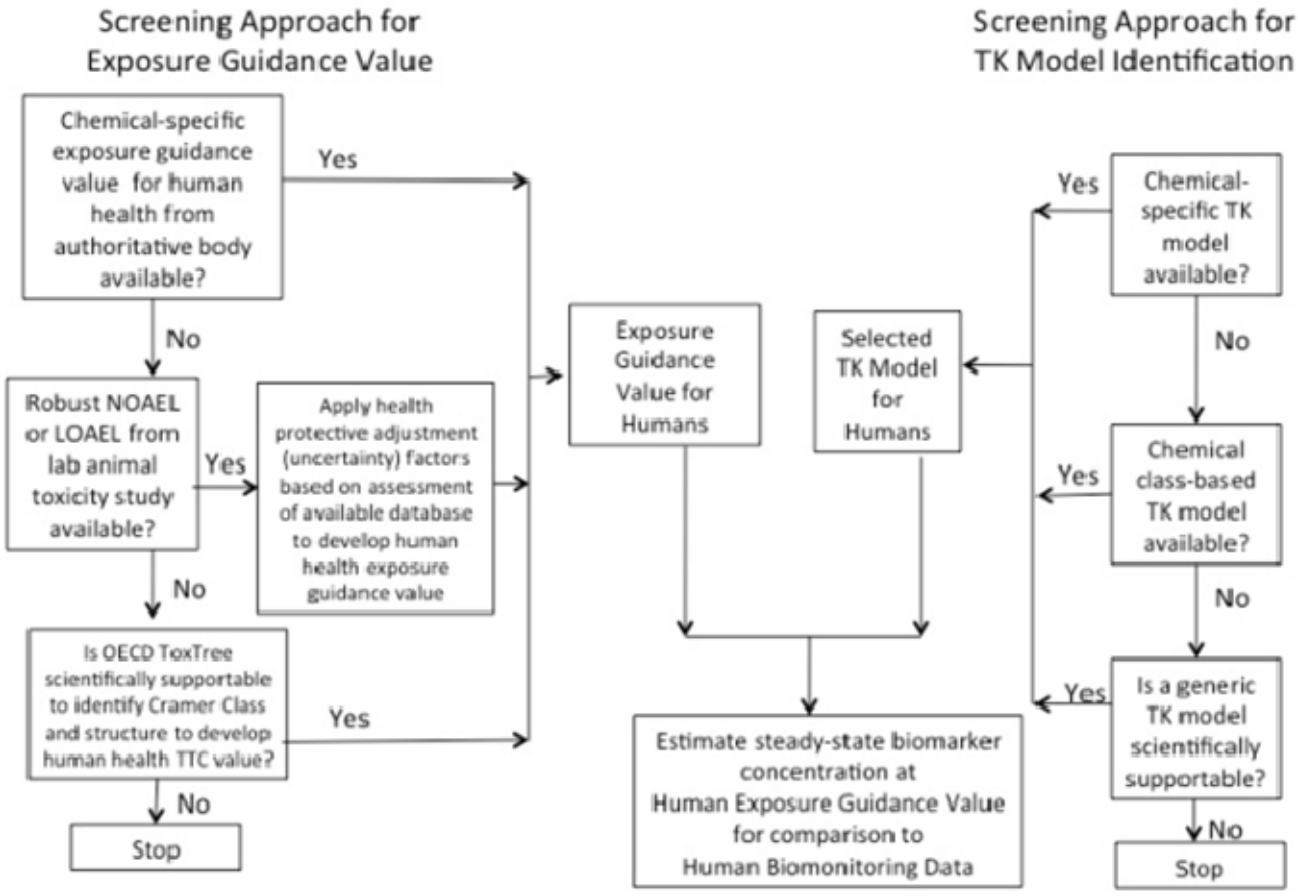
We proceeded to develop a tiered framework and decision tree (Figure 1) to guide the derivation of biomonitoring interpretation screening values which take into consideration the availability of different degrees of toxicity and toxicokinetic data.

1.2 Decision-Tree For Developing Screening Values For Interpreting Biomonitoring Results For Chemicals With Varying Levels Of Available Hazard And Toxicokinetic Data.

Figure 1 provides a general flowchart of the various approaches described here. The flowchart is conceptually similar to the tiered screening process described in a 2001 review by the Health Council of the Netherlands [Health Council of the Netherlands, 2001] with the added component of extension of the tiered approach to evaluation of biomonitoring data. These approaches should be applied in an iterative framework, with increasing refinement indicated when Margin of Safety (MOS) values are judged to be insufficient.

Of course, in a case where all-generic approaches are used to derive a provisional screening value, such a value will be highly uncertain, and would require the use of health-protective assumptions in the screening process. If chemicals being detected in biomonitoring surveys fall into this category of lacking both toxicological and toxicokinetic data, these chemicals may be candidates for early research to fill selected data gaps in order to refine the assessments for those chemicals.

Figure 1. Framework for Developing Screening Values to Interpret Human Biomonitoring Data in a Risk Context (from Becker et al, 2011).



2. Examples

In conventional risk assessment, external doses (mg/kg-day) are estimated by combining concentrations in environmental media with specific contact scenarios. When an external dose RfD or TDI (or analogous screening value such as a threshold of toxicological concern (TTC)) is available, the screening-level exposure estimate is compared directly to an external dose health-based guidance value.

If a NOAEL or other POD is used as the benchmark, then typically default adjustment factors (AFs) (synonymous with uncertainty factors UFs or safety factors) are generally used to extrapolate from animal toxicity to humans (default 10X) and to account for human variability

(default 10X); however chemical specific adjustment factors may be used in lieu of the defaults [IPCS, 2005]. Depending upon the database and quality of studies, additional AFs may be used [TERA, 2005], and if a toxicity database is not robust, use of an additional database deficiency AF should be considered.

Once the screening level health-based guidance value has been determined, then a MOS can be calculated by comparing this to the estimated daily dose rate (D):

$$MOS = \frac{\left(\frac{POD}{AFs}\right)}{D}$$

MOS values below 1 indicate that exposures exceed the screening level health-based exposure guidance value.[2]

With the Biomonitoring Equivalent approach for interpreting biomonitoring results in a risk assessment context, an external dose health-based guidance value is translated into the corresponding steady-state biomarker concentration in blood or urine (e.g., the BE_{RfD} when deriving a BE associated with the RfD). BEs provide a translational tool allowing application of the foundational risk assessment paradigm to the evaluation of exposure information provided by biomonitoring data.

As described previously [Hays et al., 2008], the guidance for development of BEs calls for use of authoritative (government) health-based guidance values (such as an RfD or TDI) as well as sufficient understanding of pharmacokinetics of the chemical in humans or key laboratory species. BEs can be used for interpreting human biomonitoring data in a risk context by calculating the margin of safety (MOS), similar to the conventional risk assessment approach described above:

$$MOS = \frac{BE}{[Biomarker]}$$

[2] If screening approaches have been used in the exposure or hazard assessment process, further refinement in those assessments may be warranted. Such refinements to provide greater certainty of potential hazards and exposures may include generation of product-specific exposure data for chemical uses with higher estimated exposure rates, conducting specific toxicity studies to address database deficiencies, or other exposure or hazard characterization refinements. Results of refined assessments can be used to identify the need for, and useful focus of, potential risk management strategies.

