"This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement N°603946

HEALS

Health and Environment-wide Associations based on Large population Surveys

FP7-ENV-2013- 603946
http://www.heals-eu.eu/

6.1 Modelling module for biomonitoring data assimilation

WP 6 Physiology based biokinetic modeling for internal dose and exposure reconstruction

Version 2 (31/03/2016)

Lead beneficiary: AUTH
Date: 31/03/2016
Nature: Other - O
Dissemination level: Public - PU

Document Information

Grant Agreement Number	ENV-603946	**Acronym**	HEALS
Full title	Health and Environment-wide Associations based on Large population Surveys		
Project URL	http://www.heals-eu.eu/		
EU Project Officer	Tuomo Karjalainen,- Tuomo.KARJALAINEN@ec.europa.eu		

Deliverable	**Number**	6.1	**Title**	Modelling module for biomonitoring data assimilation
Work Package	**Number**	6	**Title**	Physiology based biokinetic modeling for internal dose and exposure reconstruction

Delivery date	**Contractual**	M30	**Actual**	31/03/2016
Status	Draft ☐		Final ☒	
Nature	Demonstrator ☐	Report ☐	Prototype ☐	Other ☒
Dissemination level	Confidential ☐	Public ☒		

Author (Partners)	Denis A. Sarigiannis, S. Karakitsios, A. Gotti			
Responsible Author	Denis A. Sarigiannis		**Email**	denis@eng.auth.gr
	Partner	AUTH	**Phone**	+30-2310-994562

Document History

Name	Date	Version	Desciption
AUTH	29/02/2016	1	Draft report
AUTH	31/03/2016	2	Final report

TABLE OF CONTENTS

1 Introduction

1.1 Aims of HEALS internal dosimetry modelling framework

HEALS generic Physiology Based BioKinetic (PBBK) model aims at providing a comprehensive methodology for internal exposure assessment, by (a) translating external exposure into internal dosimetry and (b) allowing the identification of external and internal exposure through the assimilation of biomonitoring data. This provides very significant opportunities for advancing risk assessment towards exposure based assessments:

- Forward estimation of exposure will allow us to translate external exposure data to Biologically Effective Dose (BED) at the tissue dose; this will allow the interpretation to mechanism based hazard assessment based metrics such as the Biological Pathway Altering Dose (BPAD) (Judson et al., 2010; Judson et al., 2011). The quantity that will be calculated, the biological pathway altering dose (BPAD), is analogous to current risk assessment metrics in that it combines dose−response data with analysis of uncertainty and population variability so as to derive exposure limits. The analogy is closest when perturbation of a pathway is a key event in the mode of action (MOA) leading to a specified adverse outcome. BPADs are derived from relatively inexpensive, high-throughput screening (HTS) *in vitro* data. Use of as detailed as possible PBBK modeling is the key component so as to estimate the *in vivo* doses required to achieve the BPAD in the target tissue. Uncertainty and variability will be incorporated in both the BPAD and the PBBK parameters and then combined to yield a probability distribution for the dose required to perturb the critical pathway. Thus, the more confident and explicit we are about the biokinetic behavior of the compound, the less conservative we have to be when translating BPAD into external exposure metrics. Information about BPAD for several chemicals can be easily obtained from the publicly available ToxCastDB (USEPA).
- Assimilation of the vast amount of biomonitoring data continuously produced at the national or international level. Assimilation of biomonitoring data is greatly facilitated by exposure reconstruction (or reverse dosimetry) techniques, allowing us to estimate the actual external dose corresponding to the observed biomonitored levels. This will also allow us the possible identification of route contribution, and consequently to exposure pathways contribution identification. Also, reconstructing exposure will allow us to re-run forward our assessment, and to estimate BED, starting actually from biomonitoring data.
- Refine overall assessment of exposure, as well as assimilation of biomonitoring data, taking strongly into account the differences in physiology among different age groups, capturing the differences in BED for similar exposure situations, as well as to attribute the appropriate variance of exposure, to significantly varying biomonitoring data.

1.2 State of the art on internal exposure dosimetry

PBBK models are continuously gaining ground in regulatory toxicology, describing in quantitative terms the absorption, metabolism, distribution and elimination (ADME) processes in the human body, with a focus on the effective dose at the expected target site (Bois et al., 2010). This trend is further amplified by the continuously increasing scientific and regulatory interest about aggregate and cumulative exposure; PBBK models translate external exposures from multiple routes (Yang et al., 2010) into internal exposure metrics, addressing the effects of exposure route in the overall bioavailability (Sarigiannis and Karakitsios, 2011; Valcke and Krishnan, 2011) or the dependence on critical developmental windows of susceptibility, such as pregnancy (Beaudouin et al., 2010), lactation (Verner et al., 2008) and infancy (Edginton and Ritter, 2009). With regard to cumulative exposure, PBBK models offer the advantage of calculating the effect of the interactions among the mixture compounds at the level of metabolism, however due to the inherent difficulties arising, the existing applications are limited to VOCs (Cheng and Bois, 2011; Haddad et al., 2000; Sarigiannis and Gotti, 2008) and metals (Sasso et al., 2010). Recently, efforts have shifted also towards the integration of whole-body physiology, disease biology, and molecular reaction networks (Eissing et al., 2011), as well as integration of cellular metabolism into multi-scale whole-body models (Krauss et al., 2012).

PBBK models are also used for assimilating biomonitoring data, through exposure reconstruction, meaning the quantification of exposure components related to the observed biomarkers concentrations. Several techniques have been developed in this direction with increased level of complexity, starting from the exposure conversion factor (ECF) (Tan et al., 2006), up to combined Maximum Likelihood Estimates coupled to PBBK modeling approaches with synthetic biomarker data under a Bayesian statistics framework (Georgopoulos et al., 2009). The need for direct risk characterization based on biomonitoring data, resulted in the establishment of the biomonitoring equivalences (BEs) (Hays and Aylward, 2009). BEs values represent quantitative benchmarks of safe or acceptable concentrations of a chemical or its metabolite in biological specimens that are consistent with selected reference values, such as the ADI, TDI, MRL and RfD, using the knowledge about the biokinetic properties of the chemical (Boogaard et al., 2011). However, the use of reliable PBBK models is the most convenient way on translating external exposure reference values into BEs. Moreover, in order to utilize the capabilities of *in vitro* testing, PBBK models are used to identify the Biological Pathway Altering Dose (Judson et al., 2011), is analogous to current risk assessment metrics in that it combines dose-response data with analysis of uncertainty and population variability to arrive at conservative exposure limits. The analogy is closest when perturbation of a pathway is a key event in the mode of action (MOA) leading to a specified adverse outcome.

Considering the opportunities offered by the use of PBBK models in exposure/risk characterization, several research groups are developing generic PBBK models, either

stand-alone models such as PK-Sim (Willmann et al., 2003), Indus-Chem (Jongeneelen and Berge, 2011) or incorporated within integrated computational platforms for exposure assessment such as INTERA (Sarigiannis et al., 2011) and MENTOR (Georgopoulos et al., 2008b). The development of generic PBBK models is substantiated by the recent advances in quantitative structure–activity relationships (QSARs) and quantitative structure–property relationships (QSPRs) (Peyret and Krishnan, 2011; Price and Krishnan, 2011), providing the basis for development of relevant PBBK models regarding data-poor or new chemicals.

2 Internal exposure modeling framework

2.1 Conceptual description

The model is designed to describe as much as possible the actual ADME processes occurring in human body, so as to be easily applicable for a broad variety of chemicals under proper parameterization. The model will include the parent compound and a number of three potential metabolites. For each compound/metabolite all major organs will be included (Figure 1) and the link among the compounds and the metabolites will be through the metabolizing tissues. This is mainly the liver, but also other sites of metabolism might be considered based on the presence or not of the enzymes involved in the metabolism of the compound of interest.

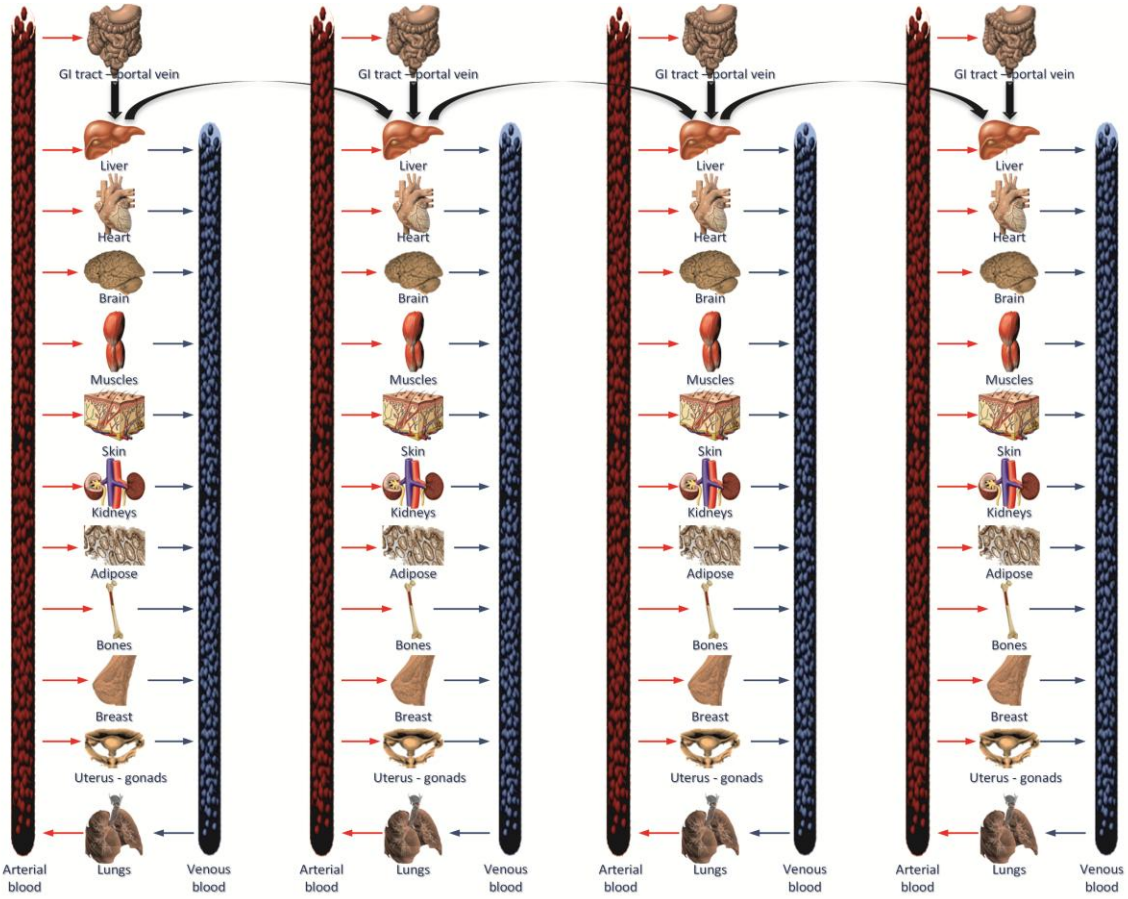


Figure 1. The generic PBTK model (parent compound and 3 metabolites)

In order to capture the in-utero exposure, the model is also replicated in order to describe the functional interaction of the mother and the developing fetus through the placenta. The anthropometric parameters of the models are age dependent, so as to provide a lifetime internal dose assessment. Details on the mathematical framework so as to describe the structure of the generic-lifetime PBTK model are given in the following sections.

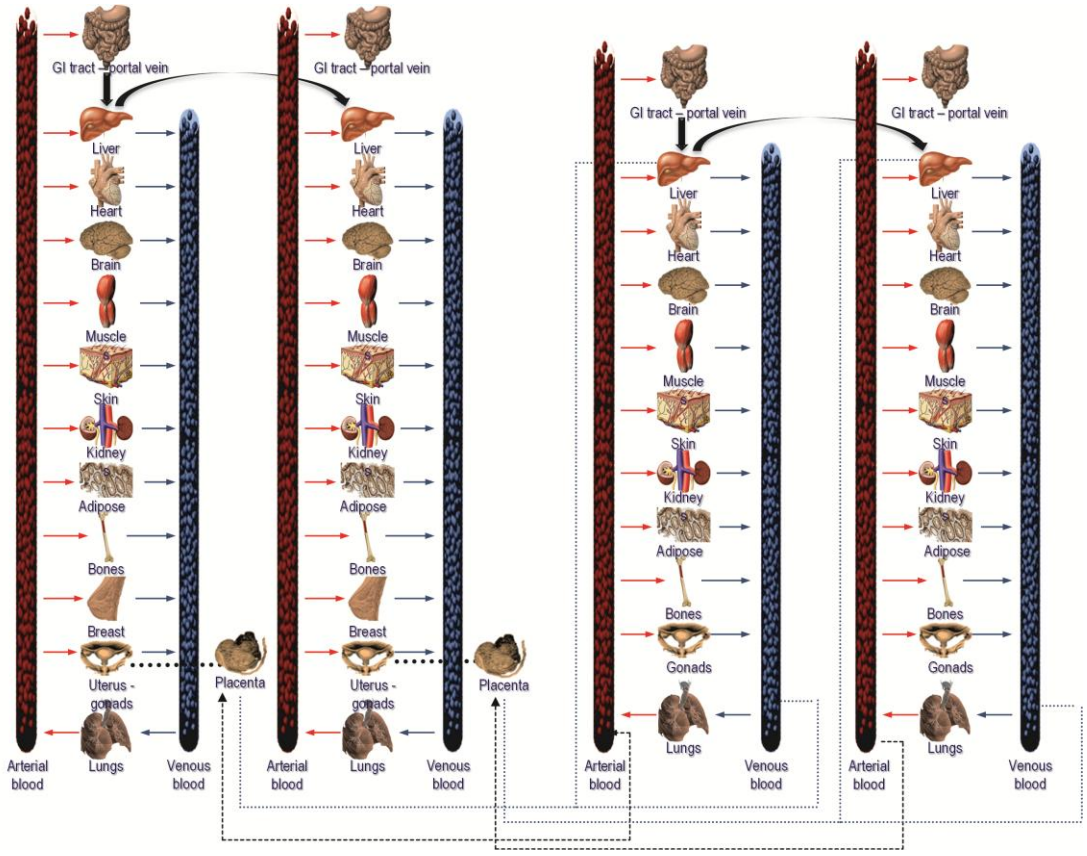


Figure 2. Conceptual representation of the Mother-Fetus PBTK model

2.2 Components of the generic life-time HEALS PBTK model

2.2.1 Formulation describing the generic multi-route PBTK model

The formulation scheme of the developed PBPK model is described below:

For non-eliminating organs:

Red blood cells

$$V_{rbc\_org}\frac{dC_{rbc\_org}}{dt} = Q_{org}\cdot HCT\cdot\left(C_{rbc\_art}-C_{rbc\_org}\right) + PS_{rbc\_org}\cdot f_u\cdot\left(C_{\text{int}\_org}-\frac{C_{rbc\_org}}{K_{rbc}}\right)$$

Where $V_{rbc\_org} = f_{vas\_org}\cdot V_{org}\cdot HCT$

Plasma +interstitial

$$V_{\text{int}\_org}\frac{dC_{\text{int}\_org}}{dt} = Q_{org}\cdot(1-HCT)\cdot\left(C_{pls\_art}-C_{\text{int}\_org}\right) - PS_{rbc\_org}\cdot f_u\cdot\left(C_{\text{int}\_org}-\frac{C_{rbc\_org}}{K_{rbc}}\right) - PS_{cell\_org}\cdot f_u\cdot\left(C_{\text{int}\_org}-\frac{C_{cell\_org}}{K_{org}}\right)$$

Where $V_{\text{int}\_org} = V_{org}\cdot\left[f_{\text{int}\_org}+f_{vas\_org}\cdot(1-HCT)\right]$

Cellular

$$V_{cell\_org}\frac{dC_{cell\_org}}{dt} = PS_{cell\_org}\cdot f_u\cdot\left(C_{\text{int}\_org}-\frac{C_{cell\_org}}{K_{org}}\right)$$

Where $V_{cell\_org} = f_{cell\_org}\cdot V_{org}$

Kidney

For kidney renal plasma clearance (*CLpls_kid)* is incorporated:

Plasma + Interstitial

$$V_{\text{int\_}kid}\frac{dC_{\text{int\_}kid}}{dt}=Q_{kid}\cdot\left(1-HCT\right)\cdot\left(C_{pls\_art}-C_{\text{int\_}kid}\right)-PS_{rbc\_kid}\cdot f_u\cdot\left(C_{\text{int\_}kid}-\frac{C_{rbc\_liv}}{K_{rbc}}\right)$$
$$-PS_{cell\_kid}\cdot f_u\cdot\left(C_{\text{int\_}kid}-\frac{C_{cell\_kid}}{K_{kid}}\right)-\frac{CL_{pls\_kid}\cdot Q_{kid}\cdot\left(1-HCT\right)\cdot C_{\text{int\_}kid}}{\left[Q_{kid}\cdot\left(1-HCT\right)-CL_{pls\_kid}\right]}$$

Portal vein

Red blood cells

$$HCT\cdot V_{pv}\frac{dC_{rbc\_pv}}{dt}=Q_{GI\_tract}\cdot HCT\cdot\left(C_{rbc\_GI\_tract}-C_{rbc\_pv}\right)$$
$$+PS_{rbc\_pv}\cdot f_u\cdot\left(C_{pls\_pv}-\frac{C_{rbc\_org}}{K_{rbc}}\right)$$

Plasma + interstitial

$$\left(1-HCT\right)\cdot V_{pv}\cdot\frac{dC_{pls\_pv}}{dt}=Q_{GI\_tract}\cdot\left(1-HCT\right)\cdot\left(C_{\text{int\_}GI\_tract}-C_{pv}\right)$$
$$-PS_{rbc\_pv}\cdot f_u\cdot\left(C_{pls\_pv}-\frac{C_{rbc\_pv}}{K_{rbc}}\right)$$

For venous blood

Red blood cells

$$HCT \cdot V_{ven} \cdot \frac{dC_{rbc\_ven}}{dt} = \sum_{org} Q_{org} \cdot HCT \cdot C_{rbc\_org} + PS_{rbc\_ven} \cdot f_u \cdot \left(C_{pls\_ven} - \frac{C_{rbc\_ven}}{K_{rbc}} \right) - Q_{lung} \cdot HCT \cdot C_{rbc\_ven}$$

Plasma +interstitial

$$(1-HCT) \cdot V_{ven} \frac{dC_{pls\_ven}}{dt} = \sum_{org} Q_{org} \cdot (1-HCT) \cdot C_{\text{int}\_org} - PS_{rbc\_ven} \cdot f_u \cdot \left(C_{pls\_ven} - \frac{C_{rbc\_ven}}{K_{rbc}} \right) - Q_{lung} \cdot (1-HCT) \cdot C_{pls\_ven}$$

For arterial blood

Red blood cells

$$HCT \cdot V_{art} \cdot \frac{dC_{rbc\_art}}{dt} = -\sum_{org} Q_{org} \cdot HCT \cdot C_{rbc\_art} + PS_{rbc\_art} \cdot f_u \cdot \left(C_{pls\_art} - \frac{C_{rbc\_art}}{K_{rbc}} \right) + Q_{lung} \cdot HCT \cdot C_{rbc\_art}$$

Plasma +interstitial

$$(1-HCT) \cdot V_{art} \frac{dC_{pls\_art}}{dt} = -\sum_{org} Q_{org} \cdot (1-HCT) \cdot C_{\text{int}\_art} - PS_{rbc\_art} \cdot f_u \cdot \left(C_{pls\_art} - \frac{C_{rbc\_art}}{K_{rbc}} \right) + Q_{lung} \cdot (1-HCT) \cdot C_{pls\_art}$$

For liver:

Red blood cells

$$V_{rbc\_liv}\frac{dC_{rbc\_liv}}{dt} = HCT\cdot\left(Q_{liv}\cdot C_{rbc\_art} + Q_{pv}\cdot C_{rbc\_pv}\right) - \left(Q_{liv}\cdot C_{rbc\_art} + Q_{pv}\cdot C_{rbc\_pv}\right) + PS_{rbc\_org}\cdot f_u\cdot\left(C_{\text{int}\_org} - \frac{C_{rbc\_org}}{K_{rbc}}\right)$$

Plasma + Interstitial

$$V_{\text{int}\_liv}\frac{dC_{\text{int}\_liv}}{dt} = \left(1-HCT\right)\cdot\left(Q_{liv}\cdot C_{pls\_art} + Q_{pv}\cdot C_{pls\_pv}\right) - \left(Q_{liv}+Q_{pv}\right)\cdot\left(1-HCT\right)\cdot C_{\text{int}\_liv} - PS_{rbc\_liv}\cdot f_u\cdot\left(C_{\text{int}\_liv} - \frac{C_{rbc\_liv}}{K_{rbc}}\right) - PS_{cell\_liv}\cdot f_u\cdot\left(C_{\text{int}\_liv} - \frac{C_{cell\_liv}}{K_{liv}}\right)$$

Cellular

$$V_{cell\_liv}\frac{dC_{cell\_liv}}{dt} = PS_{cell\_liv}\cdot f_u\cdot\left(C_{\text{int}\_liv} - \frac{C_{cell\_liv}}{K_{li}}\right) - \frac{CL_{liv\_met}\cdot C_{cell\_liv}\cdot f_u}{K_{liv}}$$

where $CL_{\_liv}$ is the intrinsic clearance and is calculated from the plasma clearance using the well stirred liver model.