To extend the BE risk-based approach for interpreting human biomonitoring data, we developed a tiered framework and decision tree for approaches to deriving BEs with varying degrees of available data (both toxicokinetics and toxicity) (Figure 1). We illustrate the application of this approach for five scenarios, each with specific examples:

Three approaches are applicable to substances for which toxicokinetics are well understood but that have different levels of toxicity data

(1) substances with established government risk assessments[3];

(2) substances with sufficient toxicity datasets but as of yet no government generated (or vetted) risk assessment; and

(3) substances amenable to the generic screening TTC approach for setting conservative tolerable intake rates.

Approaches (2) and (3) above are needed because for many chemicals in common use today, there may not be authoritative, government conducted, consensus, or "approved" chemical-specific risk assessment-based exposure guidance values available.

The additional two scenarios for deriving addressed by the framework are for instances where:

(4) chemical-specific toxicokinetic data or models are lacking; and

(5) both toxicity-based guidance values and toxicokinetic data are not readily available.

2.1. Biomonitoring Equivalents Based on Substances with Established Government Risk Assessments and Established Toxicokinetics

BE values have been derived for approximately 80 chemicals in a variety of chemical classes [See Angerer et al., 2011]. BE derivations have been published for persistent organic compounds including dioxins, hexachlorobenzene, and DDT and metabolites; for approximately 40 volatile organic compounds; for several phthalates and phenols including di-2(ethylhexyl)phthalate, bisphenol A, and triclosan; for selected pyrethroid pesticides; several metals (arsenic, cadmium, selenium) and for selected brominated flame retardant compounds. For these chemicals, screening level assessments of population biomonitoring data can be made by comparison of the

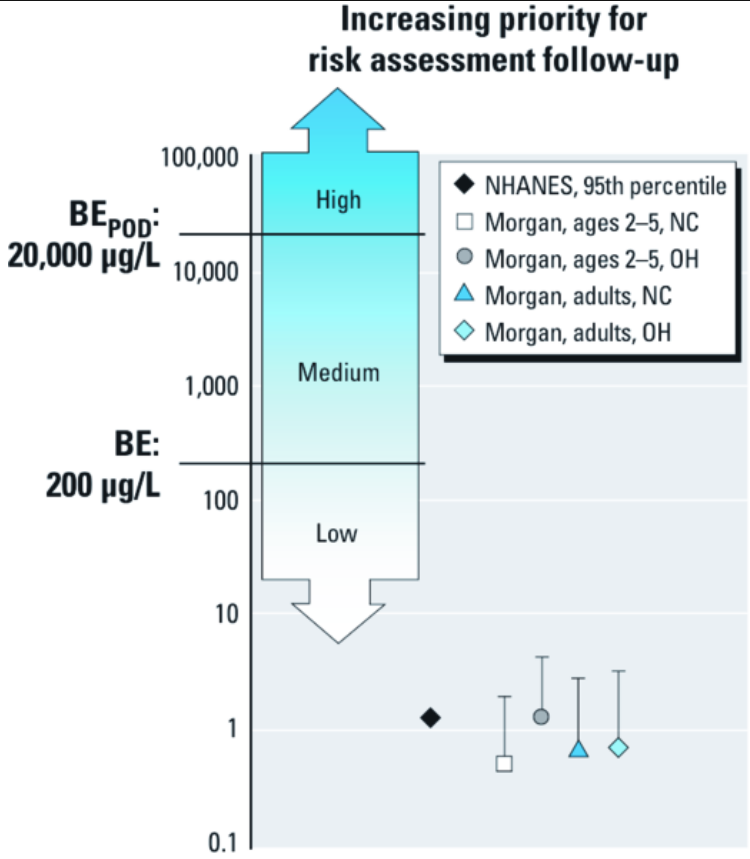
[3] For completeness, we include description of the classic BE methodology, even though it has been well-established and there have been extensive publications of both derivation and application of BEs.

data to the BE value corresponding to the risk assessment exposure guidance value deemed most appropriate.

As an example, a mass-balance approach was used to derive BE (urinary concentration µg/L) for 2,4-D based on EPA's chronic oral RfD [Aylward et al., 2010a]. This permits one to make risk-based interpretations of the general population biomonitoring studies in which 2,4-D urinary concentrations are measured, such as the U.S. National Health And Nutrition Examination Survey (NHANES). As illustrated in Figure 2, the urinary levels of 2,4-D observed in the general population are far below the BE value derived from the EPA's chronic RfD.

For some compounds for which no regulatory agency has actually derived an RfD, RfC, TDI, or MRL, etc., many times there will still be robust toxicology studies which have been conducted and which support selection of a defensible point of departure (POD). When sufficient pharmacokinetic (PK) data exists, a BE can be calculated using this screening value. Such an approach was used for deriving a BE for hexabromocyclododecane (HBCD) [Aylward and Hays, 2011].

Figure 2. Urinary 2,4-D concentrations (µg/L) in general population studies presented in the context of the BE value corresponding to the U.S. EPA RfD for general population chronic exposures.



Aylward et al. (2010). Biomonitoring Data for 2,4-Dichlorophenoxyacetic Acid in the United States and Canada: Interpretation in a Public Health Risk Assessment Context Using Biomonitoring Equivalents. Environ Health Perspect 118(2): 177-181.

2.2. Risk-Based Interpretation of Biomonitoring Based on Substances with Sufficient Toxicity Datasets but No Government-Generated (or Approved) Risk Assessment

HBCD has been the subject of numerous toxicology studies, yet no official guidance values had been derived. However, both Health Canada (HC) and the European Union (EU) identified points of departure (POD) of 10 and 20 mg/kg, respectively, based on rat repeated dose studies (see Aylward and Hays, 2011). However, neither HC nor the EU went any farther in deriving an official guidance value; neither governmental agency proposed AFs. The PODs however provide the basis for conducting a Margin of Exposure (MOE) evaluation by deriving BEs from the POD doses and then comparing these screening criteria (BE @ POD) to the available biomonitoring data.

$$MOE = \frac{BE\ @\ POD}{[Biomar\ \text{ker}]}$$

Aylward and Hays [2010] summarized the PODs available from HC and the EU, the measured or estimated lipid-adjusted concentrations of HBCD in tissues at the PODs and calculated MOEs based on central tendency and upper bound measures of HBCD in humans (Table 1).

- This example highlights how BEs can be derived for compounds which lack a formal government agency-derived exposure guidance value and how useful conclusions about MOEs can be made using these BEs along with existing human biomonitoring data.

- This approach can be applied to the many compounds for which no exposure guidance value has been derived, but yet have robust toxicology data, some knowledge of PK, and for which biomonitoring data exists.

An alternative approach would be to use a MOS evaluation. In such a case, the first step would be to select the appropriate LOAEL or a NOAEL and then derive a POD (as mg/kg-day external dose). The next step would be to apply AFs for toxicodynamics to derive a screening level health-based exposure guidance value; which is also in units of applied dose (mg/kg-bw/day). Then, in the third step, chemical-specific toxicokinetic data or models (CSTK) would be used to derive a biomarker concentration level – typically in units of concentration in blood or urine. Biomonitoring results can then be interpreted in a risk context using the MOS procedure.

Table 1: From Aylward and Hays [2011]. Points of departure identified in recent risk assessments for HBCD, with corresponding estimated lipid-adjusted concentrations. Margins of exposure (MOEs) are presented based on comparison of HBCD concentrations in population biomonitoring data to the lipid-adjusted concentrations estimated for the PODs: MOE = POD/[HBCD].

Risk Assessment	**Endpoint(s)**	**POD, administered dose (mg/kg-d)**	**POD, lipid-adjusted concentrations (ng/g lipid)[a]**	**MOE at central tendency[b]**	**MOE at upper bound[b]**
EU Draft RAR (2008)	Liver weight in female rats (BMDL for 20% change; van der Ven et al., 2006)	22.9	192,000	192,000	8,000
	Fertility (NOAEL from Ema et al., 2008)	10	121,000	121,000	6,000
HC Draft SLRA (2010)	Fertility and developmental effects (NOAEL from Ema et al., 2008)	10	121,000	121,000	6,000

[a] From regression equation provided in van der Ven et al. [2006] relating liver lipid-adjusted HBCD concentration to administered dose rate at the end of a 28 day administration period in female rats: $C_{lipid} = 33377 * D^{0.55587}$ where C_{lipid} is liver lipid-adjusted HBCD concentration (ng/g lipid) and D is the administered dose (mg/kg-d).
[b] Compared to central tendency or upper bound of general population biomonitoring data (1 or 20 ng/g lipid, respectively). For more details see From Aylward and Hays [2010].

However, when using this approach, it is important to recognize that the typical AFs of 10X for extrapolating to animals to humans and 10X to account for human variability each contain both dynamic and kinetic components. Thus, to use this method to interpret human biomonitoring data, when deriving the screening level health-based exposure guidance value from a NOAEL or POD based on an oral toxicity lab animal study, its important to use in the second step only the dynamic components of the AFs (typically 2.5X or 3.16X to extrapolate from animals to humans and 3.16X to account for human variability). Then, in a third step, where the CSTK is used to convert the applied dose into a concentration, the CSTK replaces the kinetic components of the typical AFs. If both the typical 10X for extrapolating from animals to humans and the 10X to account for human variability are applied to the lab animal toxicity NOAEL and the CSTK is also applied, "double counting" for toxicokinetics would occur.

2.3. Risk-Based Interpretation of Biomonitoring Data in the Absence of Chemical-Specific Toxicokinetic Data or Models

In some situations, an exposure guidance value may be available for a compound, but no PK data exists to allow direct calculation of a BE. In such a situation, PK information from similar

compounds may help inform bounds on the necessary PK parameters required to derive a BE (e.g., urinary excretion fraction, half-life and volume of distribution, generic PBPK models, etc.). An example of this type of approach is available from [Aylward et al., 2010b] in which steady-state solutions to PBPK models using both oral and inhalation reference values of VOCs were used to derive chemical-specific BEs for more than 35 different compounds (Figures 3 and 4, respectively). This analysis found that the range of blood concentrations of parent VOCs for a given oral or inhalation reference value was relatively narrow, especially for inhalation reference values (Figure 5). The BEs from the VOCs for which PBPK models exist (Figures 3 and 4) and the steady-state concentrations per unit dose (oral, geometric mean 12.0 µg/L per mg/kg-d; inhalation, geometric mean 3.2 µg/L per mg/m^3) can be used to inform a screening BE for a compound for which an oral or inhalation reference value exists, but no PK data are available to calculate a BE (Table 2).

This case study highlights an approach for developing BEs for compounds for which no chemical-specific PK data exists, but yet relies on PK relationships within the class of compounds. The VOCs are a data-rich class of compounds and lends themselves nicely for this type of case study. A similar approach could be envisioned for other classes of compounds, in particular a class of compounds that is biomonitored in urine and has simple metabolic schemes.

2.4. Risk-Based Interpretation of Biomonitoring Data in the Absence of Both Guidance Values and Toxicokinetic Data.

When a compound has neither a guidance value nor toxicokinetic data, a screening BE value can be derived using available BEs from within the same class of compounds. The example of VOCs could lend itself for this type of approach. For a VOC similar to those included in the VOC analysis [Aylward et al., 2010b] and which has no exposure guidance value, one could use the lower end of the BEs for the VOCs as a screening metric to interpret biomonitoring data for that compound (e.g., 0.0005 ug/L in blood – Figures 3 and 4).

Figure 3: From Aylward et al., 2010b. Estimated steady-state blood concentration screening values associated with oral reference values for 33 compounds with both oral reference values and pharmacokinetic models. Solid line and dotted lines represent the geometric mean and 95% confidence interval on the slope relating steady-state blood concentrations to oral exposure rates for 37 VOC compounds with pharmacokinetic models (see Figure 1).