For lungs:

$$V_{\text{int}\_lung}\cdot\frac{dC_{\text{int}\_lung}}{dt} = Q_{lung}\cdot\left(1-HCT\right)\cdot\left(C_{pls\_ven} - C_{\text{int}\_lung}\right) - PS_{rbc\_lung}\cdot f_u\cdot\left(C_{\text{int}\_lung} - \frac{C_{rbc\_lung}}{K_{rbc}}\right) - PS_{cell\_lung}\cdot f_u\cdot\left(C_{\text{int}\_lung} - \frac{C_{cell\_lung}}{K_{rbc}}\right) + Q_{vent}\cdot C_{amb\_air}\cdot P_{air} - Q_{vent}\cdot\left(\frac{C_{\text{int}\_lung}}{P_{air}}\cdot\left(1-V_{ds}\right) + C_{amb\_air}\cdot V_{ds}\right)$$

2.2.2 Mother-fetus interaction

For uterus:

Plasma +interstitial

$$V_{\text{int\_uterus}}\frac{dC_{\text{int\_uterus}}}{dt}=Q_{uterus}\cdot(1-HCT)\cdot\left(C_{pls\_art}-C_{\text{int\_uterus}}\right)-K_{d\_uterus\_pla}\cdot\left(C_{placenta}-C_{uterus\_M}\right)$$
$$-PS_{rbc\_uterus}\cdot f_u\cdot\left(C_{\text{int\_uterus}}-\frac{C_{rbc\_uterus}}{K_{rbc}}\right)-PS_{cell\_uterus}\cdot f_u\cdot\left(C_{\text{int\_uterus}}-\frac{C_{cell\_uterus}}{K_{rbc}}\right)$$

beside the assumption of equal diffusion flow from uterus to placenta and vice-versa during pregnancy, uterus behaves like other organs.

For placenta:

$$V_{\text{int\_placenta}}\frac{dC_{\text{int\_placenta}}}{dt}=K_{d\_uterus\_pla}\cdot\left(C_{\text{int\_placenta}}-C_{\text{int\_uterus}}\right)-K_{d\_pla\_amniot}\cdot\left(C_{\text{int\_placenta}}-C_{amniot}\frac{P_{placenta}}{P_{amniot}}\right)$$
$$+Q_{placenta\_fetus}\cdot\left(C_{\text{int\_art\_fetus}}-\frac{C_{\text{int\_placenta}}}{P_{placenta}}\right)+K_{glu\_deconj}\cdot Q_{placenta\_fetus}\cdot C_{\text{int\_placenta\_glu}}$$
$$-PS_{rbc\_placenta}\cdot f_u\cdot\left(C_{\text{int\_placenta}}-\frac{C_{rbc\_placenta}}{K_{rbc}}\right)-PS_{cell\_placenta}\cdot f_u\cdot\left(C_{\text{int\_placenta}}-\frac{C_{cell\_placenta}}{K_{rbc}}\right)$$

For breast:

$$V\frac{dC_{\_breast}}{dt}=PS_{\_cell\_breast}\cdot f_u\cdot\left(C_{\_\text{int\_breast}}-\frac{C_{\_breast}}{K_{\_breast}}\right)-L_{excr}$$

and the related excretion via lactation

$$L_{excr}=Q_{\_milk}\cdot\frac{C_{\_breast}}{K_{\_breast}}\cdot P_{\_milk/blood}$$

$$P_{\_milk/blood}=\frac{K_{ow}\cdot Fl_{\_tissue}+Fw_{\_tissue}}{K_{ow}\cdot Fl_{\_blood}+Fw_{\_blood}}$$

2.2.3 Lifetime scaling

The parameters related to organ volumes (***V***) and blood flows (***Q***) were taken from the ICRP (ICPR, 2002) report and fitted to time (***t in hours***) in order to derive continuous time depended non lineal polynomial formulas in the form of:

$$V = a \cdot t^b + c \cdot t^d + e \cdot t + f$$ for organ volumes

$$Q = a \cdot t^b + c \cdot t + d$$ for organ flows

Table 1. Regression coefficients for lifetime scaling (from conception to adulthood)

	organ volumes (mL)						**organ flows (mL/min)**			
	a	***b***	***c***	***d***	***e***	***f***	***a***	***b***	***c***	***d***
Portal vein	1.00E-01	1.00E+00	9.80E-02	9.96E-01	0.00E+00	5.70E+01	6.09E-02	-2.61E-02	1.06E+00	1.14E+02
Adipose	2.54E-02	1.00E+00	1.88E+01	5.20E-01	0.00E+00	9.06E+02	1.17E-01	1.00E-01	1.01E+00	3.00E+01
Bones	5.97E-02	1.00E+00	1.26E+00	6.10E-01	0.00E+00	4.52E+02	2.00E-02	1.04E-02	1.05E+00	3.00E+01
Brain	-5.03E-02	1.00E+00	9.07E-01	7.69E-01	0.00E+00	3.95E+02	-3.99E-01	7.10E-01	9.52E-01	1.80E+02
Gonads	8.25E-02	1.00E+00	8.31E-02	9.99E-01	0.00E+00	1.10E+00	3.57E-02	-3.56E-02	1.00E+00	3.00E-01
Heart	4.68E-02	1.00E+00	-3.81E-02	1.01E+00	0.00E+00	2.80E+01	4.08E-03	1.81E-05	1.41E+00	2.40E+01
Kidneys	3.17E-02	1.00E+00	1.44E-02	1.06E+00	0.00E+00	3.80E+01	3.86E-02	-7.90E-03	1.11E+00	1.10E+02
Liver	2.79E-03	1.00E+00	1.10E+00	6.03E-01	0.00E+00	1.60E+02	8.54E-03	-3.51E-04	1.24E+00	3.90E+01
GI tract	8.20E-02	1.00E+00	4.41E-02	1.04E+00	0.00E+00	9.00E+01	6.09E-02	-2.61E-02	1.06E+00	1.14E+02
Muscle	1.26E-01	1.00E+00	7.76E-06	1.76E+00	0.00E+00	9.50E+02	1.00E-01	-1.03E-01	9.92E-01	3.10E+01
Skin	2.88E-01	1.00E+00	2.71E-01	9.98E-01	0.00E+00	2.00E+02	1.06E-02	-2.72E-03	1.10E+00	3.00E+01
Lungs	9.74E-02	1.00E+00	6.33E-02	1.03E+00	0.00E+00	8.40E+01	-4.98E-01	9.94E-01	9.48E-01	5.58E+02
Arterial/venous blood	1.26E-01	1.00E+00	-1.25E-01	9.98E-01	0.00E+00	3.80E+01				
Total blood	1.15E-01	1.00E+00	-1.10E-01	9.92E-01	0.00E+00	1.33E+02				

	organ volumes (mL)						organ flows (mL/min)			
	a	***b***	***c***	***d***	***e***	***f***	***a***	***b***	***c***	***d***
Weight	3.90E-01	1.00E+00	8.40E+01	3.66E-01	0.00E+00	3.21E+03				
HCT	-1.00E-01	9.66E-12	-1.00E-01	1.55E-06	5.50E-07	5.80E-01				

Table 2. Gestation parameters (from conception to birth)

	organ volumes (mL)						organ flows (mL/min)			
	a	***b***	***c***	***d***	***e***	***f***	***a***	***b***	***c***	***d***
Uterus	3.90E-01	1.00E+00	8.40E+01	3.66E-01	0.00E+00	3.21E+03	3.57E-02	-3.56E-02	1.00E+00	3.00E-01
Placenta	3.90E-01	1.00E+00	8.40E+01	3.66E-01	0.00E+00	3.21E+03	3.57E-02	-3.56E-02	1.00E+00	3.00E-01
Amniotic fluid	3.90E-01	1.00E+00	8.40E+01	3.66E-01	0.00E+00	3.21E+03				

2.2.4 Description of absorption through multiple routes

2.2.4.1 Inhalation absorption model description

The ability to describe inhalation exposures is an important consideration for modeling inhalation pharmacokinetics. By modeling the lungs as a well-mixed compartment with an average, one-directional airflow in the region of gas exchange (i.e., with air moving through the lungs with a constant flow rate equal to the alveolar ventilation rate, Q_P), and with rapid equilibration between lung air and blood in the lung alveoli, the concentration in the blood exiting the lungs, C_a, can be described (Ramsey and Andersen, 1984; Reddy et al., 2005):

$$C_\alpha = \frac{Q_P \cdot C_{in} + Q_C \cdot C_{BLV}}{Q_C + Q_P / P_b}$$

where C_{BLV} is the chemical concentration in the venous blood compartment, P_b is the blood:air partition coefficient, Q_C is the cardiac output, and C_{in} is the inhaled concentration of chemical during the exposure and zero after the exposure ends.

To determine the concentration of chemical in exhaled air, C_{ex}, the concentration of chemical in alveolar air, C_{alv} (i.e., C_a/P_b), must be adjusted for the concentration of chemical in the dead space of the lungs (i.e., the volume of the lungs where gas exchange does not occur) as follows:

$$C_{ex} = F_{DS} \cdot C_{in} + (1 - F_{DS}) \cdot C_{alv}$$

where F_{DS} is the fraction of dead space in the lungs, which is about 0.33 in humans under typical physiological conditions.

The overall amount of air inhaled by an average adult is equal to about 20 m^3. It is clear that from the above formula that the contribution of the activity pattern clearly affects the overall intake and distribution within the tissues, through the ventilation rate Q_P and the cardiac output Q_C. When exercising, a higher amount of the body's blood rushes to the muscles to help fuel the physical activity. This surging of blood often results in a need for more oxygen in the blood than is present, and thus the body's breathing rate increases. Considering that the daily average respiratory rate for an adult ranges between 12-20 breaths per minute, it increases to 35-45 breaths per minute, depending on the burden of effort with a subsequent increase of the overall inhaled air and the corresponding uptake of the contaminants included.

2.2.4.2 Skin absorption model description

The multi-compartmental model is described using a unity of block diagram format. Each block represent a different compartment, since that type of programming enforces the clear conceptual divisions of the model's compartments. The skin has been modeled with a two layer structure: Stratum corneum (SC) and viable epidermis (VE). The stratum corneum has been described with a "bricks and mortar" structure (Touitou, 2002). The structure is presented in Figure 3. Current understanding of solute binding to keratin and other corneocyte constituents of the SC encompasses mostly equilibrium binding, and macro-scale parameterizations of transient binding for a few solutes. Proposed work leverages the latest and emerging results on homogenization theory and rates of binding to produce a broad mechanistic SC model that quantifies transient solute binding in terms of coexisting free and bound concentration fields, and is parameterized at the microscopic scale. This is critical to realistically describe actual chemical exposure, in which most of the penetration often occurs before reaching steady state.

The geometry of the microstructure has been investigated and described by many authors (Frasch and Barbero, 2003; Johnson et al., 1997; Mitragotri et al., 2011; Wang et al., 2006; Ya-Xian et al., 1999).
The characteristics of the skin are presented in Table 3.

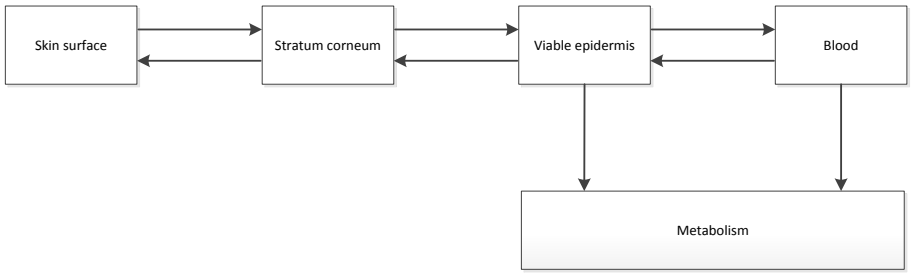


Figure 3. Skin multi-compartmental model

Table 3. Skin characteristics

Description	Symbol	Value	Unit	Reference
Number of layer	N	15	-	(Bronzino, 2000; Ya-Xian et al., 1999)
Length of corneocyte	d	30	μm	(Wang et al., 2006)
Thickness of corneocyte	t	10	μm	(Wang et al., 2006)
Length of path 1	d_1	20	μm	(Johnson et al., 1997)
Length of path 2	d_2	10	μm	(Johnson et al., 1997)
Vertical gaps	s	0.03	μm	(Johnson et al., 1997)
Horizontal gaps	g	0.03	μm	(Johnson et al., 1997)
High of viable epidermis	hv	3	mm	(Bronzino, 2000)
corneocyte edge angle	φ	90^o	degrees	(Rim et al., 2007)
Effective Diffusivity f(φ)	D_{ef}	0.002	cm^2/m	(Rim et al., 2007)

The effective tortuosity of the stratum corneum has been calculated according to the method of Johnson et al. (Johnson et al., 1997):

$$t = 1 + \frac{2g}{h} \cdot \ln\left(\frac{d}{2s}\right) + \frac{N \cdot d \cdot t}{s \cdot h} + \left(\frac{d}{1+\omega}\right)^2 \cdot \frac{\omega \cdot (N-1)}{h \cdot g}$$

Where h is the total thickness of the SC and ω is the ratio between path d1 and d2.

The model uses first order equations for the diffusion. In particular it is assumed that the layer SC retains its form and the viable epidermis is assumed as a homogenous well-mixed layer.

$$\frac{dC_S}{dt} = -D_1 \cdot C_S + D_2 \cdot C_{SC} - evap$$

$$\frac{dC_{SC}}{dt} = D_1 \cdot C_S - D_2 \cdot C_{SC} - D_3 \cdot C_{SC} + D_4 \cdot C_{ve}$$

$$\frac{dC_{ve}}{dt} = D_3 \cdot C_{SC} - D_4 \cdot C_{ve} - M - Q_b \cdot C_{ve}$$

Where C_s is the solute concentration on the skin surface that permeates through a volume of skin, C_{sc} is the solute concentration in the SC and depends on its lipid volume, lipid volume is the layer between the corneocytes and it depends on the geometry of SC, evap is the rate of the evaporation to atmosphere, C_{vs} is the permeant concentration in the viable skin, M is the rate of metabolism and Q is the blood flow rate. Moreover, D_1 and D_2 are fractions of permeate, related to the partitioning coefficient between solvent and SC, the permeant diffusivity, effective diffusivity and the geometry of SC (Nitsche et al., 2006). D_1 and D_4 are functions of the solvent/stratum partitioning coefficient, the permeant diffusivity in the SC lipid and the thickness of SC and the implementation area. D_3 and D_4 are the partition coefficient between water and SC ($K_{sc/w}$). $K_{sc/w}$ is given by the volume average (Nitsche et al., 2006):

$$K_{SC/W} = \varphi_{lip} \cdot K_{lip/w} + \varphi_{cor} \cdot K_{cor/w} \;\&\; \varphi_{lip} = 1 - \varphi_{cor}$$

Where φ_{lip} and φ_{cor} denotes the volume fractions of the lipid and corneocyte phases, respectively, with φ_{lip}=0.11 for partially hydrated SC (Nitsche et al., 2006).

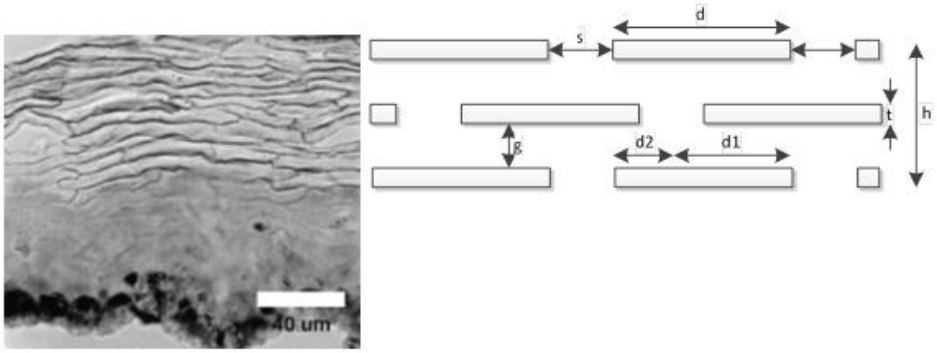


Figure 4. a) Micrograph of human stratum corneum expanded in alkaline buffer showing the stacking of corneocytes layers and the surrounding lipid matrix (Talreja et al., 2001), b) SC brick and mortal geometrical structure representation.

Also, $K_{cor/w}$ is given by volume exclusion from the fraction of the corneocyte phase occupied by keratin microfibrils. Under the assumption that there is no solute adsorption to the protein,

then $K_{o/w}=1-\varphi_f$, where $\varphi_f=0.1928$, representing the keratin microfibrils occupied fraction (Nitsche et al., 2006).

The $K_{lip/w}$ has been calculated according to (Nitsche et al., 2006). who have proposed an approximation for the predicting permeability taking into account the $K_{o/w}$:

$$K_{lip/w} = 0.43 \cdot K_{o/w}^{0.81}$$

The approximation for the predicting permeability is based on equation of Johnson et al. (1997):

$$\log k_p = 0.71 \cdot \log k_{oct} - 0.0061 \cdot MW - 6.3$$

Also, the effective lag time for the SC is given by the following relationship (Bunge and Cleek, 1995; Cleek and Bunge, 1993; Wang et al., 2006):

$$\tau_{lag} = \frac{(h_{sc})^2}{6D_{ef}}$$

Where D_{ef} is the effective diffusivity and h_{sc} is the thickness of the layer according to the geometry path length proposed by Talreja et al. (Talreja et al., 2001):

$$h_{sc} = \left(\frac{d_2}{\frac{N}{N-1}t + g} + 1 \right) \cdot h$$

where h is the total length of layer according the geometrical assumptions of the skin model.

The flow blood rate (Q) to skin layers is similar to that of a 73-kg person and it is 85.8 l/h (Abraham et al., 2005; Valentin, 2002).

Metabolism rate is given by (Abraham et al., 2005):

$$M = \left(\frac{V_{max} C_{ve}}{K_m + C_{ve}} \right)$$

Where V_{max} is the maximum rate of elimination μg/min and K_m is the permeant concentration (in venous blood) at 50% of V_{max}.

Evaporation has been calculated based on the REACH technical guidance and is given by the following equations:

$$Evaporation\ Rate = \frac{\beta \cdot MW \cdot V_p}{R \cdot T \cdot 10}$$

$$\beta = \frac{0.0111 \cdot V^{0.96} \cdot D_g^{0.19}}{v^{0.15} \cdot X^{0.04}}$$

where MW is the molecular weight of the substance, V_p is the vapor pressure of the liquid at skin temperature (in Pascal), R is the gas constant (in J/Mol/K), T is the skin temperature (in ^{0}K), β is the coefficient of mass transfer in the vapour phase (in m/h), V is the velocity of air (assumed to 0.3 m/s), D_g is the diffusivity of the liquid in the gas phase (0.05 m^2/h), v is the kinematic viscosity of air (0.054 m^2/h) and X is the length of the area of evaporation in the direction of the air stream.