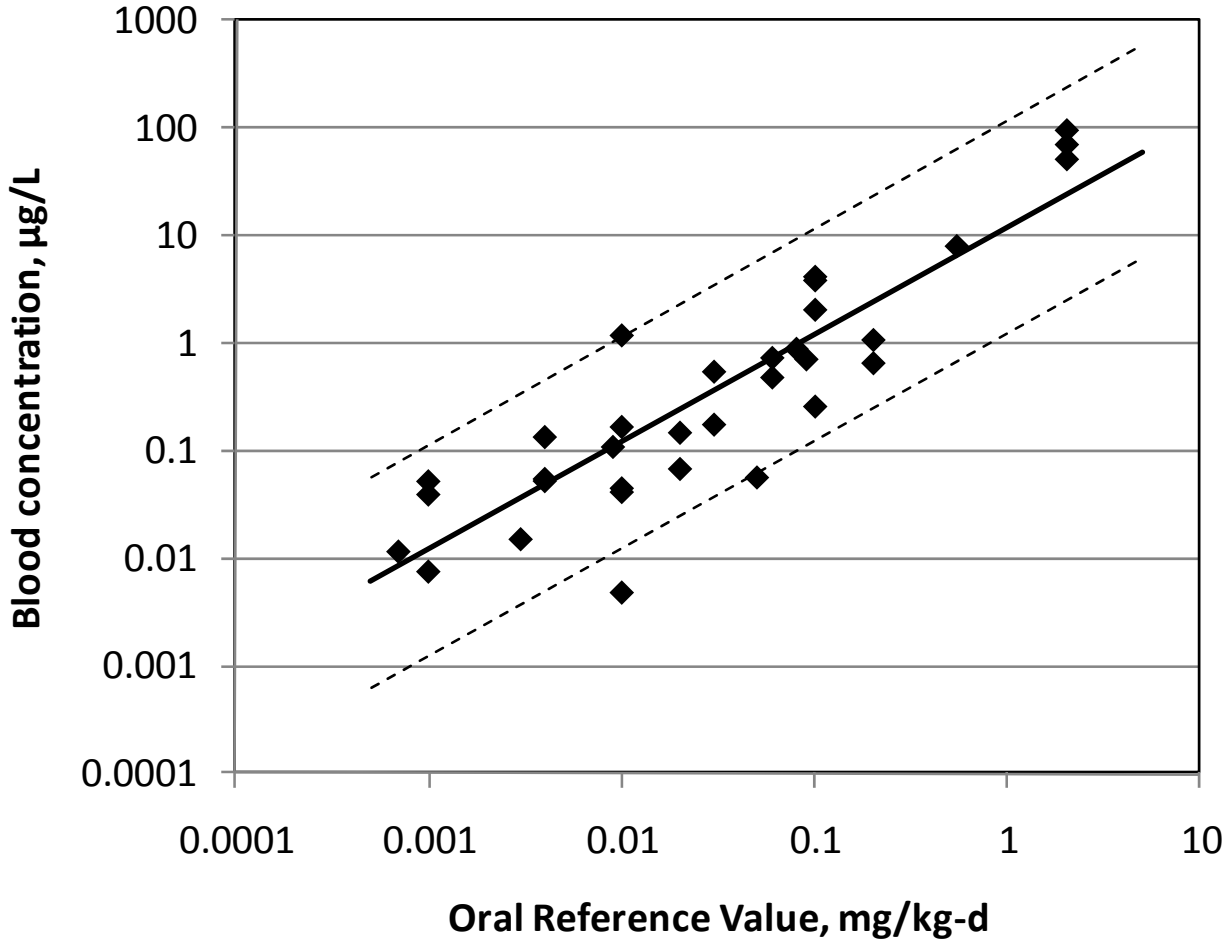


Figure 4: From Aylward et al., 2010b. Estimated steady-state blood concentration screening values associated with inhalation reference values for 23 compounds with both inhalation reference values and pharmacokinetic models. Solid line and dotted lines represent the geometric mean and 95% confidence interval on the slope relating steady-state blood concentrations to inhalation exposure levels for 38 VOC compounds with pharmacokinetic models (see Figure 1).

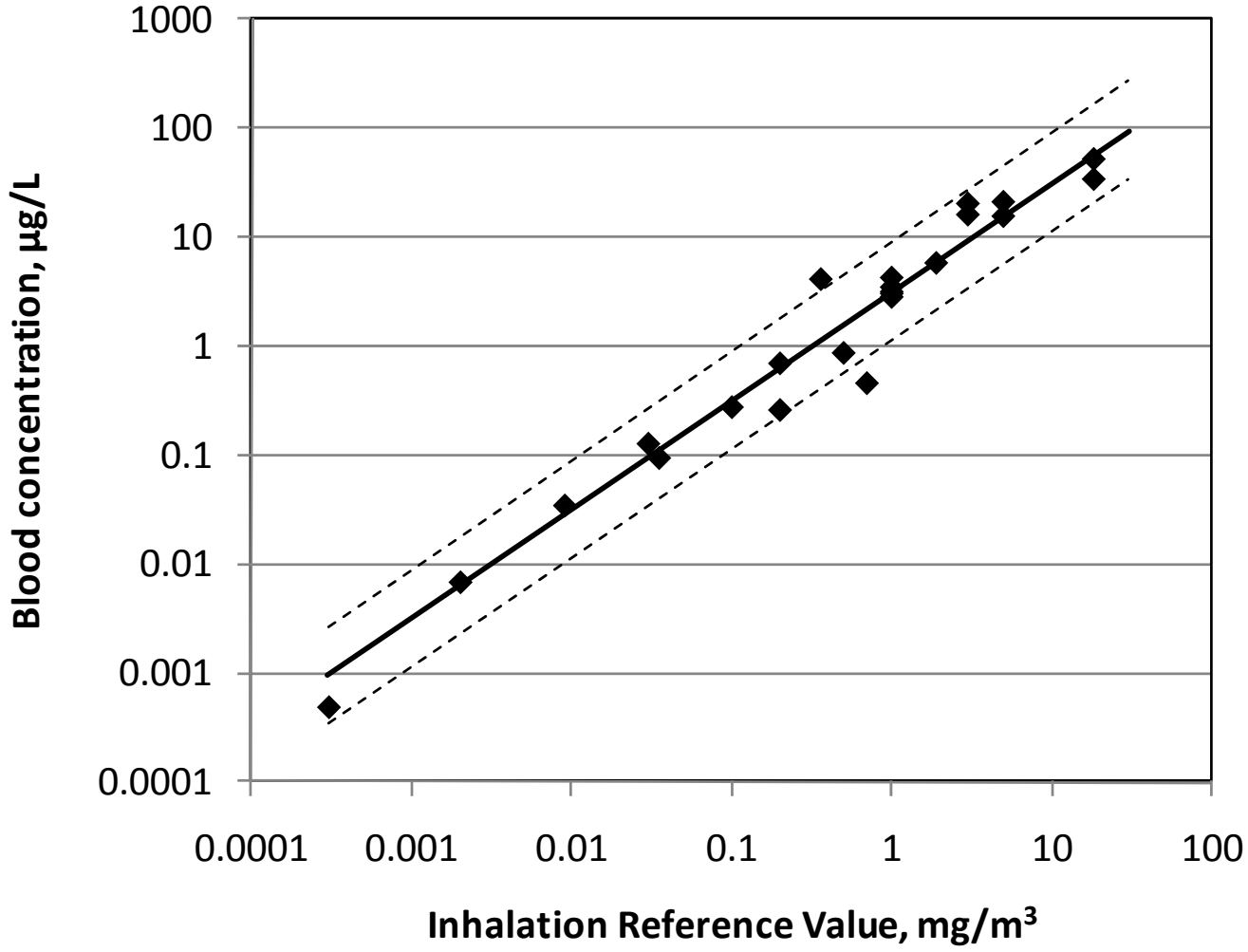


Figure 5: From Aylward et al., 2010b. Variation in estimated steady-state blood concentrations associated with a unit dose via oral (n=37) or inhalation (n=38) exposure for chemicals with pharmacokinetic models based on the steady state solutions as described in text. The symbol and bars represent the geometric mean and estimated 2.5th to 97.5th percentile of the distribution of steady-state concentrations per unit dose: oral, geometric mean 12.0 µg/L per mg/kg-d (1.2-116); inhalation, geometric mean 3.2 µg/L per mg/m^3 (1.1-9.0).

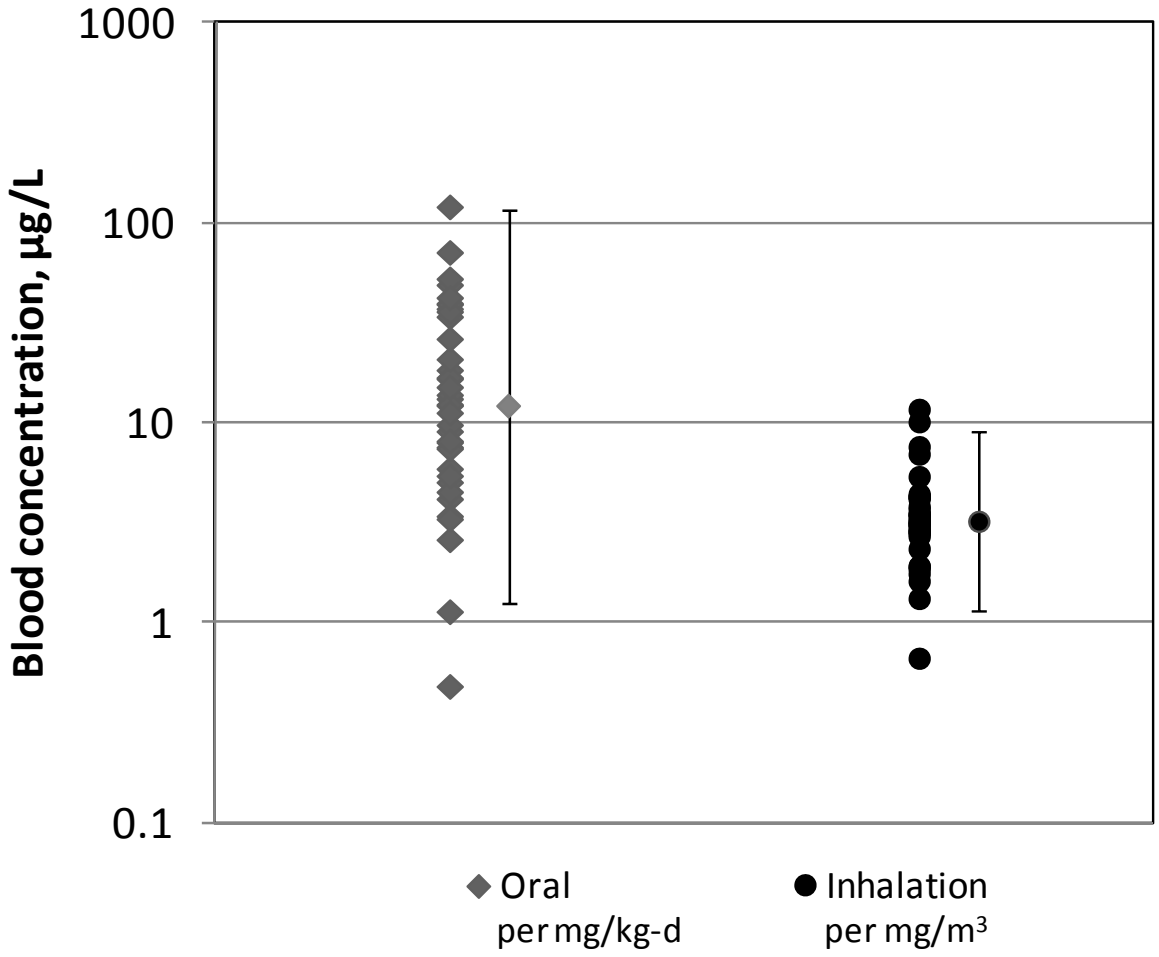


Table 2: From Aylward et al., 2010b. Extrapolated steady-state blood concentration at non-cancer chronic reference exposure levels for compounds with no identified pharmacokinetic model, extrapolated from behavior of the compounds with pharmacokinetic models. A central tendency and range of estimated blood concentrations are presented based on the geometric mean and 95% confidence interval of pharmacokinetic behavior observed among the set of 38 compounds with pharmacokinetic models identified (see Figures 3 and 4).

Chemical	Oral Reference Value (mg/kg-d)	Source	POD	POD Code	Total UF	Extrapolated steady-state blood conc., µg/L (95% CI)	Inhalation Reference Value, mg/m^3	Source	POD	POD Code	HEC POD	Total UF	Extrapolated steady-state blood conc., µg/L (95% CI)
2-Nitropropane	ND						0.02	i	16	L	16	1000	0.06 (0.02-0.2)
1,2-Dichloropropane	ND						0.004	i	12.4	L	1.3	300	0.01 (0.005-0.04)
1,2-Dibromo-3-chloropropane	ND						0.0002	i	0.17	N	0.17	1000	0.001 (0.0002-0.002)
Chlorobenzene	0.02	i	19	N	1000	0.2 (0.02-2)	0.01	HC	341	L	50	5000	0.03 (0.01-0.09)
Isopropylbenzene	0.1	i	110	N	1000	1 (0.1-10)	0.4	i	435	N	435	1000	1 (0.5-4)
Nitrobenzene	0.002	i	1.8	B	1000	0.02 (0.002-0.2)	0.009	i	0.26	B	0.26	30	0.03 (0.01-0.08)
Benzyl chloride	ND						ND						
1,4-Dichlorobenzene	ND						0.8	i	75	N	75	100	3 (0.0-7)
Carbon disulfide	0.1	i	11	N	100	1 (0.1-10)	0.7	i	20	B	20	30	2 (0.8-6)

ND- No data, no reference value found. Column headings are defined in Aylward et al., 2010b;
i, IRIS RfD; ow, EPA Office of Water RfD; h, EPA Health Effects Assessment Summary Tables (HEAST) RfD; R, RIVM TDI. POD, point of departure for derivation of reference value. POD Code: N, no-observed-adverse-effect-level (NOAEL); L, lowest-observed-adverse-effect-level (LOAEL); B, benchmark dose or concentration. HEC POD: human equivalent concentration point if departure.

2.5. Risk-Based Interpretation of Biomonitoring Based on the Thresholds of Toxicological Concern (TTC) Method

The threshold of toxicological concern (TTC) evolved from concepts initially developed by Frawley and further refined by the U.S. FDA as the Threshold of Regulation [see Felter et al., 2009; JRC, 2012]. TTC is an approach, based on chemical structure and toxicity data of structurally-related chemicals, for establishing a human exposure guidance value below which there is a very low concern for any risk to human health. A TTC has been defined as "generic human exposure thresholds for structural groups of chemicals below which no risk to human health is assumed and therefore no further testing is needed"

The TTC approach is not a formal SAR or QSAR analysis. It is based on a distribution of the toxicity potencies for chemicals that are structurally similar to the chemical of concern. To derive the TTC for a substance, its structure is evaluated[4,5] using a decision tree analytic to evaluate functional groups, and a substance is assigned to one of 3 classes:

- TTC = 0.03 mg/kg/day (Cramer Class I)
- TTC = 0.009 mg/kg/day (Cramer Class II)
- TTC = 0.0015 mg/kg/day (Cramer Class III)

For substances for which a TTC approach is used in a safety determination, extensive chemical-specific PK data are not likely to be available, making it challenging to convert a TTC value to an internal dose. By analogy to the discussion above for VOCs, it's likely that similar approaches would have to be conducted –suitable PK methods for use within a class of compounds would need to be identified (preferably for a class of compounds for which the metabolic schemes are simple or straightforward). Although we have yet to develop a specific example using a TTC, we have included such an approach in the Framework and decision tree because as TTC approaches are improved, and biomonitoring expands, this approach may be found to be appropriate and warranted under certain conditions.

[4] The ToxTree tool can be used to determine the Cramer Class of a substance http://toxtree.sourceforge.net/index.html. The EC Joint Research Centre supported the development of ToxTree, an open source computational profiling tool that can be used to assign chemicals to categories (Cramer Classes or Kroes Decision Tree) for derivation of TTCs (http://toxtree.sourceforge.net/). ToxTree has 14 modules: Cramer rules & Cramer rules with extensions, Verhaar scheme & modified Verhaar scheme, skin irritation prediction, eye irritation prediction, Benigni / Bossa rulebase for mutagenicity and carcinogenicity, START biodegradation and persistence, structure alerts for identification of Michael Acceptors, structure alerts for skin sensitization, Kroes TTC decision tree, SMARTCyp - Cytochrome P450-mediated drug metabolism and metabolites prediction, structure alerts for the in vivo micronucleus assay in rodents (ISSMIC) and structural alerts for functional group identification (ISSFUNC).