2.2.4.3 Oral absorption model description

The gastrointestinal tract is modelled with four compartments: gut, stomach and their respective lumens. The parameter $Ka_{stomach}$ and Ka_{gut} govern the diffusion of chemicals in the stomach and in the gut, respectively, and then in the systemic circulation. Absorption takes place in the stomach lumen ($Rate_{Ing}$ is the ingestion rate). In gut, chemicals can be metabolized ($QMet_{gut}$):

$$\frac{dQ_{stom_{lumen}}}{dt} = RateIng - \left(F_{stom_{lumen}} + Ka_{stomach}\right) \cdot C_{stom_{lumen}}$$

$$\frac{dQ_{stomach}}{dt} = Ka_{stomach} \cdot C_{stom_{lumen}} + F_{stomach} \cdot \left(C_{art} + \frac{C_{stomach}}{PC_{stomach}}\right)$$

$$\frac{dQ_{gut_{lumen}}}{dt} = F_{stom_{lumen}} \cdot C_{stom_{lumen}} + Ke_{bile} \cdot C_{liver} - \left(Ka_{gut} + F_{gut_{lumen}}\right) \cdot C_{gut_{lumen}}$$

$$\frac{dQ_{gut}}{dt} = Ka_{gut} \cdot C_{gut_{lumen}} + F_{gut} \cdot \left(C_{art} - \frac{C_{gut}}{PC_{gut}}\right) - QMet_{gut}$$

Outputs of spleen, pancreas, stomach and gut feed liver, as well as an arterial entry. In liver, chemicals can be eliminated via bile (Ke_{bile}) or metabolized ($QMet_{liver}$).

$$\frac{\partial Q_{liver}}{\partial t} = F_{liver_{art}} \cdot C_{art} + F_{spleen} \cdot \frac{C_{spleen}}{PC_{spleen}} + F_{pancreas} \cdot \frac{C_{pancreas}}{PC_{pancreas}} + F_{gut} \cdot \frac{C_{gut}}{PC_{gut}}$$
$$+F_{stomach} \cdot \frac{C_{stomach}}{PC_{stomach}} - F_{liver} \cdot \frac{C_{liver}}{PC_{liver}} - Ke_{bile} \cdot C_{liver} - QMet_{liver}$$

The sorting blood flow in liver is then given by:

$$F_{liver} = F_{liver_{art}} + F_{gut} + F_{pancreas} + F_{spleen} + F_{stomach}$$

3 Biokinetic interaction of mixtures

3.1 General aspects related to mixtures interactions at the level of metabolism

Single chemical exposure is an exception rather than the rule in the general and occupational environments. Humans are generally exposed to chemical mixtures resulting in a combined effect on human health. Biokinetic and toxicodynamic interactions can result effects beyond simple additivity of the mixture components. For instance, synergistic or antagonistic effects caused by perturbation of metabolism efficiency may in turn lead to enhanced or reduced toxicity of chemical mixtures compared to the toxicity expected based on knowledge of the potency and dose of the single constituents (additivity). Biokinetic interactions lead to a change in tissue dose of chemicals during exposure to mixtures. They are the most common type of interaction observed and reported in the literature and the extent of the change in tissue dose of chemicals resulting from biokinetic interactions during mixture exposures depends on the concentrations of all components and the mechanism(s) of interactions.

Table 4 provides definitions of terms used in describing the results of interactions studies according to the US Agency for Toxic Substances and Disease Registry (ATSDR).

Table 4. Interactions Terminology [a,b]

Interaction	When the effect of a mixture is different from additivity based on the dose-response.
Additivity	When the effect of the mixture can be estimated from the sum of the exposure levels (weighted for potency) or the effects of the individual components.
No apparent influence	When a component which is not toxic to a particular organ system does not influence the toxicity of a second component on that organ system.
Synergism	When the effect of the mixture is greater than that estimated for additivity on the basis of the toxicities of the components.
Potentiation	When a component that does not have a toxic effect on an organ system increases the effect of a second chemical on that organ system.
Antagonism	When the effect of the mixture is less than that estimated for additivity on the basis of the toxicities of the components.

Inhibition	When a component that does not have a toxic effect on a certain organ system decreases the apparent effect of a second chemical on that organ system.

[a] Where effect is incidence or measured response, and additivity commonly is dose or response additivity.

[b] Based on definitions in EPA (2000), Hertzberg et al. (1999), and Mumtaz and Hertzberg (1993).

The feasibility of using PBBK models to describe, predict, and extrapolate the extent and magnitude of the occurrence of interactions for various dose levels, scenarios, species, routes, and mixture complexities arises from the very nature and the scientific basis of these models. The approach we follow to develop generic multi-PBBK models of mixtures is based on sets of multi-compartment single-chemical models developed and validated that will be integrated through explicit mathematical description of the interaction processes outlined in Table 4 above and detailed in the text below.

To this aim a further Extended Literature Search will was out to identify the mechanisms of action (MoA) so as to allow us to identify the type of mechanism of interactions among the chemicals constituting the mixture and finally translate it into the appropriate mathematical formulation.

3.2 Types of interaction

Multiple types of interactions have to be taken into account, based on the combination of compounds that have to be tackled each time including:

- Physicochemical Interactions

Co-exposures can lead to modification of a chemical's solubility in lipid or water or even in its permeability across biological membranes, resulting in changes in distribution or rates of absorption. For example, some chemicals, such as antacids, can alter the pH of gastric or intestinal regions and, therefore, influence the ionization state of other acidic or basic chemicals, leading to altered absorption rates. For some chemicals, their joint presence leads to formation of complexes which are more lipophilic than any of the individual chemicals, resulting in an increase in their permeability across membranes (e.g., increased permeability across blood–brain barrier of lead-dithiocarbamate complexes).

- Reversible Protein Binding Interaction

Mechanisms of interactions at the level of protein binding might involve either competition for the binding site or induction of binding protein levels. This interaction is described by the competitive inhibitor effect on the dissociation constant.

- Irreversible Metabolic Inhibition

When an inhibitor binds irreversibly to the enzyme at the active site, it decreases the concentration of functional enzymes and thus the V_{max}. Briefly, the modeling of this type of interaction consists of describing the change in V_{max} of the substrate by taking into account the enzyme inactivation in relation to the concentration of the inhibitor. Accordingly, the concentration of the active enzyme at steady state (i.e., in the absence of inactivator) equals

the rate of enzyme synthesis (a zero order process) minus the rate of enzyme degradation (first order process that can be determined experimentally). This steady state is interrupted in the presence of an enzyme inactivator and the levels of enzyme is reduced by a rate of enzyme inactivation that is directly related to the concentration of the inactivating chemicals in the liver and other metabolism sites specific to the compounds of interest.

- Induction

Enzyme induction increases the V_{max} of its substrate, due to increased enzyme synthesis and/or decreased enzyme degradation.

- Reversible Metabolic Inhibition

Among all the mixture PBBK models published to date, reversible metabolic inhibition is by far the most frequently encountered type of interaction. There are three types of reversible enzyme inhibition, namely competitive, non-competitive, and uncompetitive. Competitive inhibition occurs when chemicals compete for the enzyme active site, resulting in decreased apparent affinity (i.e., increased K_m) and, therefore, reduced rate of metabolism at lower substrate concentration. Non-competitive inhibition occurs when a chemical binds to the enzyme (free or complexed with substrate) at a site that is away from the catalytic active site. This binding changes the conformation of the enzyme, resulting in a decreased catalytic activity (i.e., decreased V_{max}). The less frequently encountered uncompetitive inhibition occurs when a chemical binds to the enzyme-substrate complex. The catalytic function is affected without interfering with substrate binding. The inhibitor causes a structural distortion of the active site and inactivates it. This has the effect of reducing the available enzyme for the reaction (i.e., lowering V_{max}) and also driving the reaction (E+S→ES) to the right (i.e., decreasing K_m).

3.3 Modelling interactions

We shall follow a three-prone approach to the development of TK models for human health risk assessment of multiple chemicals:

3.3.1 Identifying interaction terms

The first and main approach for modeling biokinetic interactions in mixtures involves identifying and linking all individual chemical models via interaction terms. In this light PBBK models for mixtures of any level of complexity can then be created as long as the quantitative information on the mechanism for each interacting chemical pair is available or can be hypothesized. According to this methodology, only plausible binary interactions need to be characterized to model the kinetics of the interactions between components of complex mixtures. In a mixture of three chemicals, for example, there are three two-way interactions (Figure 5). The first step will be to properly parameterize the generic model for each component of the mixture. Then the single-chemical models should be interconnected at the binary level by modifying the appropriate equations. For example considering competitive metabolic inhibition as the mechanism of interaction in a quaternary mixture of chemicals A,

B, C and D, the equation for calculating the rate of the amount metabolized (RAM) of each component will be modified from the case of a single chemical:

$$RAM_A = \frac{V\max_A \cdot Cvl_A}{Km_A \cdot Cvl_A}$$

to the followings equations:

$$RAM_A = \frac{V\max_A \cdot Cvl_A}{Km_A \cdot (1 + \frac{Cvl_B}{Ki_{BA}} + \frac{Cvl_C}{Ki_{CA}} + \frac{Cvl_D}{Ki_{DA}}) + Cvl_A}$$

$$RAM_B = \frac{V\max_B \cdot Cvl_B}{Km_B \cdot (1 + \frac{Cvl_A}{Ki_{AB}} + \frac{Cvl_C}{Ki_{CB}} + \frac{Cvl_D}{Ki_{DB}}) + Cvl_B}$$

$$RAM_C = \frac{V\max_C \cdot Cvl_C}{Km_C \cdot (1 + \frac{Cvl_A}{Ki_{AC}} + \frac{Cvl_B}{Ki_{BC}} + \frac{Cvl_D}{Ki_{DC}}) + Cvl_C}$$

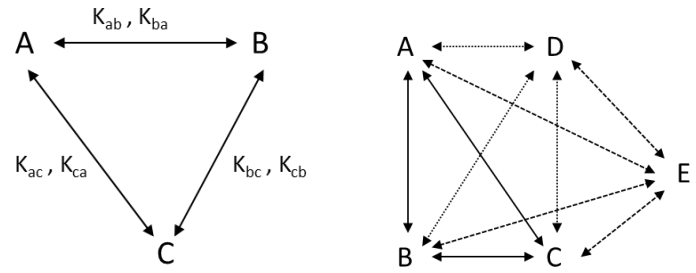
$$RAM_D = \frac{V\max_D \cdot Cvl_D}{Km_D \cdot (1 + \frac{Cvl_A}{Ki_{AD}} + \frac{Cvl_B}{Ki_{BD}} + \frac{Cvl_C}{Ki_{CD}}) + Cvl_D}$$

where Cvl_I *is the* venous blood concentrations of chemical I leaving liver compartments, Ki_{IJ} is the constant describing competitive inhibition of the metabolism of chemical J by chemical I, $Vmax_I$ is the maximal velocity of metabolism of chemical I, Km_I is the Michaelis affinity constant of chemical I.

Logically, this PBBK modeling approach should be applicable to predict higher order interactions in mixtures of any complexity, invoking various kinds of mechanisms of interactions among components. It is important to note that all linkages involving mixture components are of binary nature only.

This approach is where the unique usefulness of PBBK modeling becomes evident. Let us assume that the binary chemical interaction between A and B has been modeled. Following the addition of another chemical, C, the PBBK model not only simulates the binary interactions involving C (i.e., C-A, C-B) but also the modulatory effect of C on the interaction between A and B. Once we describe for instance the inhibitory effect of C on B, this would result in a reduction in the rate in which B gets metabolized and consequently an increase in its concentration in the venous blood leaving the metabolism site (e.g. the liver). Because the

exposure to chemical increases concentration in venous blood leaving the liver, this translates into a modification of the magnitude of the interactive effect of B on A. Similarly, C may also affect the concentration of A, which would then result in a change in the magnitude of the interactive effect of A on B. The PBBK model framework can also simulate similar phenomena affecting the concentration of C because all components of the mixture are linked. Based on this analogy it will be possible to predict the influence of the addition of another chemical, D, to the ternary mixture, and so forth. When D is added to an existing ternary mixture PBBK model for chemicals A, B, and C, we need to consider only three additional binary interactions (Figure 5a). By doing this, the modulating effect of D on the C-A and B-A interactions will be automatically simulated because all components are linked with each other within the PBBK framework. The effect of D on the kinetics of A will in turn affect the kinetics of B, C, and D. Any modulation of a binary interaction will affect the kinetics of other chemicals that are part of the network of binary interactions present in the mixture. The same considerations are applicable when another chemical, E, is added to the quaternary mixture. After adding the four new binary interactions (E-A, E-B, E-C, and E-D), chemical E will become an integral part of the network of the components of the mixture and any modulation of a binary interaction involving E will have repercussions on all the others (Figure 5b). The interesting thing about this approach is that it only requires data on binary interaction mechanisms for predicting the magnitude and consequence of higher order interactions within complex mixtures.



(a) (b)

Figure 5: Interactions among the components of a three (a) and a five (b) compound mixture.

Logically, this PBBK modeling approach should be applicable to predict higher order interactions in mixtures of any complexity, invoking various kinds of mechanisms of interactions among components. It is important to note that all linkages involving mixture components are of binary nature only, yet they allow tackling mixtures of increasing complexity encompassing a relatively large number of constituents (on the order of 20-30 compounds; the actual level of mixture complexity that can be handled with the binary interaction approach is limited by the computational power of the IT system used).

3.3.2 Lumping approach

When the mixture involves a larger number of chemicals, which does not easily allow the application of the framework illustrated above, (i.e. beyond 20-30 compounds) a top-down approach will be used. The top-down or lumping approach will start with considering the mixture as a single entity and proceed to splitting out the chemicals for which specific information is required, such as the biokinetics of a chemical that is the main cause of the effect of interest. This approach simplifies the problem by isolating specific components for which a description is helpful and by treating the remaining components as a single, lumped chemical. This approach has been already applied to a complex mixture of hydrocarbons (Dennison et al., 2004) showing that although model accuracy is reduced to some extent with a chemical-lumping approach, the model still provides a good representation of the kinetics of five isolated chemicals (n-hexane, benzene, toluene, ethylbenzene, and o-xylene) during exposure to various levels of two different gasoline blends. When appropriate kinetics data are available for model development, the approach can be applied to other mixtures especially when the relevant properties of each component are similar enough to be described as a whole.

4 Generic PBBK model parameterization

PBBK models are data intensive, requiring several parameters (mainly partition coefficients and clearance kinetics) for proper parameterization. These parameters are chemical specific, and they vary among the different chemical classes, as well as among the several chemicals of the same class. For data-rich chemicals, where human PBPK models exist, data will be stored in the HEALS database, or the parameters might be retrieved via external databases. For data poor chemicals, data will be calculated by Quantitative Structure–Activity Relationship (QSAR models).

4.1 Use of databases

Specific data for PBPK models parameterization could be retrieved from the following databases:

4.1.1 PoPGen

PopGen aims to generate data that predict realistic anatomical and physiological variation in human populations. Organ volumes and blood flows are determined for virtual individuals from both a priori distibutions of anthropometric parameters such as body mass, height, and body mass index, and from measured data from existing studies. The main aim is to describe and to incorporate inter-individual variability in the PBPK models.
Link: http://xnet.hsl.gov.uk/popgen/

4.1.2 Simcyp

The free toolbox of SimCyp provides essential information for PBBK models parameterization, such as
- the concentration at which the maximum intrinsic clearance occurs as well as the ratio of this intrinsic clearance to the corresponding intrinsic clearance based on Michaelis-Menten kinetics.

- the fraction of the compound unbound in plasma
- age dependent hepatic scaling factors

Link: http://www.simcyp.com/

4.1.3 EpiSuite

This database includes data on the physicochemical properties of the compounds such as octanol-water partition coefficient, Henry's Law constant, vapor pressure. These parameters are essesntial for QSARs calculation for PBBK models parameterization.
Link: http://www.epa.gov/opptintr/exposure/pubs/episuite.htm

4.2 Use of Quantitative Structure–Activity Relationship (QSAR models)

4.2.1 Rationale

QSAR models are regression or classification models used in the chemical and biological sciences and engineering. QSARs form a relationship between biological effects and chemistry of each chemical and comprise three parts:

- The activity data to be modeled,
- The data with which to model and
- A method to formulate the model.

The biological effects are normally the property to be modeled, which are linked with the physical or structural chemistry of the molecules. QSAR methods are used particularly for the estimation of physicochemical properties, biological effects as well as understanding the physicochemical features governing a biological response.

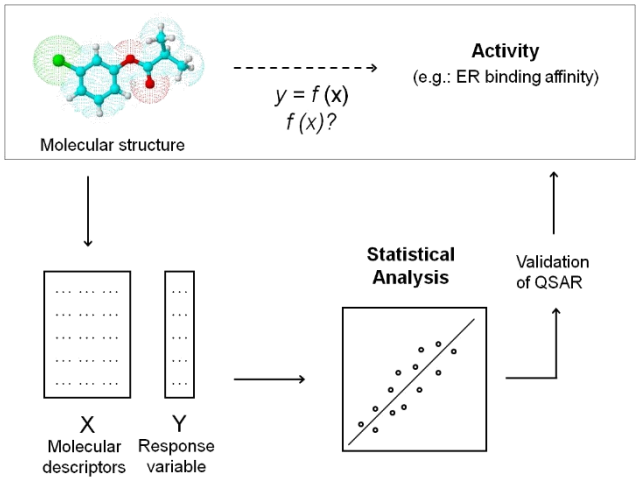


Figure 6. Basic methodological scheme for deriving a QSAR

The input parameters required for solving the set of PBBK model equations are either species- specific or chemical-specific and should reflect biological or mechanistic determinants of absorption, distribution, metabolism and elimination (ADME) of the chemical being modeled. The species-specific parameters, for example, relate to alveolar ventilation rate (Q_p), cardiac output (Q_c), tissue blood flow rates (Q_t) and tissue volumes (V_t) should be within the documented range for the particular species and life stage. The chemical-specific input include partition coefficients (blood:air (P_{ba}), tissue:air (P_{ta}) or tissue:blood (P_{tb})) as well as metabolic parameters such as the maximal velocity (V_{max}) and Michaelis affinity constant (K_m) or the intrinsic clearance (V_{max}/K_m). These physicochemical parameters should have been obtained on the basis of independent measurements (in vitro, in vivo) or using algorithms in the valid domain of application, like QSARs. In particular, in the case of a chemical for which biokinetic parameter database is either incomplete or lacking, the internal dose cannot be reliably estimated. The internal dose measure associated with a particular exposure scenario could vary from anywhere between zero (theoretical minimum) and the

potential dose (theoretical maximum). This large uncertainty is due to the fact that there is a lack of precise knowledge regarding the key chemical-specific determinants of ADME. Since these parameters, together with the physiology of the animal species, determine the biokinetics of chemicals in biota, integrated QSAR-PBBK models can effectively predict or identify the possible range of internal dose.

4.3 QSARs models

4.3.1 Peyret, Poulin and Krishnan algorithm for estimating tissue:blood partition coefficients

A unified algorithm by Peyret, Poulin and Krishnan has been examined, which is applied both for environmental chemicals and drugs and predict the rat tissue:blood, tissue:plasma water and tissue:plasma partition coefficients for liver, muscle and adipose tissue (Peyret et al., 2010). The tissue:blood, tissue:plasma water and tissue:plasma partition coefficient can be computed from matrix:water partition coefficients as follows:

$$P_{tb} = \frac{P_{ct} \cdot F_{ct} + P_{it} \cdot F_{it}}{P_p \cdot F_p + P_e \cdot F_e}$$

$$K_{pu} = P_{ct} \cdot F_{ct} + P_{it} \cdot F_{it}$$

$$K_p = \frac{P_{ct} \cdot F_{ct} + P_{it} \cdot F_{it}}{P_p \cdot F_p}$$

where P_{ct} is the tissue cell:water partition coefficient, P_{it} is the tissue interstitial fluid:water partition coefficient, P_p is the plasma:water partition coefficient, P_e is the erythrocyte:water partition coefficient, F_{ct} is the fractional content of cells in tissue, F_{it} is the fractional content of interstitial fluid in tissue, F_p is the fractional content of plasma in blood and F_e is the fractional content of erythrocyte in blood.

The matrix:water partition coefficients can be calculated by the following equation:

$$P_{mw} = \frac{(1+I_m) \cdot F_{wm} + P_{ow} \cdot F_{nlm} + I_m \cdot P_{aplw} \cdot F_{aplm} + (1+I_m) \cdot P_{prw} \cdot F_{prm}}{(1+I_w)}$$

where P_{mw} is the matrix:water partition coefficient, I_m and I_w is the ionization term for the matrix and water, respectively, F_{wm} is the fractional volume of water equivalent in the matrix, F_{nlm} is the fractional volume of neutral lipids equivalent in the matrix, F_{aplm} is the fractional volume of acidic phospholipids in the matrix, F_{prm} is the fractional volume of binding proteins in the matrix, P_{aplw} is the acidic phospholipids:water partition coefficient, P_{prw} is the protein:water partition coefficient and P_{ow} is the vegetable oil:water, which is calculated using the n-octanol:water partition coefficient (Bartels et al., 2012):

$$\log P_{oil:water} = -4.653 + \frac{7.972}{\left(1+10^{(0.1175-0.2849 \cdot \log P_{oct:water})}\right)}$$

The ionization term I_m depends on the category of the chemical, if it is neutral, monoprotic base, diprotic acid and is calculated by the Henderson-Hasselbach equations, while the protein:water partition coefficient is equal to zero for tissue cells because it is assumed that binding to the proteins is of minor relevance.

For interstitial fluid and plasma, it is computed by the equation below:

$$P_{prw} = \left(\frac{1}{fu_p} - 1 - \frac{P_{ow} \cdot F_{nlp}}{(1 + I_p)} \right) \cdot \frac{1}{F_{prp}}$$

where f_{up} is the unbound fraction in plasma, F_{nlp} is the fractional content of neutral lipids equivalent in plasma and F_{pr} is the fraction of albumin for the acidic compounds and weak bases or lipoprotein for the neutral compounds in plasma.

For erythrocyte, P_{prw} refers to the hemoglobin:water partition coefficient only for relatively hydrophobic VOCs, with $\log P_{ow}>1$ and a low molecular volume,

$$P_{Hbw} = \frac{\left(\frac{P_{bw} - F_p \cdot P_p}{F_e} \right) \cdot (1 + I_w) - \left((1 + I_e) \cdot F_{we} + P_{ow} \cdot F_{nle} + I_e \cdot P_{aplw} \cdot F_{aple} \right)}{F_{pre}}$$

where P_{bw} is the blood:water partition coefficient, which is determined by the ratio between blood:air partition coefficient and water:air partition coefficient.

The term of acidic phospholipid:water partition coefficient of tissue cells is used only for strong basic compounds and basic zwitterions with at least one pKa>7 and is computed using the following equation:

$$P_{aplw} = \left[P_{ew} - \frac{(1 + I_e) \cdot F_{we} + P_{ow} \cdot F_{nle}}{1 + I_p} \right] \cdot \frac{1 + I_p}{I_e \cdot F_{aple}}$$

The plots of four of the chemicals examined are presented at the following figure. They are both environmental chemicals, such as Benzene and DEHP and drugs, such as MeEtBarbiturate and Cyclosporine.

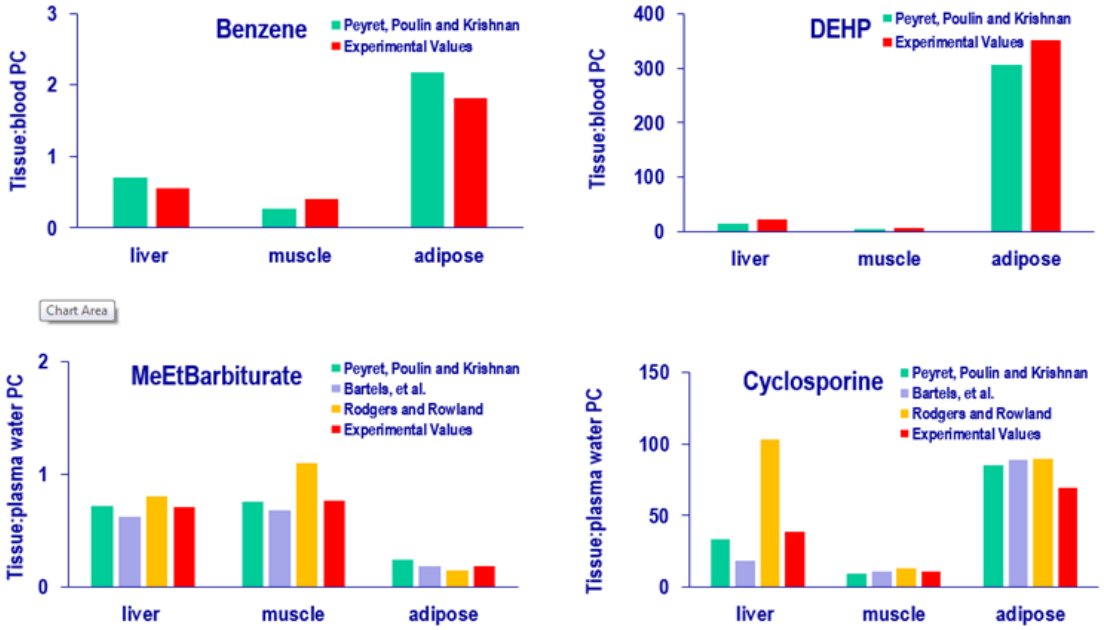


Figure 7. Partition coefficients of the examined chemicals.

The predicted values are very close to the experimental ones obtained from literature (Ballard et al., 2003; Keys et al., 1999; Rodgers and Rowland, 2006; Tanaka et al., 2000; Zhang, 2004). Generally, the algorithm applies quite well to acidic compounds and neutrals. It should be mentioned that there have to be improvements regarding the prediction of strong bases' parameters.

4.3.1.1 QSARs based on molecular fractions for estimating biological properties

The QSARs coefficients for V_{max} and K_m are determined based on data for 53 volatile organic compounds (VOCs). Following the methodology proposed by Béliveau et al. (Béliveau et al., 2003), QSAR fragment contributions (C_i) are determined by Artificial Neural Networks (ANN), using experimental values of V_{maxc} and K_m (Price and Krishnan, 2011), as follows:

$$\log V_{\max c} = \sum f_i \cdot C_i$$

where f_i is the frequency of occurrence of the fragment i (CH_3, CH_2, CH, C, C=C, H, Cl, Br, F, benzene ring, and H in benzene ring structure) in a given molecule and C_i is the contribution of each fragment.

In the case of V_{max}, it should be mentioned that the QSAR modelling is based on the maximal velocity normalized to the body weight (BW) of the animal, using the following allometric relationship:

$$V_{\max} = V_{\max c} \cdot BW^{0.75}$$

The average body weight of human was considered equal to 75kg. The experimental and predicted values of the logarithm of the body weight normalized maximal velocity of enzymic metabolism, V_{maxc}, are presented in Figure 8.

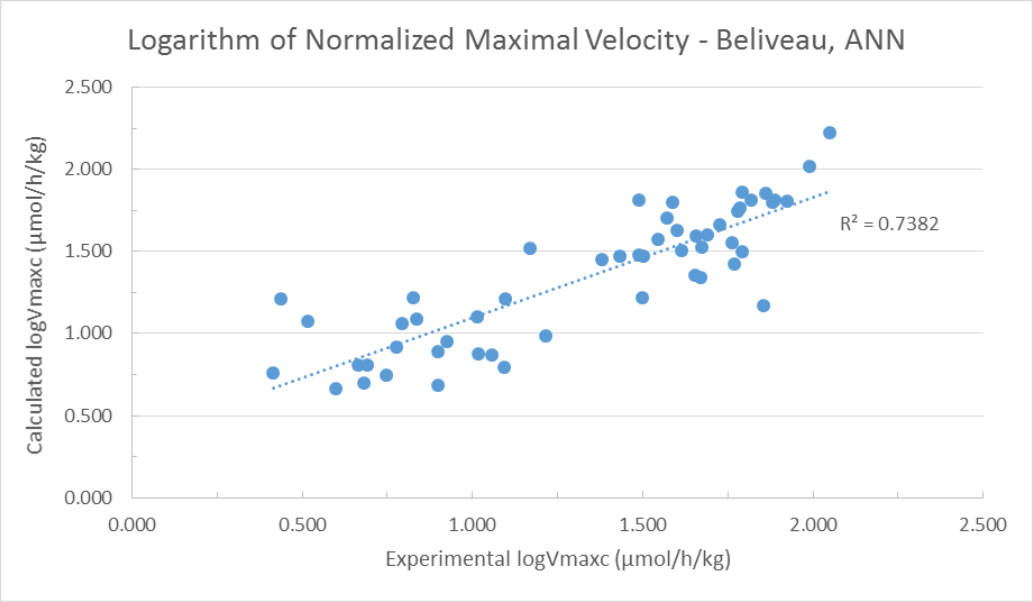


Figure 8. Experimental vs predicted logarithm of body weight normalized maximal velocity for selected volatile compounds.

4.3.1.2 Abraham's solvation equation for estimating biological properties

Abraham's solvation equation is used to calculate several biological properties for a series of solutes:

$$\log SP = c + r \cdot R_2 + s \cdot \pi_2^H + a \cdot \Sigma\alpha_2^H + b \cdot \Sigma\beta_2^H + v \cdot \log V_x$$

where SP is a biological property, R_2 is an excess molar refraction that can be determined simply from a knowledge of the compound refractive index, π_2^H is the compound dipolarity/polarizability, $\Sigma\alpha_2^H$ is the solute effective or summation hydrogen-bond acidity, $\Sigma\beta_2^H$ is the solute effective or summation hydrogen-bond basicity, V_x is the McGowan characteristic volume that can trivially be calculated for any solute simply from a knowledge of its molecular structure (Abraham, 1993).

Because the constants in Abraham's equation represent quite specific properties of the phase or receptor area, they should follow correct chemical principles. Thus, Abraham's equation is not simply a statistical fitting procedure, but is substantive equation expressing not only the effect of solutes on some particular process, but also the properties of the solvent phase or receptor area involved. QSARs derived from Abraham's equation have to be examined with regard to goodness-of-fit, as is the case for any QSAR, but also with regard to general chemical principles. This latter test is highly unusual in QSAR work, but is

very important in that strict application of the test leads to QSARs that are chemically firmly based, and are not just fitting equations to a given data set (Abraham, 1993).

4.4 Coupled Abraham's solvation equation parameters to Non Linear Regression (NLR) and Artificial Neural Networks (ANNs) models

4.4.1 Method description

The Abraham's equation can be solved by statistical methods to yield the constants c, r, s, a, b and v. These constants are important and can be used to characterize the receptor area involved. The r-constant gives the propensity of the phase to interact with solute x; - and n - electron pairs, the s-constant is the phase-area dipolarity/polarizability, the a-constant is the phase-area basicity (because a basic phase will interact with acid solutes), similarly the b-constant is the phase-area acidity, and the v-constant is a measure of the phase-area lipophilicity.

The chemoinformatic methods used in building QSAR models can be divided into three groups, i.e., extracting descriptors from molecular structure, choosing those informative in the context of the analyzed activity, and, finally, using the values of the descriptors as independent variables to define a mapping that correlates them with the activity in question. The selected statistical methods include non-linear models, which extend the structure-activity relationships to non-linear functions of input descriptors. Such models may become more accurate, especially for large and diverse datasets. However, usually, they are harder to interpret. Complex, non-linear models may also fall prey to overfitting, i.e., low generalization to compounds unseen during training (Dudek et al., 2006).

The method of Non Linear Regression (NLR) is used in order to calculate the constants of Abraham's equation. In statistics, nonlinear regression is a form of regression analysis in which observational data are modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables. The physicochemical properties of Abraham's equation, R_2, π_2^H, $\Sigma\alpha_2^H$, $\Sigma\beta_2^H$ and V_x, are obtained from literature for 34 chemical compounds and are estimated as inputs of the model. The experimental values of the tissue:blood partition coefficients and the kinetic parameters are also found for the same compounds and are used as targets of the model. The method of least squares (LS) is chosen in order to find the constants of Abraham's equation and consequently the predicted values of the chemical compounds. The LS tries to indirectly obtain knowledge on the latent variables by decomposing the input matrix of descriptors into two components, the scores and the loadings. The scores are orthogonal and, while being able to capture the descriptor information, allow also for good prediction of the activity. The estimation of score vectors is done iteratively. The first one is derived using the first eigenvector of the activity descriptor combined variance-covariance matrix. Next, the descriptor matrix is deflated by subtracting the information explained by the first score vector. The resulting matrix is used in the derivation of the second score vector, which followed by consecutive deflation, closes the

iteration loop. In each iteration step, the coefficient relating the score vector to the activity is also determined (Dudek et al., 2006).

Another examined statistical method is Artificial Neural Networks (ANN), which is biologically inspired prediction method based on the architecture of a network of neurons. In this occasion, the whole data set is divided into training and test sets (85% and 15% respectively of the total number of compounds) and the models developed from training set is externally validated using the test set. The examined QSAR model has been analyzed with perceptron-based method, which falls into the category of feed-forward network, in which, during the prediction, the information flows only in direction from the input descriptors, through a set of layers, to the output of the network. The Multi-Layer Perceptron (MLP) model consists of a layered network of interconnected perceptrons, i.e., simple models of a neuron. Each perceptron is capable of making a linear combination of its input values and, by means of a certain transfer function, produce a binary or continuous output. A noteworthy fact is that each input of the perceptron has an adaptive weight specifying the importance of the input. In training of a single perceptron, the inputs of the perceptron are formed by the molecular descriptors, while the output should predict the activity of the compound. To achieve this goal, the perceptron is trained by adjusting the weights, to produce a linear combination of the descriptors that optimally predicts the activity. The adjusting process relies on the feedback from comparing the predicted with the expected output. That is, the error in the prediction is propagated to the weights of the descriptors, altering them in the direction that counters the error. While a single perceptron is a linear model, a network consisting of layers of perceptrons, with output of one layer connected to inputs of neurons in consecutive layer, allows for non-linear prediction. Multi-layer networks contain a single input layer, which consists simply of the values of molecular descriptors, one or more hidden layers, which process the descriptors into internal representations and an output layer utilizing the internal representation to produce the final prediction (Dudek et al., 2006). The statistical qualities are judged and selected by the square of correlation coefficient (R^2).

4.4.2 Results – comparison to existing methods

As mentioned above, the Abraham's equation is solved by Non Linear Regression (NLR) and Artificial Neural Networks (ANN), on the basis of experimental values (Baláž and Lukáčová, 1999; Paixão et al., 2013; Pelekis and Krishnan, 2004; Price and Krishnan, 2011), in order to predict the tissue:blood partition coefficients and Michaelis – Menten constants of a large variety of chemicals. Figure 9 and Figure 10 present the experimental and predicted values of the logarithm of liver:blood partition coefficients by using NLR and ANN, respectively.

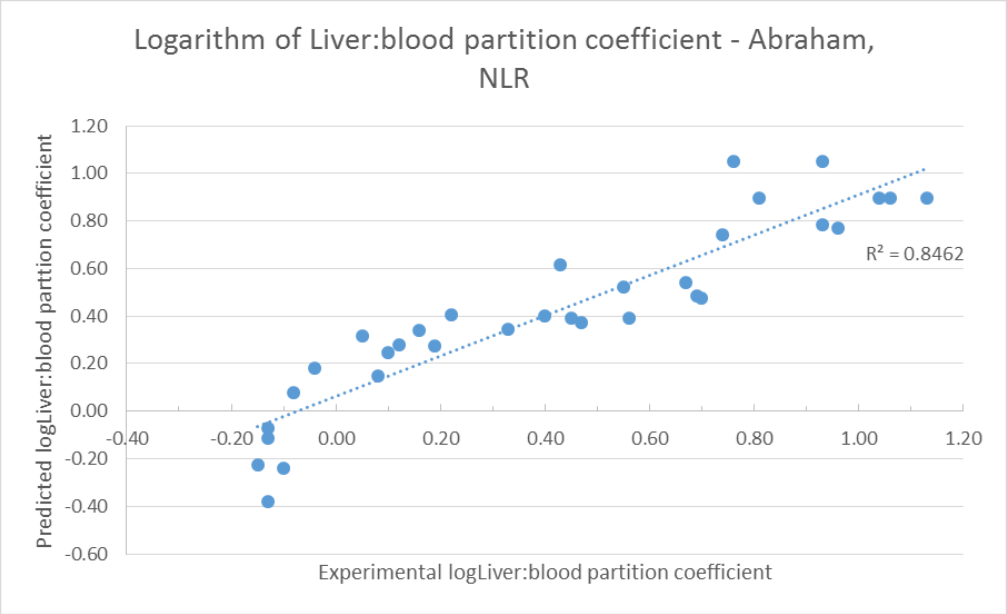


Figure 9. Experimental vs predicted logarithms of liver:blood partition coefficients for selected low volatile compounds, using NLR.

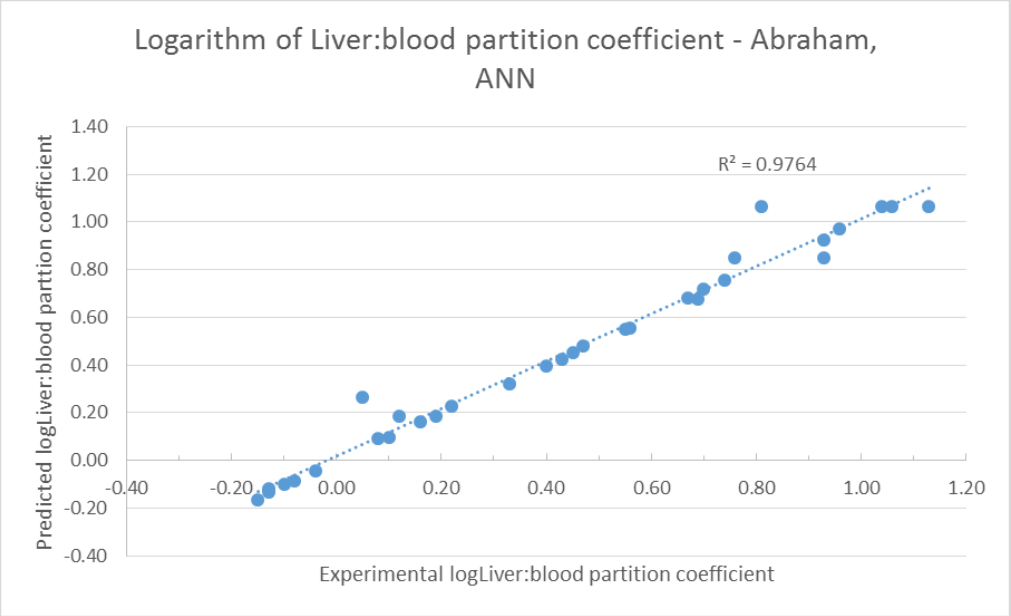


Figure 10. Experimental vs predicted logarithms of liver:blood partition coefficients for selected low volatile compounds, using ANN.

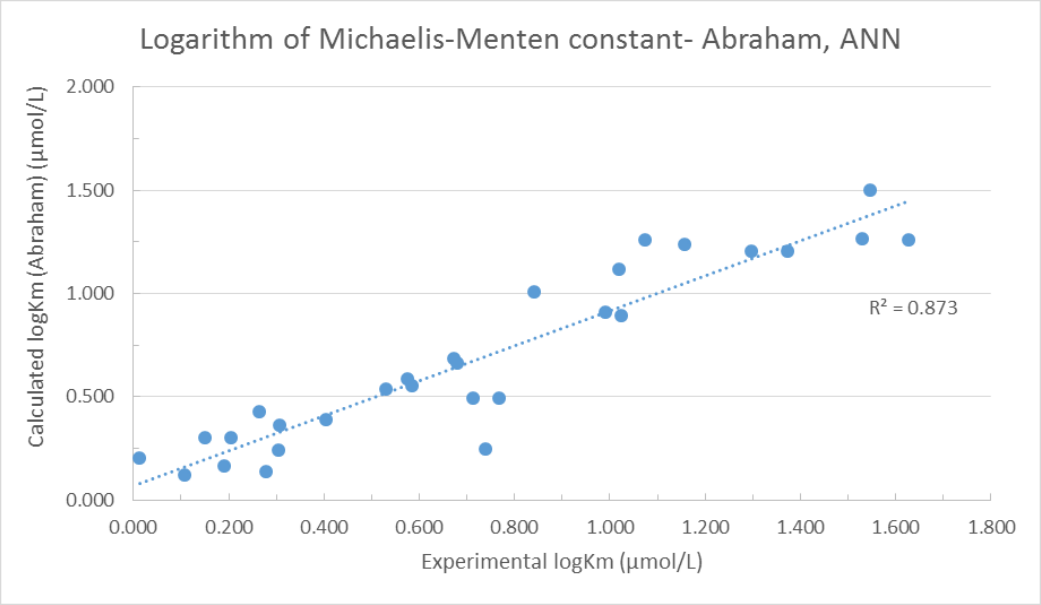


Figure 11. Experimental vs predicted logarithm of Michaelis - Menten constant for selected low volatile compounds.

As shown in the figures above, the correlation between the experimental and predicted values obtained using Artificial Neural Networks is better than that using Non Linear Regression. Although, both the NLR and ANN give high values of the square of correlation coefficient, R^2.

In Figure 11, the experimental and predicted values of logarithm of Michaelis - Menten constant K_m are presented. Up to now, these QSARs seem to perform well for a number of chemical families, relevant to the aims of HEALS.

A major breakthrough came from the use of Artificial Neural Networks coupled to Abraham's solvation equation for predicting biological/biochemical properties such as tissue:blood partition coefficients, maximal velocity of enzymic metabolism (V_{max}) and Michaelis - Menten constant (K_m).

This is a remarkable advance, since till now, prediction capability of the Michaelis - Menten constant was rather poor (R^2 up to 0.35); with our coupled ANN - Abraham's solvation equation method for the investigated group of chemicals R^2 went up to 0.87 (Figure 11), which is by far higher to any other existing methodology.

The improved performance of ANN-Abraham's equation combination can be ascribed to its capacity to represent mathematically the complex interactions of biochemical micro-processes, which are lumped into the Michaelis-Menten constant.

A training set of 33 chemicals is used in order to predict the partition coefficients for four major parts of the human body, the adipose, kidney, liver and brain tissue. The selected training set resulted in predicted values with marginal differences compared to the experimental ones. The square of the correlation factor, R^2, ranges from 0.83 for lung to 0.99

for adipose tissue. The test set is consisted of 150 chemicals, whose experimental tissue:blood partition coefficients are unknown.