[5] Kroes et al, 2004 and Barlow, 2005 specify that the following chemical groups should be excluded from the general TTC approach because of unique properties / toxicities: heavy metals and polyhalogenated dibenzo-p-dioxins, polyhalogenated dibenzofurans and polyhalogenated biphenyls, high molecular weight chemicals, organophosphates, proteins, steroids.

If a TTC value were to be used to interpret human biomonitoring data, the TTC applied dose would need to be converted to an internal dose concentration (the corresponding biomarker concentration under the assumption of chronic steady-state exposure at the TTC). As discussed above, in converting to an internal dose concentration, attention must be paid to proper application of AFs for dynamics and kinetics to avoid "double counting." This typically will entail review of the derivation of TTC, removal of the default AF used for toxicokinetics, then applying chemical-specific or class-specific toxicokinetic data or models to obtain an internal biomarker concentration level equivalent to the TTC. Biomonitoring results could then be interpreted in a risk context using the MOS procedure.

3. Discussion

There continues to be a pressing need to develop risk-based tools to interpret human biomonitoring data. The framework presented here provides a decision tree, hierarchical approach for using available data and knowledge of toxicology and toxicokinetics to develop the benchmark values for interpreting human biomonitoring results in a health risk context.

From the perspective of scientific confidence, in general, the highest degree of confidence arises when the two elements used to derive BE values are validated substance-specific toxicokinetic data and/or models and up-to-date definitive exposure guidance values promulgated from authoritative sources. However, even in the absence of these two elements, the tools developed for chemical risk assessment can be applied to derive screening level benchmarks for comparison to human biomonitoring data to calculate a margin of safety. We have demonstrated, using the framework and decision tree, how its possible to derive interpretive metrics for placing human biomonitoring results into a health risk context even in cases where knowledge of toxicity and toxicokinetics is less than perfect. Obviously, in such cases the scientific certainty will be lower, and this needs to be factored in to decisions reached on the adequacy of a margin of safety derived from use of such screening approaches.

The confidence in the BE will be related to uncertainties associated with all the components of the science that underlie the BE derivation. The guidelines for development and communication of BEs [Hays et al., 2008; LaKind et al., 2008] discuss the confidence in chemical-specific BEs, noting that this requires consideration of the toxicity data, the understanding of ADME and mode of action that determines the relationship between the measured biomarker and the critical dose metrics related to adverse effects of the chemical, and the robustness of the pharmacokinetic data and models. In external risk assessment the exposure is calculated for an average daily dose (or a lifetime average daily dose). Whereas biomonitoring data are generally only spot samples and, depending on the substance, the half-life and the frequency of exposure, these may not represent

the average exposure over the period of interest. Dose metrics related to peak concentrations, which may be important for many chemical-specific toxic responses, are generally not considered, although they also not generally dealt with directly in risk assessments based on external dose.

The appropriate uses and limitations of BE values, particularly for compounds with relatively short biological and urinary half-lives, have been discussed (Hays et al. 2008). For modeling exposures, assuming a steady-state also leads to uncertainties and limitations in using BEs. Since spot samples are routinely used in population surveys, variability in concentrations should be factored when compared to BEs (which are calculated assuming steady-state continuous exposures). For many of these types of chemicals, the parent compounds (or metabolites) have relatively short elimination half-lives, such that concentrations within individuals would be expected to vary substantially with a given day and across days. Thus, extremes in sampled concentrations in an individual may not reflect longer term exposure levels consistent with the chronic risk assessment basis of the toxicity value which the BE has been based upon.). Therefore, for such substances, its best to use the central tendency of measured values in a population-based study for comparing to a BE value. Use of 24-h urine composite samples can serve to minimize this uncertainty.

When deriving RfDs with less than complete datasets, EPA has set a rule that the combined adjustment factor should not exceed 3,000 [U.S. EPA, 2002], although some early EPA criteria were based on a total adjustment factor of 10,000. By analogy, in the case of a substance with an incomplete toxicity dataset, if toxicokinetic components have been adequately accounted for in deriving a screening level value for interpreting human biomonitoring data using a NOAEL (or POD), one could make the case that a MOS = 300[6] (which corresponds to a combined adjustment factor of 10X for toxicodynamics (TD) and 30X for database uncertainty) would provide the same degree of confidence as a MOS =1 for a substance with a rich toxicity database (where the adjustment factor is 10X for TD).

BEs and the other biomonitoring screening values described in this case study can be used to provide a risk context to biomonitoring data in just the same way that RfDs or NOAELs or TTCs are used in evaluating external dose information, since the biomonitoring interpretation values are simply external doses transformed to their corresponding internal concentrations. And, just as RfDs are not diagnostic of a disease or health state, neither are these biomonitoring interpretation values. BEs are not diagnostic medical biomarkers. BEs are not like serum cholesterol or serum liver enzymes, which have been shown through extensive scientific studies to be biomarkers indicative of (or correlated with) a specific health status or health outcome risk factor in humans. Similarly, RfDs are not medical diagnostic tools either. That said, RfDs have

[6] AF_{AH} = (3.16 TK x 3.16 TD); AF_{H} = (3.16 TK x 3.16 TD; AF_{DB}= 30.

been used as decision tools in the evaluation of potential health risks to both individuals and populations, and BEs and related biomonitoring interpretation tools can be used in a similar manner.

For example, an RfD can be used to evaluate potential individual risk at the point of contact by an individual consuming drinking water with a measured or modeled concentration of a contaminant. The external dose of the contaminant received by consumption of the water is calculated for that individual and compared to an RfD. Similarly, at an individual's point of contact with a measured or modeled concentration of a contaminant in air, the air concentration dose at the point of contact (in mg/m3) can be compared to an RfC. By comparing these external intake doses to the RfD or RfC, an evaluation can be made about potential risk to individuals and populations. Such information is central to risk management evaluations of different actions to address exposures, including the no-action alternative. Again, this is not a medical diagnostic decision, but rather a risk-based public health decision. And by analogy, when using a BE or similar screening tool to interpret internal dose (concentration), an individual's biomonitoring data – or a population's results -- can be interpreted and a risk-based public health decision reached for individuals and the population. The BE is just an extension of the use of the RfD in a manner that allows one to use the RfD to interpret not only potential risks measured by point of contact intake doses but also to interpret internal doses. Similarly, the BEs derived using the framework and decision tree presented here are not any different than approaches that are used to determine individual and population risks based on derivation of screening level external dose benchmarks, such as applied dose NOAELs or applied dose TTCs.

The confidence one will have in the use of these approaches to interpret human biomonitoring data depends on the underlying confidence in the toxicity datasets and the toxicokinetic data or models used for a specific substance. There will be greater certainty when interpreting biomonitoring data for substances with extensive toxicity datasets and CSTK data or models. When using class data or generic models, concerns can arise regarding domains of applicability, and these need to be evaluated prior to using such methods.