In order to compare the results obtained from Abraham's equation, a simpler QSAR model is used (Poulin and Krishnan, 1995):

$$P_{tissue/blood} = \frac{P_{ow} \cdot Fl_{tissue} + Fw_{tissue}}{P_{ow} \cdot Fl_{blood} + Fw_{blood}}$$

where P_{ow} is the octanol:water partition coefficient, Fl_{tissue} is the fractional content of lipids in tissue, Fw_{tissue} is the fractional content of water in tissue, Fl_{blood} is the fractional content of lipids in blood and Fw_{blood} is the fractional content of water in blood.

The comparison between the tissue:blood partition coefficients, predicted with a) Pow equation and b) Abraham' equation, is presented in Figures 12-15.

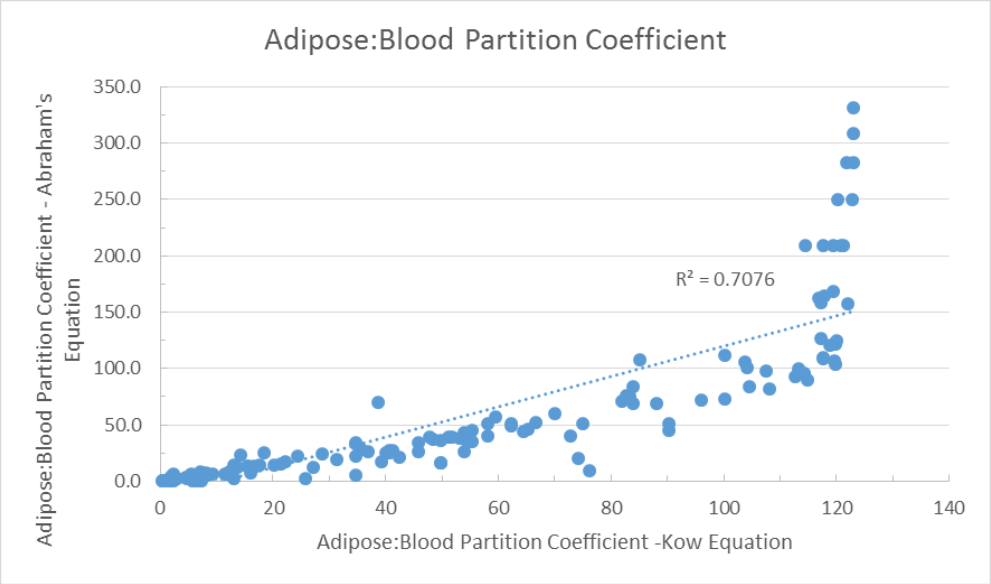


Figure 12. Predicted adipose:blood partition coefficients using Pow equation and Abraham' equation.

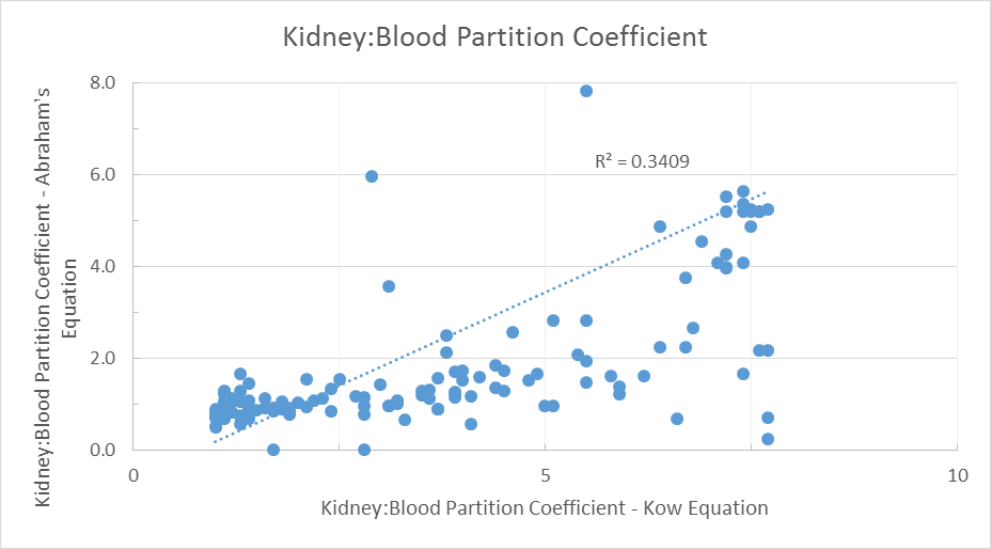


Figure 13. Predicted kidney:blood partition coefficients using Pow equation and Abraham' equation.

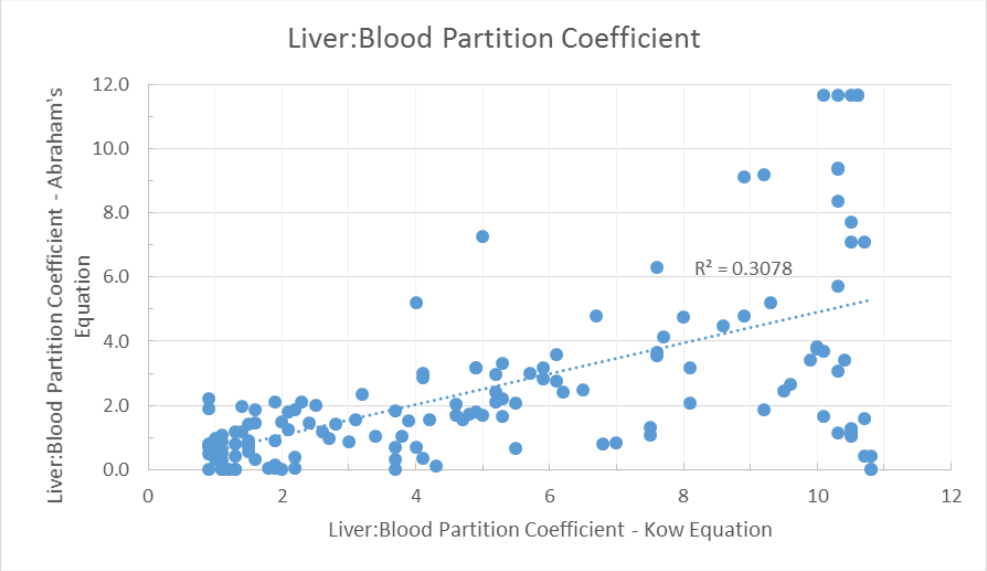


Figure 14. Predicted liver:blood partition coefficients using Pow equation and Abraham' equation.

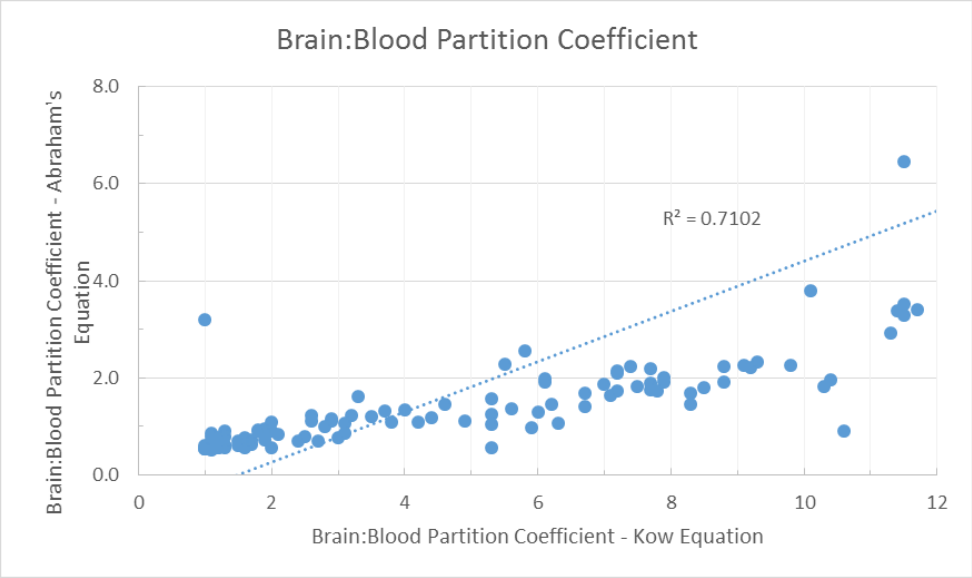


Figure 15. Predicted brain:blood partition coefficients using Pow equation and Abraham' equation.

The chemicals of the test set are grouped by their chemical structure. The tissue:blood partition coefficients for each chemical group are calculated with the two methods described above. The estimated correlation coefficient for each group is presented in the following figures (Figures 16-19) for every tissue.

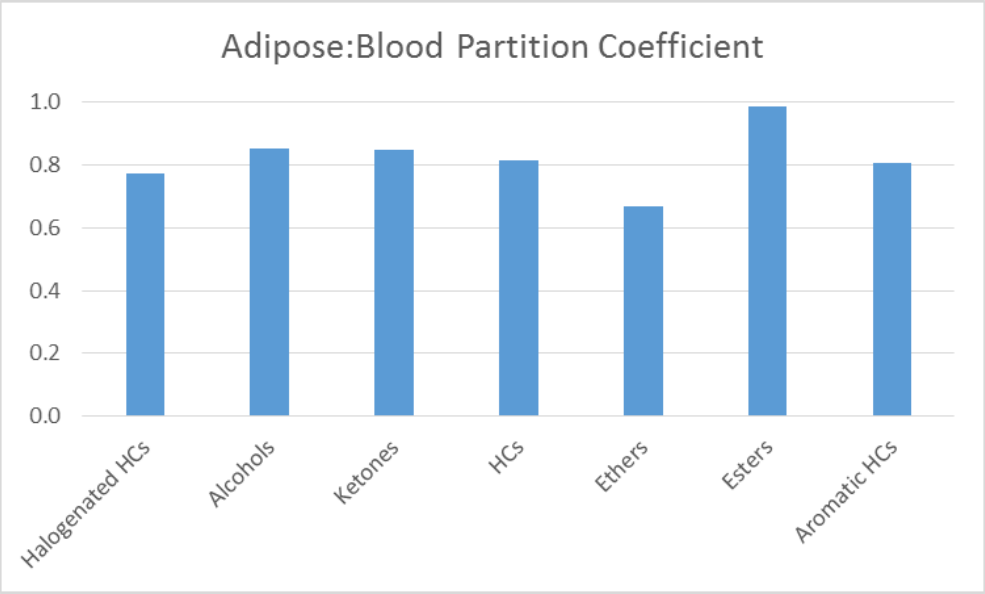


Figure 16. Correlation coefficient for adipose:blood partition coefficient for several chemical groups.

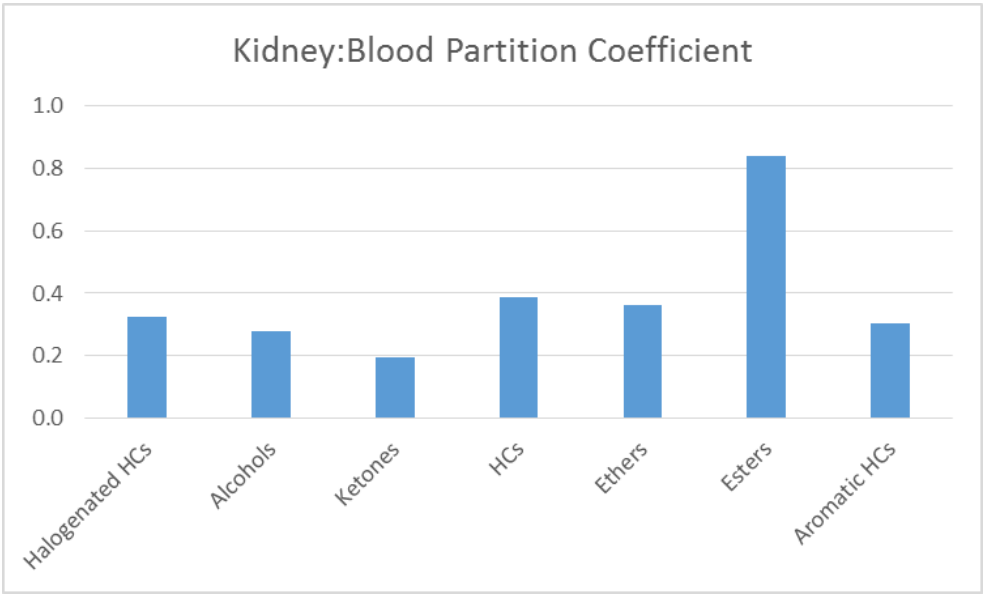


Figure 17. Correlation coefficient for kidney:blood partition coefficient for several chemical groups.

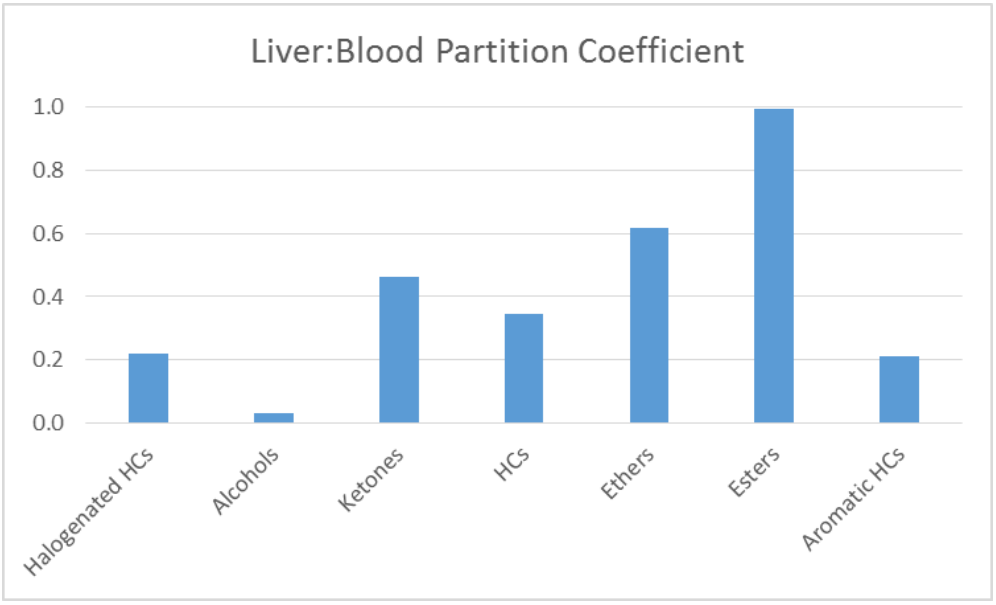


Figure 18. Correlation coefficient for liver:blood partition coefficient for several chemical groups.

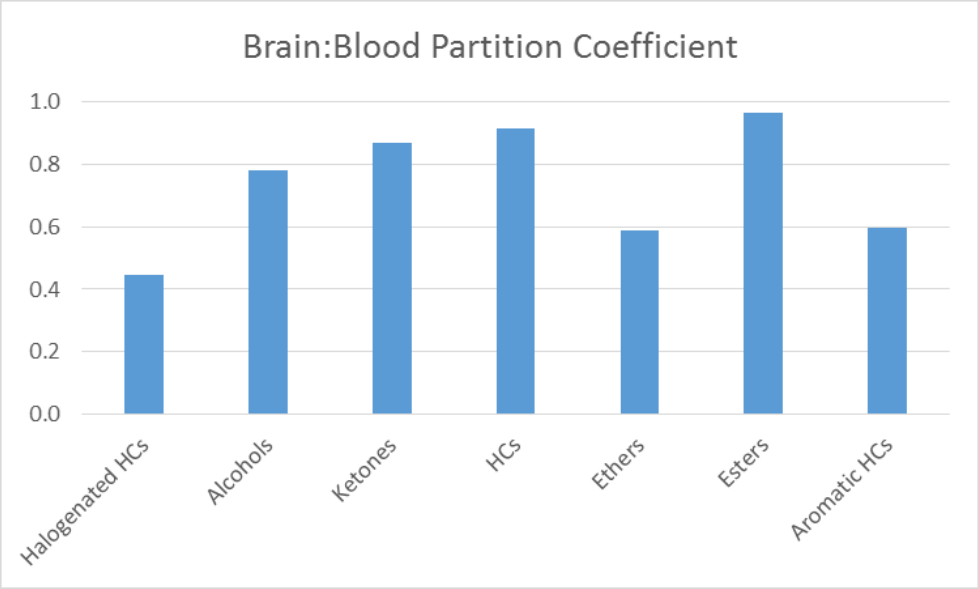


Figure 19. Correlation coefficient for brain:blood partition coefficient for several chemical groups

5 Methodological framework for optimal use of HBM data in assessing population exposure

5.1 Introduction

One particularly well-suited source of information on exposure to environmental agents is human biomonitoring (HBM). Human biomonitoring can be defined as “the method for assessing human exposure or their effect to chemicals by measuring these chemicals, their metabolites or reaction products in human species, such as blood or urine” (CDC, 2009). HBM includes (1) biomarkers that allow assessment of exposure to a chemical on the basis of its measurement in a biological matrix (biomarker of exposure) , (2) changes that have occurred in the biochemical or physiological makeup of an individual because of this exposure (biomarker of effect), or (3) biomarkers that assess a person’s susceptibility to alter the progression along the exposure-effect continuum (biomarker of susceptibility) (NRC, 2006).

Most likely the main achievement of HBM data is that it provides an integrated overview of the pollutant load any participant is exposed to, and hence serves as an excellent approximation of aggregate exposure. The internal dose of a chemical, following aggregate exposure has a much greater value for environmental health impact assessment as the internal body concentration is much more relevant to the impact on human health than mere exposure data (direct EDR-relationship in Figure 20).

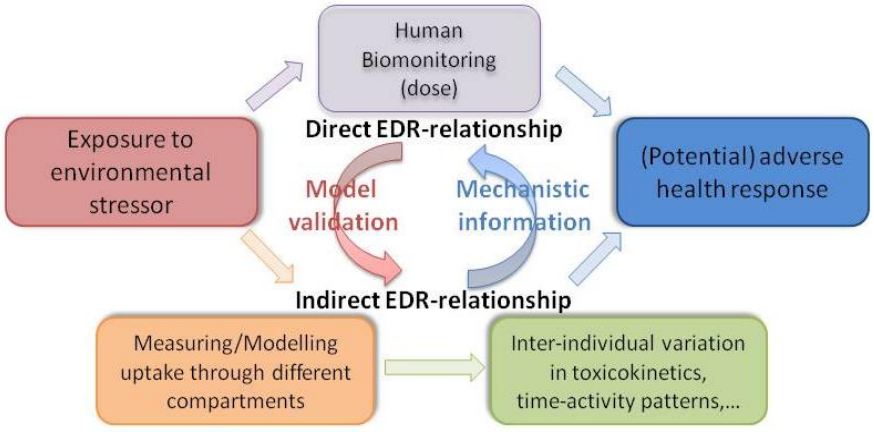


Figure 20. The Exposure-Dose-Response Triad to evaluate the potential adverse health effects of exposure to environmental agents (adapted from Smolders and Schoeters (2007))

However, it needs to be stressed that HBM in itself cannot replace environmental monitoring and modeling data. Most often, environmental monitoring data for different environmental

compartments (air, water, food, soil, settled dust) provide better insight into potential sources, hence allowing the development of more informed and appropriate risk reduction strategies. At the same time, mathematical approaches to describe the pharmacokinetic and toxicokinetic behavior of environmental agents (generally referred to as Physiologically-based Toxicokinetic - PBTK models) offer a more mechanistic insight into the behavior and fate of environmental agents following exposure (Indirect EDR-relationship in Figure 2). As biomarker data also reflect individual accumulation, distribution, metabolism and excretion (ADME) characteristics of chemicals, HBM data offer an excellent opportunity for the validation of these PBTK models. Ultimately, combining both lines of evidence to assess exposure prove to be optimal for relating complex exposure to environmental agents to potential adverse health effects assessment.

There are three approaches (Figure 21) for linking biomonitoring data to health outcomes: direct comparison to toxicity values, forward and reverse dosimetry.

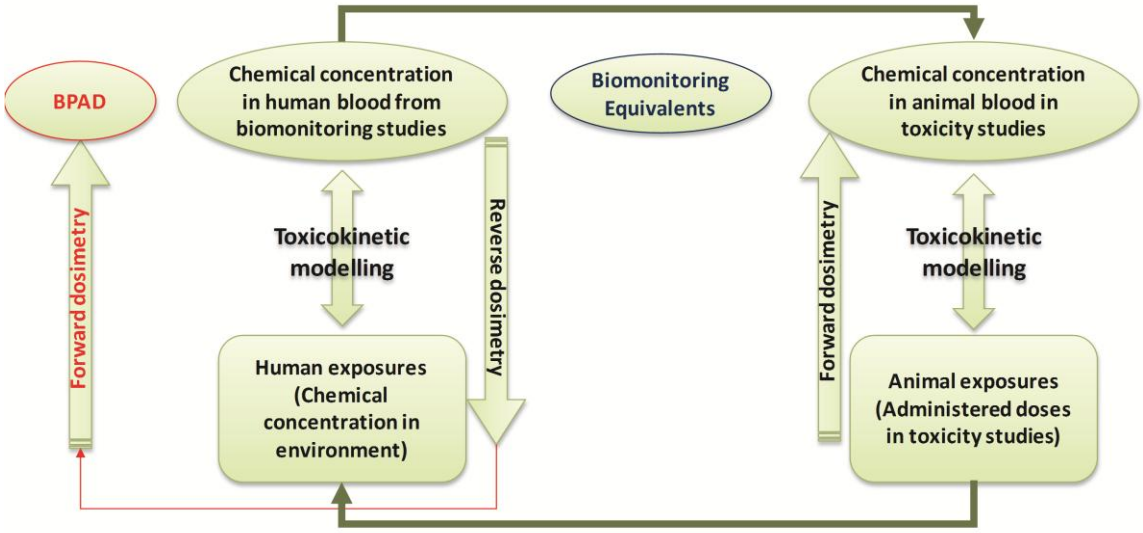


Figure 21. Interpretation of biomonitoring data

Biomonitoring data can be directly compared to toxicity values in the case where the relationship of the biomarker to the health effect of concern has been characterized in the human. In forward dosimetry, pharmacokinetic data in the experimental animal can be used to support a direct comparison of internal exposure in humans derived through the application of PBTK models, providing an estimate of the Margin of Safety (MoS) in humans. It is possible to determine the relationship between biomarker concentration and effects observed in animal studies. An evolution of this concept is the biomonitoring equivalents.

Alternatively, reverse dosimetry can be performed to estimate the external exposure that is consistent with the measured biomonitoring data through the backward application of PBTK

models. In this case the PBTK model is geared with reverse modeling algorithms in order to reconstruct exposure from human biomonitoring (HBM) data. Assimilation of human biomonitoring data and their translation into intake distribution amounts to a computational inversion problem, where the objective is to identify the specific input distributions that best explain the observed outputs while minimizing the residual error. Inputs involve spatial and temporal information on micro-environmental media concentrations of xenobiotics and corresponding information on human activities, food intake patterns or consumer product use that result in intakes; outputs are the observed biomonitored levels. The error metric can be defined in terms of population variation (the latter has to be lower than the intra-individual variation, which may be associated to measurement or other random error source).

More in detail, a computational framework was developed based on Bayesian Markov Chain Monte Carlo (MCMC) combined with the generic Physiological Based Pharmacokinetic (PBTK) model aiming at performing accurate exposure reconstruction (ER). The ER framework developed consists of 3 basic steps:

- At first the prior parameter distribution, the joint probability distribution, the population model and the determination of the measurement model have to be specified.
- At the next step exposure is calculated using MCMC simulation considering the observed biomonitoring data.
- Finally, the evaluation of the results is realized using MC simulation, with emphasis to the comparison of prior and posterior distribution as well as parameter independence

In a more elaborate scheme, the reconstructed exposure, could be used to run the PBTK model in forward mode, so as to estimate the Biologically Effective Dose (BED) at the target tissue. The estimated BED can be evaluated against the respective biological pathway altering dose (BPAD), which is analogous to current risk assessment metrics in that it combines dose−response data with analysis of uncertainty and population variability so as to derive exposure limits (Judson et al., 2010; Judson et al., 2011). The analogy is closest when perturbation of a pathway is a key event in the mode of action (MoA) leading to a specified adverse outcome. BPADs are derived from relatively inexpensive, high-throughput screening (HTS) *in vitro* data, publicly available from the Toxcast 21 database.

5.2 Exposure reconstruction modelling framework

5.2.1 Methods for Exposure Reconstruction related to Population Biomonitoring Studies

Because human biomonitoring typically is an integrative measure of different exposure episodes along various routes and over different time scales, it often is very difficult to

reconstruct the primary exposure routes from HBM data alone. This uncertainty often limits the interpretative value of biomarker data. However, several mathematical approaches have been developed to reconstruct exposures related to population biomonitoring studies, and can be subdivided in a number of different approaches. Exposure reconstruction techniques combined with PBBK model can be subdivided into Bayesian and non-Bayesian approaches (Georgopoulos et al., 2008a). Moreover, the computational inversion techniques (and exposure reconstruction techniques as well), can be classified as deterministic or stochastic (Moles et al., 2003) based on the identification of a global minimum of the error metric, the input parameters and the model setup.

The deterministic methods aim to convergence on a global minimum. The problem is solved using an “objecting function” based on biomarkers. Additionally, constraints such as bounds, equalities and inequalities are incorporated. The deterministic models have been used on several biological applications using different methods. Muzic Jr and Christian (2006) have applied a regression technique in order to estimate biokinetic parameters. Moreover, a gradient method has been used by Isukapalli et al. (2000) calculating the uncertainty in PBBKs. A maximum likelihood method has been carried out for short and long term for exposure reconstruction using a PBBK for chloroform (Roy et al., 1996).

In contrast, the stochastic methods aim to provide a reasonable solution. A probabilistic framework for inverse computation problem is the Bayesian approach which is based on Bayes theorem:

$$p(x \mid \mathrm{y}) = \frac{p(\mathrm{y} \mid \mathrm{x})\, p(\mathrm{x})}{\int p(\mathrm{y} \mid \mathrm{x})\, p(\mathrm{x})\, \mathrm{dx}}$$

Where ***x*** is the possible exposure and ***y*** is the amount of the biomarker **P(X)** which is the available prior information. The relationship between **X** and **Y** and inherently the relationship between the prior and theoretical knowledge is given by

$$p(x \mid \mathrm{y}').$$

Moreover,

$$p_{theory}(\mathrm{y} \mid \mathrm{x}) = p_{model}(\mathrm{y} \mid \mathrm{x})$$

$$p\ (\mathrm{y} \mid \mathrm{x}) = p_{model}(\mathrm{y} \mid \mathrm{x})$$

The posterior distribution of the biomarker measurements will be

$$p_{inferred}(\mathrm{x} \mid \mathrm{y}')$$

and

$$p_{prior}(\mathrm{x} \mid \mathrm{y}') = p_{\mathrm{inferred}}(\mathrm{x} \mid \mathrm{y}')\, p_{prior}(\mathrm{y}')$$

Hence taken into account,

$$p_{prior}(\mathrm{x}) = \int p_{prior}(\mathrm{x} \mid y)\,\mathrm{d}\,y = p_{prior}(\mathrm{x}) \int p_{theory}(y \mid \mathrm{x})\,\mathrm{d}\,y$$

Therefore,

$$p_{posterior}(\mathrm{x} \mid \mathrm{y}') = \frac{p_{theory}(\mathrm{y}' \mid \mathrm{x})\, p_{prior}(\mathrm{x})}{\int p_{theory}(\mathrm{y}' \mid \mathrm{x})\, p_{prior}(\mathrm{x})\,\mathrm{dx}}$$

And

$$p_{theory}(\mathrm{y} \mid \mathrm{x}) = \int p_{error}(\mathrm{y} \mid \mathrm{m})\, p_{\mathrm{model}}(\mathrm{m} \mid \mathrm{x})\,\mathrm{dm}$$

Where $p_{error}(y \mid m)$ is the probability of measuring y when the true value is m.

Therefore,

$$p(\mathrm{x} \mid \mathrm{y}) = \frac{p(x) \int p_{error}(\mathrm{y} \mid \mathrm{m})\, p_{\mathrm{model}}(\mathrm{m} \mid \mathrm{x})\,\mathrm{dm}}{\int p(\mathrm{x})dx \int p_{error}(\mathrm{y} \mid \mathrm{m})\, p_{\mathrm{model}}(\mathrm{m} \mid \mathrm{x})\,\mathrm{dm}}$$

Bayesian MCMC has been used to for the exposure reconstruction of intakes in combination with PBBK (McNally et al., 2014). Holmes et al. (2000) applied genetic algorithms on PBBK models for biokinetic of nicotine to optimize the parameters of the model. Also, fast equivalent operational models (FEOMs) such as the deconvolution technique has been used (Sparacino et al., 2002) for exposure reconstruction in a model combined with PBBK.

5.2.2 Bayesian Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) techniques are numerical approximation algorithms. They originated in statistical physics and they were used in Bayesian inference to sample from probability distributions by constructing Markov chains. In Bayesian inference, the target distribution of each Markov chain is a marginal posterior distribution. Each Markov chain begins with an initial value and the algorithm attempting to maximize the logarithm of the un-normalized joint posterior distribution and eventually arriving at each target distribution by multiple iterations. Each iteration is considered a state. A Markov chain is a random process with a finite state-space where the next state depends only on the current state, not on the past one.

The implemented methodologies are based on Bayesian Markov Chain Monde Carlo (Gelman and Rubin, 1996; Gilks et al., 1996). The method requires defining the prior distributions, the biomonitoring data, as well as a likelihood function defining the likelihood of the data given a set of forward model parameters. The MCMC approach takes into account an acceptance criterion that considers the likelihood of the data given parameters. Also, the MCMC samples using algorithms based on Metropolis Hastings (M-H) or on differential evolution.

Several studies have used MCMC techniques combined with PBBK models for inverse modeling (Chen et al., 2010; Georgopoulos et al., 2009; Lyons et al., 2008; McNally et al., 2012; McNally et al., 2014)

5.2.2.1 Metropolis Hastings (M-H)

The Metropolis Hastings is the sampling algorithm of the MCMC method that has been selected. Given a target density **F** that is associated with a working conditional density **Q(Y|X)**, a Markov kernel **K** is created with stationary distribution **F** and according this kernel a Markov chain **(X(T))** is generated. The limiting distribution of the Markov chain is **F** and integrals can be approximated according to the Ergodic Theorem. M-H is used for deriving and constructing of a kernel **K** that is associated with an arbitrary density **F** (Robert and Casella, 2010). Thus, the proposed distribution typically depends on the current sample and the acceptance of the sample depends on the criteria of M-H. Then, the acceptance of the samples leads the samples to be the next element in the chain, otherwise the previous element is added again in the chain.

The acceptance probability is calculated according the following ratio:

$$a(X \mid X^{(t-1)}) = \min\left\{1, \frac{f(X)\,\mathrm{q}(X^{(t-1)} \mid X)}{f(X^{(t-1)})\,\mathrm{q}(X \mid X^{(t-1)})}\right\}$$

Where $\mathbf{Q(X^{(T-1)}|X)}$ is the Gaussian proposal density and $\mathbf{Q(X|X^{(T-1)})}$ its equal symmetric, $\mathbf{F(X)}$ and $\mathbf{F(X^{(T-1)})}$ are the calculated values for the probabilities for the current and for the candidate point. It has to be mentioned that the Metropolis sampler must have symmetric proposed distributions because the use of Markov Chain draws samples under the condition **of** reversibility (Robert and Casella, 2010).

The process ends when the chain has converged to its stationary distribution or enough samples have been collected in order to perform the desired statistical analysis. The chain is expected to eventually converge to the stationary distribution, which is also the target distribution but typically requires a burn-in period. The burn-in period is the number of iterations that have to be performed before the collected samples. The determination of convergence is based on the diagnostic of Gelman-Rubin technique (Gelman and Rubin,

1992) that examines multiple MCMC chains by dividing each chain up into batches and by examining the variance between the chains.

The sampling techniques and the generation of the proposed samples used on calculation are determined by a particular permutation of the update mode, the adaptive proposal and the delay reduction.

5.2.2.1.1 Update mode

The update mode is based on Multivariate as well as on Component-wise.

A multivariate proposal allows to each iteration the generation of proposed distributions that take into account the correlation from a multivariate normal distribution and from a proposed covariance matrix (Genz and Bretz, 2009; Roberts and Rosenthal, 2009). Hence, multivariate normal sampling proposes the generation of a sample by drawing from a multivariate normal distribution with dimension equal to the number of parameters, mean equal to the previous sample and a covariance matrix determined either by a previously converged chain, or by the computed covariance matrix of the sample chains gathered so far in the run.

Component-wise proposals indicate that a proposal is made for each parameter without considering correlation and it has to evaluate the model a number of times equal to the number of parameters, per iteration. Hence, component-wise update mode samples only one parameter at a time, holding the other fixed (in a Gibbs sampling scheme). The proposed is a univariate normal with a mean equal to the last sample value and a standard deviation computed either by sampling the priors, or by adaptive tuning as the run progresses to achieve the desired acceptance rate.

5.2.2.1.2 Adaptive Proposal Variance

The adaptive MCMC algorithm corresponds to the case where a finite dimensional parameter θ depends on the whole history of the chain $(X_0,\ldots,X_n, \theta_0,\ldots, \theta_n)$ though in practice it is often the case that the pair process $[f(X_n; \theta_n); n > 0]$ is Markovian. The adaptive mode provides the ability of the sample to explore the parameter space and collect samples which are indicative of the target distribution. The acceptance rate is determined by the variance used in the proposal distribution. The amount of the variance controls the size of the steps between points and also it has influence to time of exploration of the parameter space. An effective proposal distribution using a random walk Metropolis algorithm has been done using the an Adaptive Proposal Variance (Haario et al., 2001).

5.2.2.1.3 Delayed Rejection

The Metropolis – Hastings algorithm can be improved by the delaying rejection mechanism (Tierney, 1994) in that the resulting estimates have, uniformly, a smaller asymptotic variance on a sweep by sweep basis. When a Markov chain retains the same position over subsequent time and a candidate sample generated from the rejected proposal sample, the estimates obtained by averaging along the chain trajectory become less efficient. The

solution to that problem is the reduction of the number of rejected proposals based on Mira (2001) methodology. In particular, when a sample is rejected by the Metropolis-Hastings criteria, delaying rejection technique generates a new proposal sample with smaller variance. Thus, delayed rejection is a technique wherein if a sample is rejected when applying the Metropolis-Hastings criteria, another sample is immediately generated by using a proposal with a smaller variance. If this second sample is accepted, it is appended to the chain instead of a repeat of the previous sample. This technique provides the generation of well-mixed chains at the expense of more evaluations of the likelihood on each MCMC iteration. Moreover, it can be used as an alternative technique in case strong correlations exist between the parameters.

5.2.2.2 MCMC algorithms

The goal of MCMC is to design a Markov chain such that the stationary distribution of the chain is exactly the distribution that we are interesting in sampling from. The combination of the sampling technique settings leads to existing Metropolis Hasting techniques. Table 5 presents the available MCMC algorithms based on Metropolis Hasting sampling that can been used.

Table 5. MCMC algorithms based on Metropolis Hasting

MCMC algorithms	Update mode:	Reference
Delayed Rejection Metropolis (DRM)	Multivariate	(Mira, 2001)
Delayed Rejection Adaptive Metropolis (DARM)	Multivariate	(Haario et al., 2006)
Adaptive Metropolis(AM)	Multivariate	(Haario et al., 2001)
Componentwise Metropolis (CHM)	Componentwise	(Haario et al., 2005)
Random-Walk Metropolis (RWM)	Componentwise	(Gilks and Roberts, 1996)

The Delayed Rejection Metropolis (DRM or DR) algorithm is a Random-Walk Metropolis (RWM) (Mira, 2001). Whenever a proposal is rejected, the DRM selects one or more alternate proposals and corrects for the probability of this conditional acceptance. The delaying rejection enforces the decreased autocorrelation in the chains and the algorithm is encouraged to move. The additional calculations increase the computational cost of each iteration of the algorithm in which the first set of proposals is rejected, but the major benefit is the faster convergence to the optimal solution.

The Delayed Rejection Adaptive Metropolis (DRAM) algorithm is merely the combination of both Delayed Rejection Metropolis (DRM) and Adaptive Metropolis (AM) (Haario et al., 2006). DRAM has been demonstrated to be robust in extreme situations where DRM or AM fail

separately. Haario et al. (2006) present an example involving ordinary differential equations in which least squares could not find a stable solution, and DRAM did well.

The Adaptive Metropolis (AM) algorithm of Haario et al. (2001) is an extension of Random-Walk Metropolis (RWM) that adapts based on the observed covariance matrix from the history of the chains. The algorithm is specified under adaptation and periodicity. Thus, the beginning of the iteration and the frequency in the periodicity in adaption have to be set. The adaption has to be controlled and immediate adaption has to be avoided since the algorithm is based on the observed covariance matrix of historical and accepted samples. Hence, a valid covariance matrix before adaptation has to be composed with a large number of samples. However, at the beginning of the algorithm, a small covariance matrix is commonly used to encourage a high acceptance rate.

The Componentwise Metropolis (CHM) is based on the Single Component Adaptive Metropolis (SCAM) that has been developed by Hario et al. (2005) and on the single component Metropolis – Hasting algorithm. In the SCAM the adaption is performed component by component. The chain is no more Markovian, but it remains ergodic. The SCAM can be used in many moderately high dimensional problems. Also, the algorithm does not need detailed prior knowledge of the target distribution and it can be used in numerous problems typically solved using pre-runs and hand tuning (Haario et al., 2005). Also, the algorithm resembles basic single component Metropolis algorithm with Gaussian proposal distributions, the only exception being that the variances of the one-dimensional proposal distributions depend on time and the variance is been computing by a simple recursive formula. Moreover, in high dimension the updating of the proposal distribution performed demands only computations of component-wise variances. Hence, the additional computation brought in by the adaptiveness is negligible. Additionally, component-wise proposals usually indicate that a proposal is made for each parameter, without considering correlation. In case of that parameters are correlated, the problem of the distribution is faced with the rotation of the proposal distribution. Thus, the covariance matrix of the chain is computed and the principal vector direction is determined and it is used as sampling directions in the SCAM-algorithm. After the burn-in period of the algorithm, the proposal direction is fixed and the sampling is continued by only updating the size of the one-dimensional Gaussian proposal distribution. Hence, the SCAM is characterized as fully automatic algorithm. SCAM is widely applicable and general-purpose algorithm. It is appropriate to be performed to models with a small to medium number of parameters since the proposal covariance matrix grows with the number of parameters and the computation cost simultaneous increases.

The random walk algorithm of Metropolis is known to be an effective Markov chain Monte Carlo method for many diverse problems (Metropolis et al., 1953). The proposed Random-Walk Metropolis (RWM) is a multivariate extension of Metropolis-within-Gibbs (MWG) (Gilks and Roberts, 1996). RWM is an algorithm the initials specification are not necessary though

blockwise sampling. In fact RWM is a generic algorithm to draw a sample from a d-dimensional target distribution from a probability density function.The optimal scale of the proposal covariance is based on the asymptotic limit of infinite-dimensional Gaussian target distributions that are independent and identically-distributed (Gelman et al., 1996). In case of multiple parameters the existence of correlations occurrences is very common. Hence, MCMC algorithms attempt to estimate multivariate proposals from a multivariate normal distribution taking into account correlations through the covariance matrix. The convergence of the algorithm is related with the proposal density. A small variance leads to slowly converge and conversely, if the variance is too large, the Metropolis algorithm will reject too high a proportion of its proposed moves (Roberts et al., 1997).

5.2.3 Differential Evolution Monte Carlo

Differential Evolution (DE) is a genetic algorithm for numerical global optimization and it is a population Markov Chain Monte Carlo algorithm, in which parallel run for several chains is applied (Ter Braak, 2006). The combination of DE and MCMC is called Differential Evolution Monte Carlo (DEMC) and the field has been explored among others by Liang and Wong (2001) Liang (2002) and Laskey and Myers (2003). DEMC provides solutions to the choosing and the orientation of the jumping of the distribution that is an important practical problem in random walk Metropolis. In fact DEMC algorithm is based on a Metropolis Hasting and it is combined with a genetic algorithm called Differential Evolution (DE) with multiple chains and each chain learn from another parallel chain. The crucial idea behind DE is an innovated generation of parameter vectors. DE adds a weighted difference vector between two population members in order to generate vectors. The vector yields an objective function value. Then the value is compared with the predetermined population and if the resulting value is lower than the existent, the new vector replaces the compared vector. Moreover, the evaluation of each generation can be done with the best parameter vector in order to retain track of progress during the minimization process. The DE is described in detail by Storn and Price (1995) (1997) and the adaption of DE in MCMC is described and proofed by Ter Braak (2006).

5.2.3.1 DEMC algorithms

The applied DEMC algorithm is based on the Ter Braak (2006) algorithm.