In addition to the data availability limitations, concern can arise when a BE is derived from a government health guidance value that is old, out-dated or otherwise inconsistent with current scientific best practices. . In such situations, the interpretation of the biomonitoring data may have the appearance of a high level of confidence or certainty, where in actuality, due to deficiencies in the underlying health guidance value, there is considerably lower confidence and greater uncertainty. In such cases, it may be preferable derive a BE by using the more up to date scientific data and, if available, knowledge of mode of action, instead of using the out-dated government health guidance value.

With respect to sensitive populations, for BEs based upon RfDs the derived BEs cover human variability and sensitive subpopulations in the same manner as the RfD. EPA defines the RfD (http://epa.gov/risk/glossary.htm#r) as follows:

> *Reference dose (RfD): An estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime. It can be derived from a NOAEL, LOAEL, or benchmark dose, with uncertainty factors generally applied to reflect limitations of the data used. Generally used in EPA's noncancer health assessments. [Durations include acute, short-term, subchronic, and chronic and are defined individually in this glossary].*

When other approaches are used to derive a screening level BE, then care must be taken to assure the population for which the interpretation of the biomonitoring data is to be applied is consistent with the populations covered by the derivation of the POD and the populations covered by use of the specific toxicokinetic approach.

References:

Aylward et al., 2010a. Biomonitoring Data for 2,4-Dichlorophenoxyacetic Acid in the United States and Canada: Interpretation in a Public Health Risk Assessment Context Using Biomonitoring Equivalents. Environ Health Perspect 118(2): 177-181.

Aylward et al.,2010b. Chemical-specific screening criteria for interpretation of biomonitoring data for volatile organic compounds (VOCs)—application of steady-state PBPK model solutions. Regulatory Toxicology and Pharmacology, vol. 58, no. 1, pp. 33–44.

Aylward, L.L., S.M. Hays. 2011. Biomonitoring-based risk assessment for hexabromocyclododecane (HBCD). Int. J. Hyg. Environ. Health 214: 179–187.

Barlow, 2005. Threshold of Toxicological Concern (TTC) – A Tool for Assessing Substances of Unknown Toxicity Present at Low Levels in the Diet. (ILSI Europe Concise Monograph Series, Brussels), pp. 1–32.

Becker et al, 2011. Development of Screening Tools for the Interpretation of Chemical Biomonitoring Data. Journal of Toxicology Volume 2012 (2012), Article ID 941082, 10 pages; doi:10.1155/2012/941082; http://www.hindawi.com/journals/jt/2012/941082/ref/

Felter et al., 2009. Refining the Threshold of Toxicological Concern (TTC) for Risk Prioritization for Trace Chemicals in Food. Fd. Chem. Toxicol., 47: 2236-2245.

Hays et al., 2008. Guidelines for the derivation of biomonitoring equivalents: report from the Biomonitoring Equivalents Expert Workshop. Regulatory Toxicology and Pharmacology, vol. 51, no. 3, pp. S4–S15.

Health Council of the Netherlands, 2001. Toxicity Testing: A more efficient approach. http://www.gezondheidsraad.nl/sites/default/files/01@24E_0.pdf.

IPCS, 2005. Chemical-Specific Adjustment Factors For Interspecies Differences And Human Variability: Guidance Document For Use Of Data In Dose/Concentration–Response Assessment http://whqlibdoc.who.int/publications/2005/9241546786_eng.pdf.

JRC, 2012. Applicability of the Threshold of Toxicological Concern (TTC) approach to cosmetics – preliminary analysis. http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/doc/ReqNo_JRC68188_LB-NA-25162-EN-N_-PDF.pdf

LaKind et al., 2008. Guidelines for the communication of biomonitoring equivalents: report from the Biomonitoring Equivalents Expert Workshop. Regulatory Toxicology and Pharmacology, vol. 51, no. 3, pp. S16–S26.

TERA, 2005. The Scientific Rationale for Deriving Database and Toxicodynamic Uncertainty Factors for Reproductive or Developmental Toxicants http://www.tera.org/Peer/UFD/PC2%20UFD%20Final%208-24-05.pdf.

U.S. EPA, 2002. A Review of the Reference Dose and Reference Concentration Process. EPA/630/P-02/002F. 01 Dec 2002. U.S. EPA, Risk Assessment Forum, Washington, DC. Available at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=55365

Bibliography of Published BE Papers

Aylward, L.L., S.M. Hays. 2011. Biomonitoring-based risk assessment for hexabromocyclododecane (HBCD). Int. J. Hyg. Environ. Health 214: 179–187.

Hays SM, Pyatt DW, Kirman CR, Aylward LL. 2012. Biomonitoring Equivalents for benzene. Regul Toxicol Pharmacol. 2012 Feb;62(1):62-73. Epub 2011

Hays SM, Aylward LL, Kirman CR, Krishnan K, Nong A. 2011. Biomonitoring equivalents for di-isononyl phthalate (DINP). Regul Toxicol Pharmacol. 2011 Jul;60(2):181-8.

Aylward LL, Krishnan K, Kirman CR, Nong A, Hays SM. 2011. Biomonitoring equivalents for deltamethrin. Regul Toxicol Pharmacol. 2011 Jul;60(2):189-99.

Aylward and Hays, 2010. Human biomonitoring assessment values: approaches and data requirements. International Journal of Hygiene and Environmental Health, vol. 214, no. 5, pp. 348–360.

Hays SM, Aylward LL, Gagné M, Nong A, Krishnan K. Biomonitoring equivalents for inorganic arsenic. Regul Toxicol Pharmacol. 2010 Oct;58(1):1-9.

Aylward and Hays, 2011. Biomonitoring-based risk assessment for hexabromocyclododecane (HBCD). International Journal of Hygiene and Environmental Health, vol. 214, no. 3, pp. 179–187.

Becker et al, 2011. Development of Screening Tools for the Interpretation of Chemical Biomonitoring Data. Journal of Toxicology Volume 2012 (2012), Article ID 941082, 10 pages; doi:10.1155/2012/941082; http://www.hindawi.com/journals/jt/2012/941082/ref/

Boogaard, P.J., S.M. Hays, L.L. Aylward. 2011. Human biomonitoring as a pragmatic tool for health risk management of chemicals - Examples under the EU REACh programme. Reg. Toxicol. Pharmacol. 59(1):125-32 http://dx.doi.org/10.1016/j.yrtph.2010.09.015

Aylward, L.L., S.M. Hays, M. Gagné, A. Nong, K. Krishnan. 2010. Biomonitoring Equivalents for hexachlorobenzene. . 58(1): 25-32 Reg. Toxicol. Pharmacol. doi:10.1016/j.yrtph.2010.06.003

Hays, S.M., L.L. Aylward, M. Gagné, A. Nong, K. Krishnan. 2010. Biomonitoring Equivalents

for inorganic arsenic. Reg. Toxicol. Pharmacol 58(1):1-9. doi:10.1016/j.yrtph.2010.06.002