Table 6. MCMC algorithms based on Differential Evolution method

MCMC algorithms	Update mode:	Reference
Differential Evolution Monte Carlo (DEMC)	Multivariate	(Ter Braak, 2006)

DEMC is similar with Metropolis-within-Gibbs (MWG) but the main different consist in that DEMC updates by chain. The algorithm is specified under the number of chain that should be at least three and the thinning factor. The thinning factor provides the reduction of storage requirements and enhances the convergence of the chain to posterior distribution. In particular, the sampling is realized randomly and without replacement from a possibly thinned chain. Moreover, an adaptive step size can be used (ter Braak and Vrugt, 2008) with the same contribution as it has been described to section 1. Also the snooker update fraction (Gilks et al., 1994) (Liang and Wong, 2001) (ter Braak and Vrugt, 2008) can been specified providing to the sampler the ability to update along each coordinate axis in turn one axis at a time, with the specificity that this axis does not need to run parallel to the coordinate axes. Finally, it can be set the randomly uniform offset distribution that added to the creation of the DEMC proposal distribution.

5.2.4 Methodology in HEALS and selected algorithm

The exposure reconstruction approach to be applied in the HEALS methodology and computational platform relies upon the concepts initially described by Georgopoulos et al. (2009).

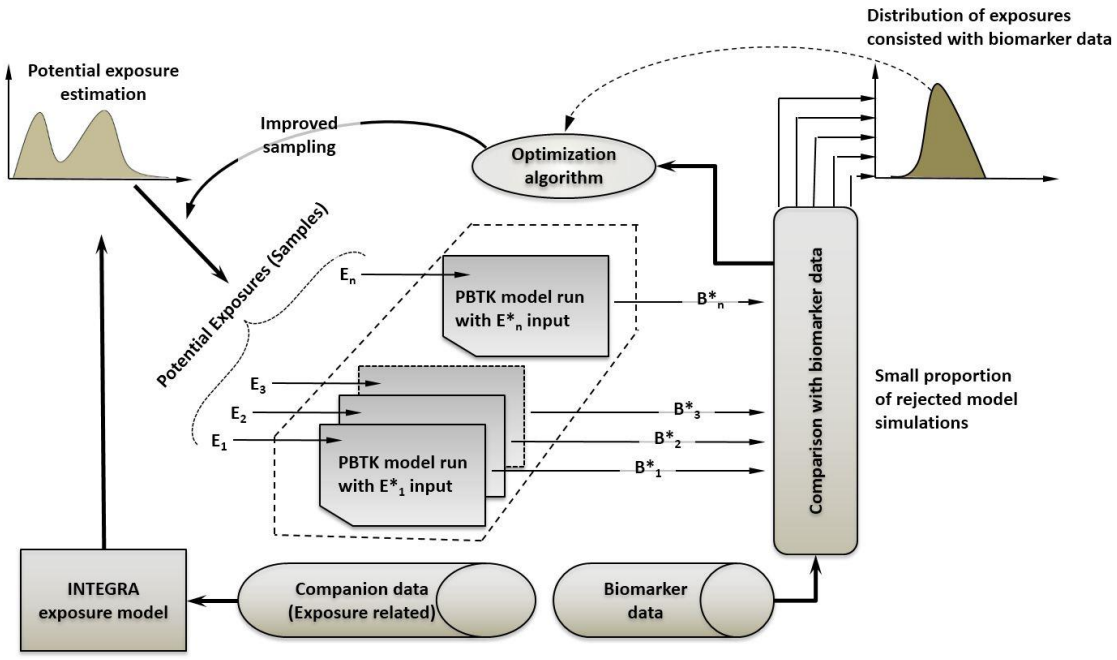


Figure 22. Optimization-aided exposure reconstruction based on HBM data using time-evolving PBTK models (figure adapted from Georgopoulos et al. (2009)).

The analysis of the exposure reconstruction problems based on the MCMC and DEMC technique is realized according to the following steps:

1. The process starts from exposure related data which are fed into the exposure model;

2. This in turn provides input to the PBTK model, taking into account the duration and the magnitude of exposure from all the exposure routes (inhalation, skin and oral route);
3. The result of the PBTK model simulation (taking also into account the distribution of PBTK parameters, e.g. inter-individual variability in clearance), is then evaluated against the human biomonitoring data distributions. Based on the outcome of the comparison, the optimization algorithm changes the exposure model input parameters following each iteration, so as to achieve the convergence to biomonitoring data;
4. More detailed information on exposure parameters reduces uncertainty in back-calculating doses from biomarker information, resulting in faster and more efficient convergence;
5. Several iterations are repeated, until minimizing the error between the predicted and the actual biomonitored data.

The framework shown in Figure 22 is not limited to exposure reconstruction. It can also be used for estimating distributions of physiological and biochemical PBTK model parameters (under well-defined exposure conditions) for individuals and populations that are consistent with available biomarker data (typically study-specific data where exposures are adequately characterized) by combining the data with prior estimates of the parameters.

The Bayesian Markov Chain Monte Carlo technique described above simulates and calculates the investigated exposure conditions. The sampling is set appropriately according to the problem and to the available data for the proposal function. The flowchart diagram of the whole process is shown in Figure 23.

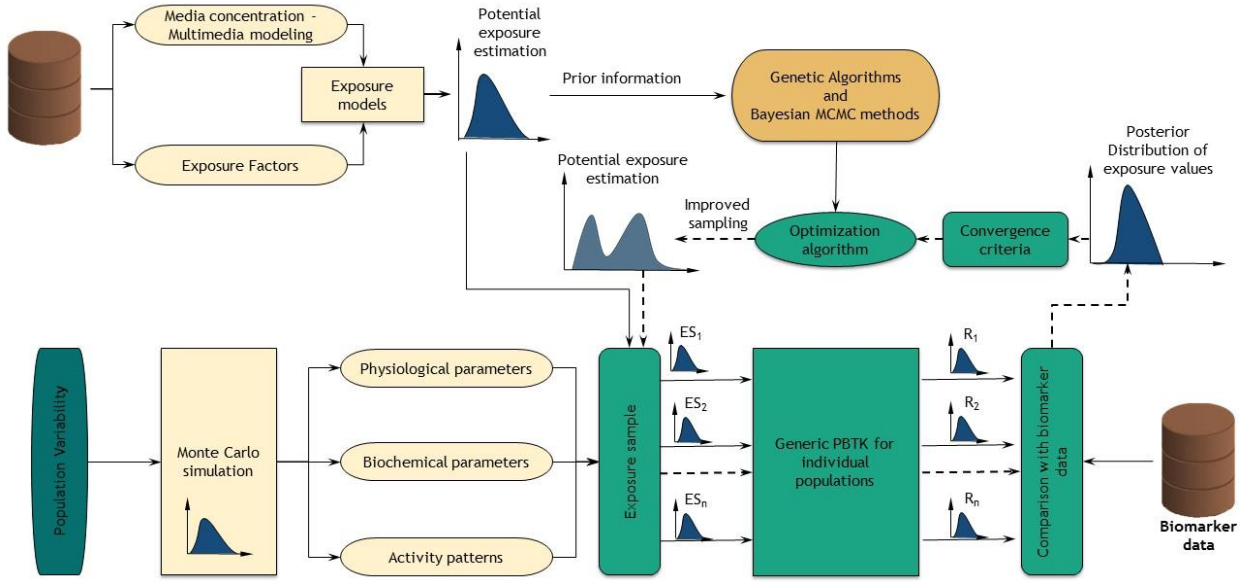


Figure 23. Exposure reconstruction flowchart procedure

The model has been developed in acslX®. The user can choose between component wise or multivariate update mode. The adaptive mode as well as the delay rejection can be set in the M functions.

5.3 Application

5.3.1 Average daily intake exposure reconstruction starting from spot urine samples

The described exposure reconstruction methodology has been applied for BPA and DEHP under 2 realistic exposure scenarios for both chemical substances. The scenarios were investigated using both the MCMC algorithm with RWM and the DEMC algorithm. The algorithms have been performed to be adaptive and to have a burn-in period of 50 iterations. The total number of iteration was set to 1000.

5.3.1.1 Exposure to BPA – single route exposure

Bisphenol A (BPA) (4,4'-(propane-2,2-diyl)diphenol) is one of chemicals with the highest industrial production volume worldwide (Bailin et al., 2008). The major volume of BPA is used for the production of polycarbonate plastics as well as a basic component in production of the epoxy resin (Vandenberg et al., 2009). Various common consumer products contain or are made using polycarbonate plastics such as household electronics and baby bottles (Liao and

Kannan, 2011). Epoxy resin is used in the majority of food and beverage cans (Erickson, 2008). Moreover, BPA is commonly used in paper industry and particularly as color developer in thermal and copy paper (Biedermann et al., 2010; Liao and Kannan, 2011; Mendum et al., 2011; Viñas et al., 2012).

The first scenario referring to BPA is common and usual for adult generic population that is exposed 3 times per day to BPA. The exposure conditions include three oral doses (dietary exposure) during breakfast at 7:00 AM (dose 1), lunch at 2:00 PM (dose 2) and dinner at 7:00 PM (dose 3). The exposure concentration doses have been set to 14, 28 and 14 ug respectively. In addition, the exposure dose boundaries for the prior distributions are ranging between 10 and 40 ug.

The generic PBBK model developed in the frame of HEALS was parameterized based on literature data (Edginton and Ritter, 2009). Moreover, the biomonitoring data have been set up using measurement from morning urine samples. The biomonitoring data referred to BPA-glucuronide, a common type of metabolite quickly formed by liver metabolism, which is the only metabolite that has been detected in urine and blood after controlled exposure (Matthews et al., 2001; Völkel et al., 2002). The prior knowledge of the distribution of the exposure time is based on the actual dietary schedule of the generic population. The time of the three basic dietary meals follows normal distribution and in particular, the breakfast is at 7:00 AM and the dinner is at 7:00 PM with standard deviation of 30 minutes for both and the lunch is at 2:00 PM with standard deviation of 30 minutes. The total time of the iteration was 885 seconds.

The results of the exposure reconstruction (Figure 24 (a), (b) and (c)) show that the posterior distributions include the actual exposure doses. Moreover, the posterior distributions have a reduced standard deviation and a mean value closer to the real. However, the MCMC model using the prior knowledge cannot achieve a posterior distribution with a sufficient confidence interval (CI) for actual exposure value. The dose 2 and 3 appear a better prediction and a high frequency of the predictive closer to the actual exposure value.

The use of the DEMC algorithm demonstrated that the predictions have a similar behavior compared to the MCMC (Figure 25 (a), (b) and (c)).

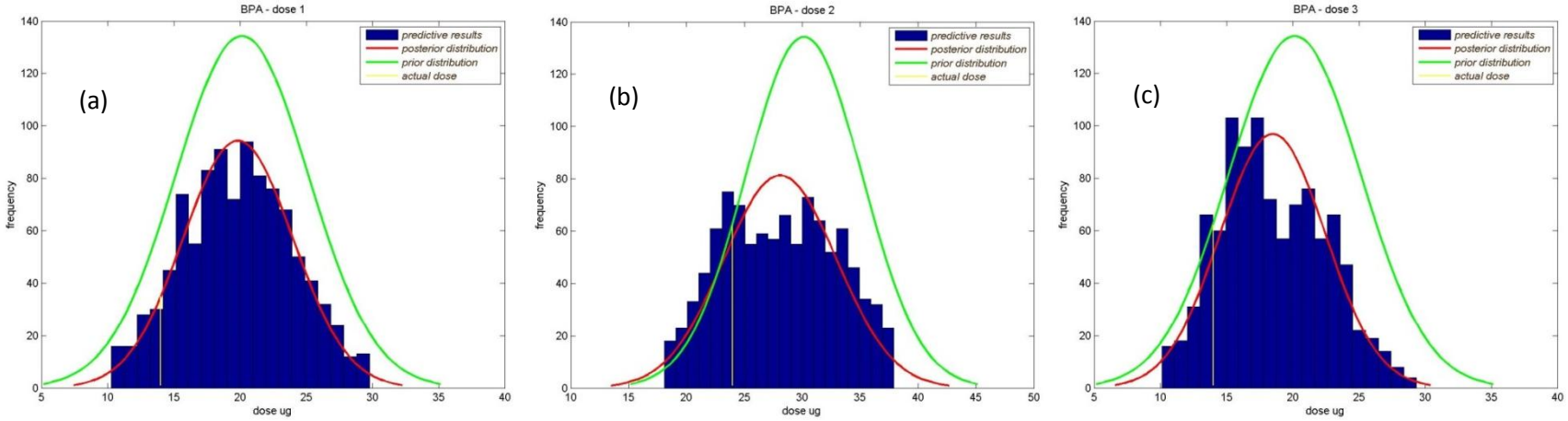


Figure 24. Reconstruction of oral exposure to BPA from urine concentration to metabolite glucuronide for dose 1 (a), dose 2 (b) and dose 3 (c) through MCMC

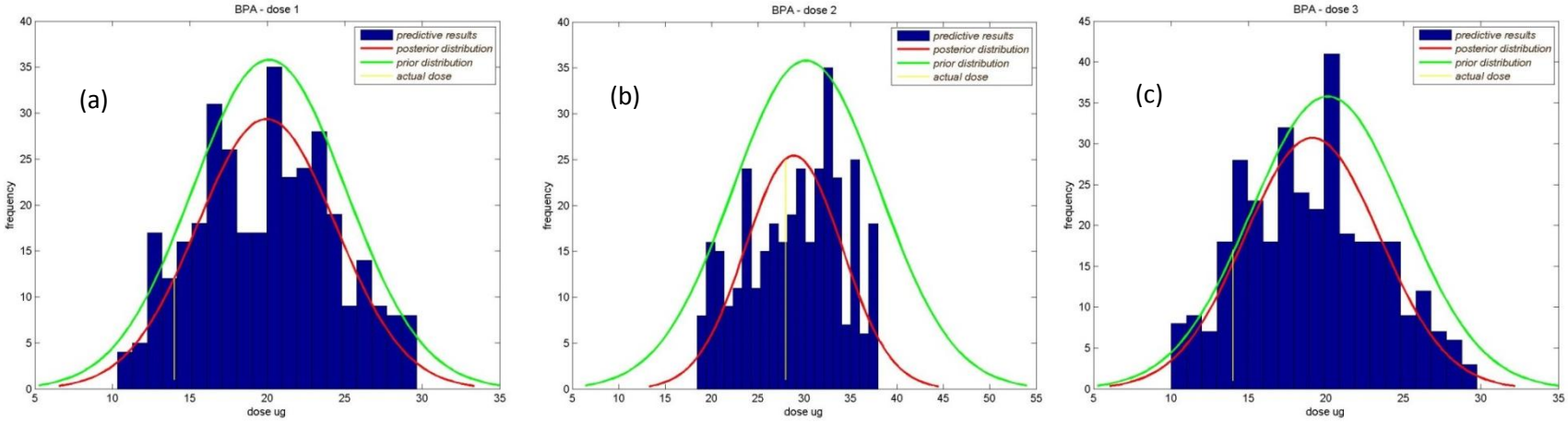


Figure 25. Reconstruction of oral exposure to BPA from urine concentration to metabolite glucuronide for dose 1 (a), dose 2 (b) and dose 3 (c) through DEMC

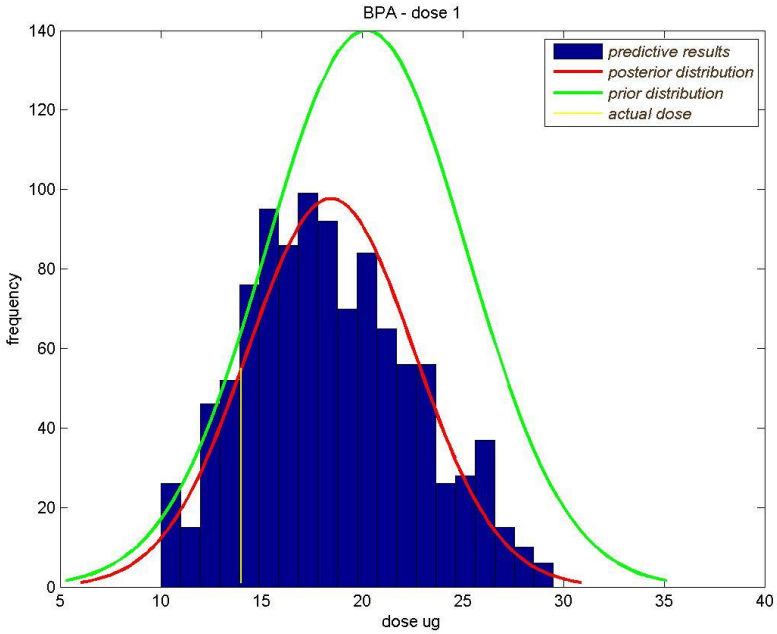


Figure 26. Reconstruction of oral exposure to BPA from urine concentration to metabolite glucuronide – dose 1 – DEMC (using one additional urine sample taken at 17:00)

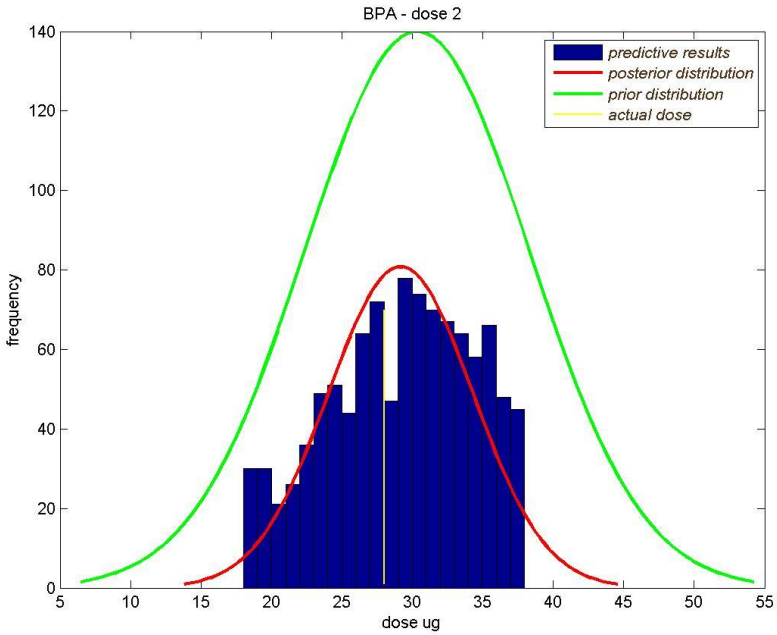


Figure 27. Reconstruction of oral exposure to BPA from urine concentration to metabolite glucuronide – dose 2 – DEMC (using one additional urine sample taken at 17:00)

The same scenario has been tested using biomonitoring data of two different spot sample measurements through the DEMC algorithm. The first one taken at 10:00 AM that is 3 hours after the first dietary exposure while the second one has been kept the same at 7:00 AM.
A third simulation was realized for exposure reconstruction of BPA using the DEMC algorithm. The biomonitoring data used in the simulation were increased by adding one sampling point between the 2nd and the 3rd exposure at 17:00. The simulation results of that the reconstructed distributions of the 1st and 2nd exposure were improved. This is due to the fact that the new distances of the prior information of the time exposure from the available biomonitoring data were shorter than the first simulation, as well as the new data point gave a better determination of the exposure scenario. Hence, the new information between the second and third exposure gave the ability to the Metropolis Hasting algorithm to create a better set of covariance matrix, which in turn resulted in improved predictions according to the prior information. The comparison of Figure 25 (a) with Figure 26 as well as of Figure 25 (b) with Figure 27 confirmed that addition of biomonitoring data enhanced the prediction capability.
The predictive result distribution of the histograms shows that density of the results is near to the actual exposure. Moreover, the mean value of the distribution of the second dose in Figure 27 is almost identical with the mean value of the posterior distribution.

5.3.1.2 Exposure to DEHP – exposure from two exposure routes

DEHP (Bis(2-ethylhexyl) phthalate) is plastic-softening phthalate of widespread use that is used to enhance the flexibility of rigid polyvinylchloride (PVC) (Lorz et al., 2007). The content of DEHP in flexible polymer materials varies but is often around 30% (w/w) (Program, 1982). The DEHP can migrate, leach, or evaporate into indoor air and atmosphere from building materials, daily and common used products such clothes, accessories, toys and also inside automobiles from plasticized components leading to exposure of phthalate esters via ingestion, inhalation and dermal pathways (Becker et al., 2004; Fromme et al., 2007; Wormuth et al., 2006). Moreover, among 25 different phthalate esters DEHP is the most common used in the production of medical devices (Tickner et al., 2001) such blood bags and dialysis equipment. However, human exposure to DEHP is age and lifestyle dependent (Franco et al., 2007). DEHP is mainly used in PVC products profiles and hoses as well as in film, wall- and roof covering and flooring. Hence, the wide use of DEHP gives rise to many possible scenarios of human exposure.
The scenario investigated refers to:

- non-dietary oral exposure from dust (hand to mouth behavior), accounting for DEHP dust concentration with prior distribution N (400, 70) ug/g dust and range between 200 and 800.
- exposure via inhalation, with a prior distribution of N (3.5, 0.7) ug/kg/day, ranging between 1 and 6.