Krishnan, K., M. Gagné, A. Nong, L.L. Aylward, S.M. Hays. 2010. Biomonitoring Equivalents for triclosan. Reg. Toxicol. Pharmacol. 58(1):10-17 doi:10.1016/j.yrtph.2010.06.004

Krishnan, K., M. Gagné, A. Nong, L.L. Aylward, S.M. Hays. 2010. Biomonitoring Equivalents for bisphenol A (BPA). Reg. Toxicol. Pharmacol. 58(1):18-24 doi:10.1016/j.yrtph.2010.06.005

Aylward, L.L., C.R. Kirman, B.C. Blount, S.M. Hays. 2010. Chemical-specific screening criteria for interpretation of biomonitoring data for volatile organic compounds (VOCs) - Application of steady-state PBPK model solutions. Reg. Toxicol. Pharmacol. 58(1):33-44. http://dx.doi.org/10.1016/j.yrtph.2010.05.011

Aylward, L.L, M.K. Morgan, T.E. Arbuckle, D.B. Barr, C.J. Burns, B.H. Alexander, S.M. Hays. 2010. Biomonitoring data for 2,4-dichlorophenoxyacetic acid in the US and Canada: Interpretation in a public health risk assessment context using Biomonitoring Equivalents. Environ. Health Perspect. 118:117-181. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2831914/?tool=pubmed

LaKind, J.S., D.Q. Naiman, S.M. Hays, L.L. Aylward, B.C. Blount. 2009. Public health interpretation of trihalomethane blood levels in the United States: NHANES 1999-2004. J. Exp. Sci. Environ. Epidem. 19(4):435-441. http://www.nature.com/jes/journal/v20/n3/full/jes200935a.html

Hays, S.M., L.L. Aylward, M. Gagné, K. Krishnan. 2009. Derivation of Biomonitoring Equivalents for cyfluthrin. Regul. Toxicol. Pharmacol. 55(3):268-75. doi:10.1016/j.yrtph.2009.09.002

Aylward, L.L., S.M. Hays, M. Gagné, K. Krishnan. 2009. Derivation of Biomonitoring Equivalents for di(2-ethylhexyl)phthalate (CAS No. 117-81-7). Regul. Toxicol. Pharmacol. 55(3):249-58. doi:10.1016/j.yrtph.2009.09.001

Aylward, L.L., S.M. Hays, M. Gagné, K. Krishnan. 2009. Derivation of Biomonitoring Equivalents for di-n-butyl phthalate (DBP), benzylbutyl phthalate (BzBP), and diethyl phthalate (DEP). Regul. Toxicol. Pharmacol. 55(3):259-267. doi:10.1016/j.yrtph.2009.09.003

Hays, S.M. and L.L. Aylward. 2009. Using biomonitoring equivalents to interpret human biomonitoring data in a public health risk context. J. Appl. Toxicol. 29(4):275-288.

http://www3.interscience.wiley.com/journal/121588561/abstract?CRETRY=1&SRETRY=0

LaKind, J.S., S.M. Hays, L.L. Aylward, and D.Q. Naiman. 2009. Perspective on serum dioxin levels in the United States: An evaluation of the NHANES data. J. Exp. Sci. Environ. Epidemiol. 19:435-441. http://www.nature.com/jes/journal/v19/n4/full/jes200863a.html

Aylward, L.L., J.S. LaKind, and S.M. Hays. 2008. Derivation of Biomonitoring Equivalent (BE) values for 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and related compounds: A screening tool for interpretation of biomonitoring data in a risk assessment context. J. Toxicol. Enivron. Health Pt. A. 71(22):1499-1508. http://www.informaworld.com/smpp/content~db=all?content=10.1080/15287390802361755

Hays, S.M., L.L. Aylward, and J.S. Lakind. 2008. Introduction to the Biomonitoring Equivalents Pilot Project: Development of guidelines for the derivation and communication of Biomonitoring Equivalents. Reg. Toxicol. Pharmacol. 51:S1-S2. doi:10.1016/j.yrtph.2008.02.007

Hays SM, Aylward LL, LaKind JS, Bartels MJ, Barton HA, Boogaard PJ, Brunk C, DiZio S, Dourson M, Goldstein DA, Kilpatrick ME, Krewski D, Krishnan, K, Lipscomb J, Nordberg M, Okino M, Tan Y-M, Viau C, Yager JW. 2008. Guidelines for the derivation of Biomonitoring Equivalents: Report from the Biomonitoring Equivalents Expert Workshop. Reg. Toxicol. Pharmacol. 51:S4-S15. doi:10.1016/j.yrtph.2008.05.004

LaKind, J.S., L.L. Aylward, C. Brunk, S. DiZio, M. Dourson, D.A. Goldstein, M.E. Kilpatrick, D. Krewski, M.J. Bartels, H.A. Barton, P.J. Boogaard, J. Lipscomb, K. Krishnan, M. Nordberg, M. Okino, Y.-M. Tan, C. Viau, J.W. Yager, and S.M. Hays. 2008. Guidelines for the communication of Biomonitoring Equivalents: Report from the Biomonitoring Equivalents Expert Workshop. Reg. Toxicol. Pharmacol. 51:S16-S26. doi:10.1016/j.yrtph.2008.05.007

Hays, S.M., and L.L. Aylward. 2008. Biomonitoring equivalents (BE) dossier for acrylamide (AA) (CAS No. 79-06-1). Reg. Toxicol. Pharmacol. 51:S57-S67. doi:10.1016/j.yrtph.2008.05.010

Hays, S.M., M. Nordberg, J.W. Yager, and L.L. Aylward. 2008. Biomonitoring equivalents (BE) dossier for cadmium (Cd) (CAS No. 7440-43-9). Reg. Toxicol. Pharmacol. 51:S49-S56. doi:10.1016/j.yrtph.2008.05.008

Aylward, L.L., J.S. LaKind, and S.M. Hays. 2008. Biomonitoring equivalents (BE) dossier for

trihalomethanes. Reg. Toxicol. Pharmacol. 51:S68-S77. doi:10.1016/j.yrtph.2008.05.005

Aylward, L.L., and S.M. Hays. 2008. Biomonitoring equivalents (BE) dossier for 2,4-dichlorophenoxyacetic acid (2,4-D) (CAS No. 94-75-7). Reg. Toxicol. Pharmacol. 51:S37-S48. doi:10.1016/j.yrtph.2008.05.006

Aylward, L.L., H.A. Barton, and S.M. Hays. 2008. Biomonitoring equivalents (BE) dossier for toluene (CAS No. 108-88-3). Reg. Toxicol. Pharmacol. 51:S27-S36. doi:10.1016/j.yrtph.2008.05.009

LaKind, J.S., L. Barraj, N. Tran, and L.L. Aylward. 2008. Environmental chemicals in people: Challenges in interpreting biomonitoring information. J. Environ. Health. 70(9):61-64. http://www.biomonitoringequivalents.net/LaKind_et_al_JEH_May_2008_Guest_Commentary.pdf

Hays, S.M., R. Becker, H.-W. Leung, L.L. Aylward, and D.W. Pyatt. 2007. Biomonitoring equivalents: A screening approach for interpreting biomonitoring results from a public health risk perspective. Reg Toxicol Pharmacol. 47(1):96-109. doi:10.1016/j.yrtph.2006.08.004