The generic PBBK model developed in the frame of HEALS was parameterized and validated based on the toxicokinetic data from Cahill et al. (2003) and Lorber et al. (2010). Moreover, the biomonitoring data have been taken from morning measurements of urine samples. The biomonitoring data referred to the metabolite MEHP.

The MCMC analysis shows that the posterior distributions include the actual exposure doses via inhalation and oral route. Moreover, the posterior distribution has a reduced standard deviation and a mean value close to the actual exposure value even though in case of inhalation exposure the movement of the mean value toward the actual one is not at a sufficient level. The results obtained applying the DEMC algorithm show how the actual exposure level is very close to the posterior distribution. In addition, the density of the predictive results from the DEMC shows a peak close to the actual exposure. The results are presented in Figure 30 and Figure 31.

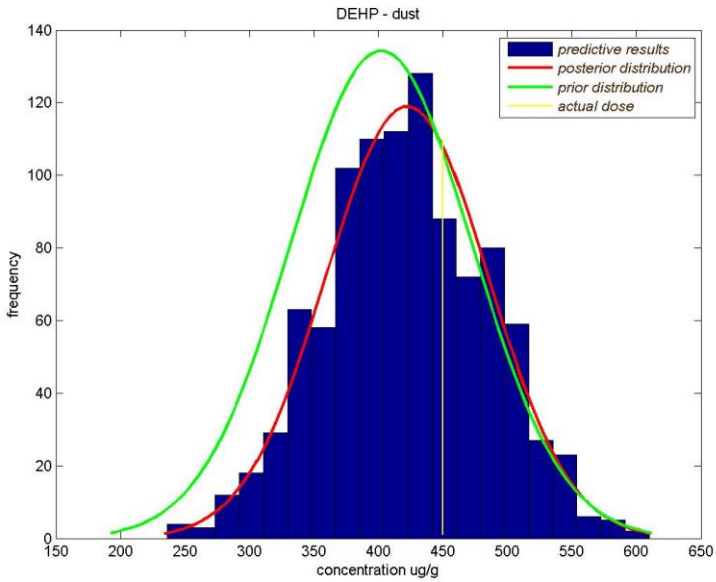


Figure 28. Reconstruction of dust exposure to DEHP from urine concentration to metabolite MEHP – dust - MCMC

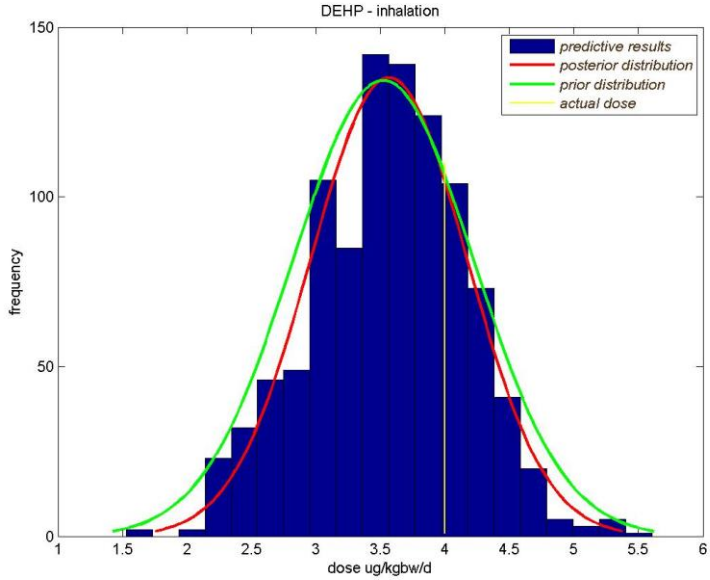


Figure 29. Reconstruction of inhalation exposure to DEHP from urine concentration to metabolite MEHP – inhalation – MCMC

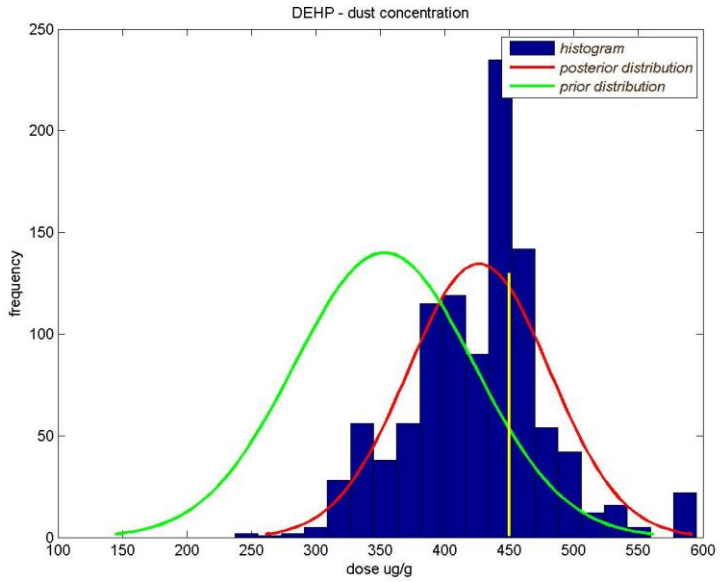


Figure 30. Reconstruction of inhalation exposure to DEHP from urine concentration to metabolite MEHP – inhalation - DEMC

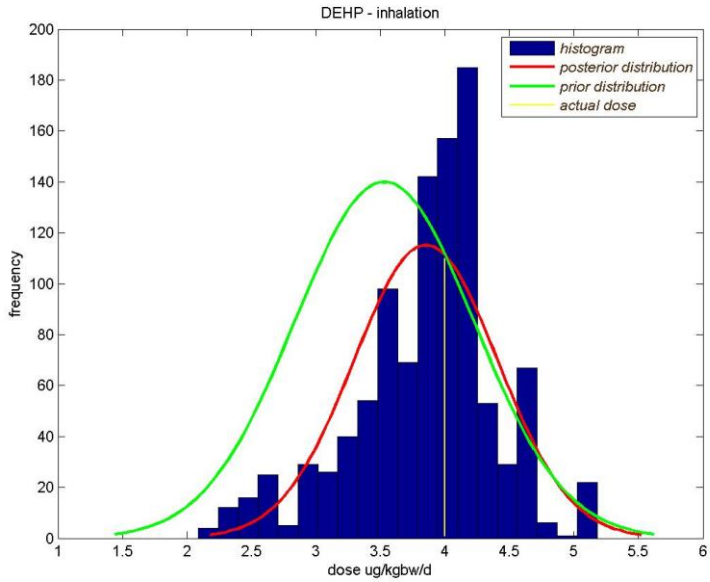


Figure 31. Reconstruction of inhalation exposure to DEHP from urine concentration to metabolite MEHP – inhalation – DEMC

5.3.1.3 Exposure reconstruction of trichloromethane (TCM)

Aiming at investigating the effect of domestic cleaning activities on children passive exposures to trichloromethane from their mere physical presence at home, we evaluated urinary TCM data from children and matched-mothers. In practice, using the children urinary chloroform levels, indoor air background chloroform concentrations were reconstructed. These concentrations were used for estimating mother exposure. Re-running forward the model using these concentrations levels as exposure for the mothers, their urinary chloroform was predicted, as well as the respective chloroform blood levels (internal exposure). Regarding our study on the effect of domestic cleaning activities, our analysis showed the valid use of urinary chloroform levels as a proxy to internal exposure to chloroform, but only if background exposure concentrations were considered. Given that chloroform are metabolized and excreted rather rapidly, their levels in morning urine reflect primarily indoor air concentration and, to a smaller extent, drinking water levels. Activities that lead to significant increase in chloroform release into the indoor air such as dishwashing, bleaching, showering, bathing etc. affect the observed biomarker levels, by raising the uptake rate of chloroform from the indoor air. Based on the urinary levels and by reconstructing exposure so as to fit the measured biomonitoring data, blood and exhaled breath chloroform levels were also calculated for the matched-mothers' and children (Figure 32).

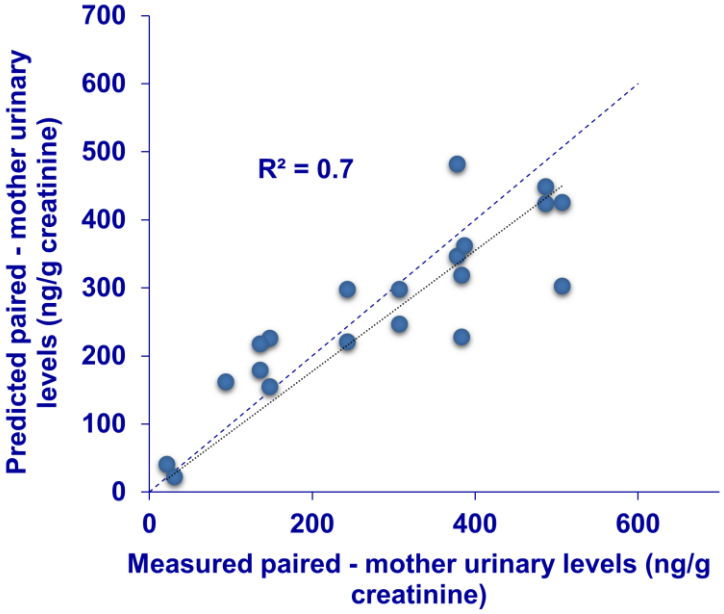


Figure 32. Measured vs. predicted urinary chloroform levels for the paired mothers, based on indoor concentrations derived from exposure reconstruction of paired-children data

5.3.2 Reconstruction of timely variable exposure from multiple biomonitoring samples

Translating urinary concentration into exposure levels, allowed us to estimate internal exposure as well. Cleaning activities resulted in chloroform blood concentrations close to 100 ng/L, while mopping seems to be associated to higher internal exposure levels; this is the result of the higher intensity of activity during mopping, when intake rate is increased due to increased inhalation rate.

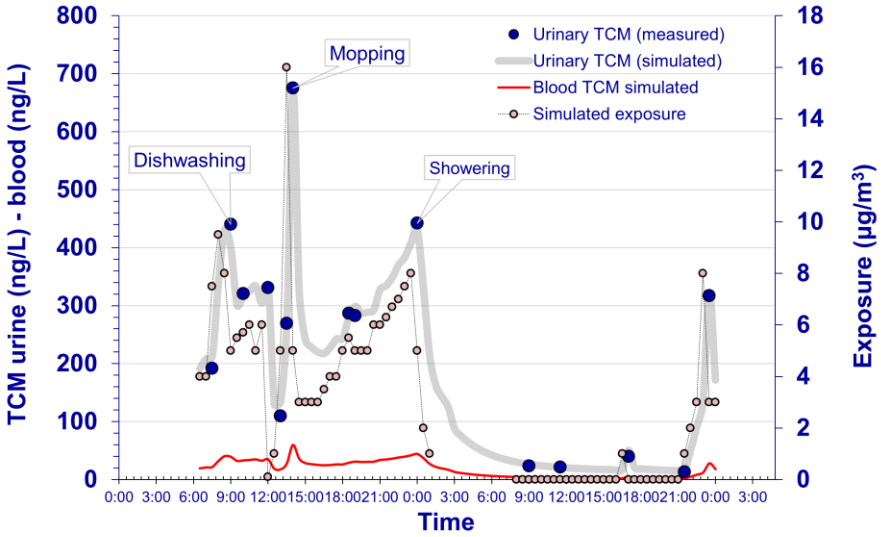


Figure 33. Measured urinary TCM (blue dots) and modelled chloroform levels in urine (grey line), blood (red line) and ambient air (dashed line)

However, potential differences in consumer product-related exposure (amount of product use, chlorine concentration of the product) and housing conditions (air exchange rate) act as confounders prohibiting the derivation of robust conclusions about the relative significance of the respective activities. In any case, we need to highlight that the use of a validated PBBK model allows us to use a biomarker acquired by a non-invasive technique (urinary chloroform), which is also one magnitude of order higher than the respective blood biomarker. This allows us to better differentiate exposure conditions and thus identify the contribution of cleaning activities in the overall exposure to chloroform.

5.3.2.1 Exposure reconstruction of triclosan

Another application of validating the approach was the estimation of triclosan exposure levels during teeth brushing. Seven volunteers were writting in a time-activity diary the time of teeth brushing and the amount of toothpast used, while all day urinary voids were collected and analyzed. Based on the urinary triclosan concentrations, and knowing the timing that exposure events occurred, the amount of triclosan intaken per brushing was successfully estimated. The results of the simulation for a typical individual are illustrated in Figure 34. Starting from the measured urinary triclosan (black dots) and knowing the moment that the individual was exposed to triclosan, the dose received in each brushing was estimated (green dots). The accurate prediction of the dose is shown by the very good fit of the measured urinary concentrations against the ones predicted by the model. This further allows us to estimate the actual internal dose, meaning the concentration of triclosan in blood (red line) and eventually to potential target tissues.

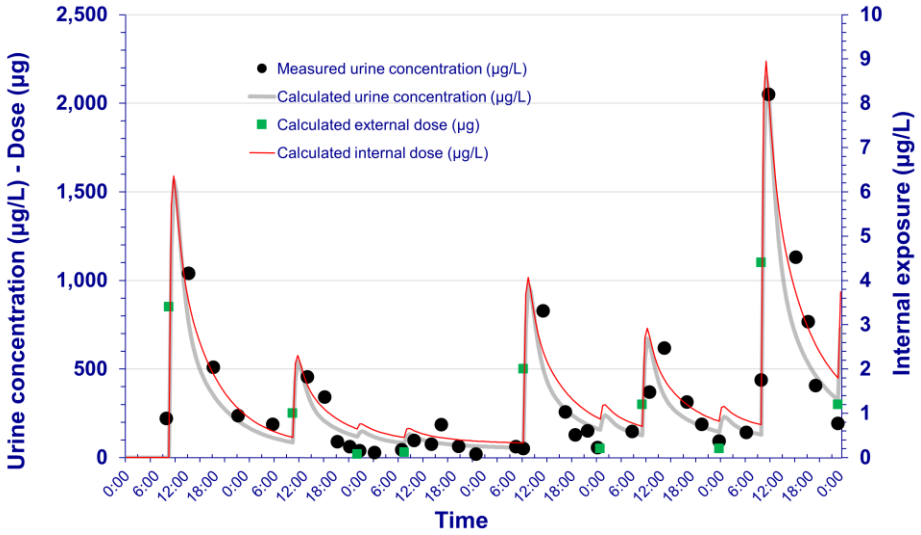


Figure 34. Measured (black dots) and modelled (grey line) urinary triclosan levels, modelled levels in blood (red line) and predicted dose (green dots)

5.3.2.2 Exposure reconstruction of bisphenol A

Similarly, diurnal exposure to bisphenol A through food and drink items was estimated starting from urinary biomonitoring data. The results indicated that overall daily exposure to

bisphenol A remains below 0.1 μg/kg_bw/d, while internal dose of free plasma bisphenol A was in the range of few pg/L (Figure 35).

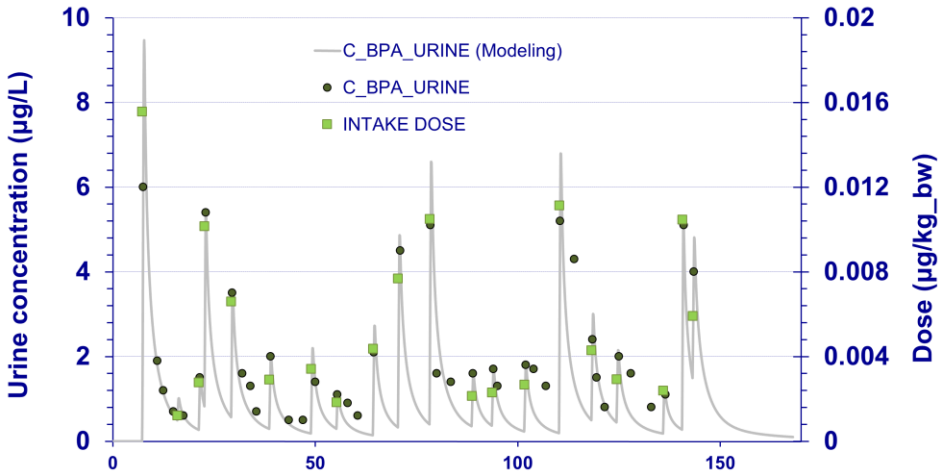


Figure 35. Measured (black dots) and modelled (grey line) urinary bisphenol A levels, and predicted dose (green dots)

6 Use of health-based guidance values as screening tool for HBM data

6.1 The German HBM-I and -II values

Biomonitoring data can be interpreted by comparing the measured biomarker levels to health relevant biomonitoring reference values. In this context, the German Human Biomonitoring Commission has derived health-based guidance values for several compounds (Schulz et al., 2007). These are HBM values determined either based on exposure-effect relationships (for cadmium, lead, mercury and pentachlorophenol) or derived from tolerable daily intake values (as for DEHP). There are two levels of HBM values derived by the German Human Biomonitoring Commission, namely HBM-I and HBM-II:

- The HBM-I value represents the concentration of a substance in human biological material below which – according to the knowledge and judgment of the German HBM Commission – there is no risk for adverse health effects and, consequently, no need for action;
- The HBM-II value represents the concentration of a substance in a human biological material above which there is an increased risk for adverse health effects and, consequently, an acute need for exposure reduction measures and the provision of biomedical advice. The HBM-II-value should thus be regarded as an intervention or action level.

At a concentration level higher than the HBM-I- and lower than the HBM-II-value the result should be verified by further measurements. If these measurements confirm the initial result a search for potential sources of exposure should be undertaken. Exposure to such sources should be minimized or eliminated where necessary and achievable with an acceptable level of input. The HBM-I-value should thus be regarded as a verification or control value.

Up to 2014 the derivation of toxicologically founded HBM values was based on studies which allowed a correlation between the concentration of a substance or its metabolites in human body fluids und the occurrence of adverse effects. Yet, as such studies are lacking for most chemicals, the German HBM Commission decided to derive also HBM values on the basis of toxicologically justified tolerable daily intakes or other suitable parameters from animal studies. Being well aware of the uncertainties of such derivation and estimates, the HBM Commission considers this new approach a possibility to derive urgently needed HBM values for substances or their metabolites for which no appropriate studies on health effects of low dose environmental exposure are currently available. A comprehensive overview of these values is given in Table 7.

Table 7. German Human Biomonitoring Commission reference biomonitoring values for various compounds (Schulz et al., 2011). The HBM-I value is more of a control value while the HBM-II value is defined as an action level

Parameter and medium	Population group (age range)	HBM I value	HBM II value
Based on epidemiological studies			
Cadmium in urine	Children and adolescents	0.5 µg/L	2 µg/L
	Adults	1 µg/L	4 µg/L
Lead in blood	General population incl. children <12 years, women of reproductive age	Suspended	Suspended
Mercury in urine	Children and adults	7 µg/l 5 µg/ g Cr	25 µg/l 20 µg/g Cr
Mercury in blood	Children and adults* Derived from women of reproductive age. The value is recommended for other groups	5 µg/L	15 µg/L
Thallium in urine	General population	5 µg/l	/
Pentachlorophenol in serum	General population	40 µg/L	70 µg/L
Pentachlorophenol in urine		25 µg/L 20 µg/g Cr	40 µg/L 30 µg/g Cr
Sum of PCBs (138+153+180) in serum	Infants, small children and women of child-bearing age	3.5 µg/l	7 µg/l
Based on TDIs			
Sum of the metabolites of di(2-ethylhexyl)phthalate DEHP: 5-oxo- and 5-OH-MEHP in urine	Children aged 6–13	500 µg/L	/
	Women of reproductive age	300 µg/L	/
	Males ~14 years, general population	750 µg/L	/
Bisphenol A in urine	Children Adults	1.5 mg/l 2.5 mg/l	/
Glykolether which are metabolized to methgoxy acetic acid (MAA)	General population	0.4 mg MAA/G Cr	1.6 mg MMA/g Cr
∑ DINCH-metabolites (OH-MINCH + cx-MINCH)	Children Adults	3 mg/l 4.5 mg/l	/
∑ DPHP-metabolites OH-MPHP + oxo-MPHP	Children Adults	1 mg/l 1.5 mg/l	/

6.2 Biomonitoring equivalents

Biomonitoring Equivalents (BEs) are defined as the concentration of a chemical or metabolite in a biological matrix (blood, urine, human milk, etc.) consistent with defined exposure guidance values or toxicity criteria, including reference doses and reference concentrations (RfD and RfCs), minimal risk levels (MRLs) and tolerable daily intakes (TDI), using the knowledge about the toxicokinetic properties of the chemical (Boogaard et al., 2011). The application of BEs is based on the assumption that intake and excretion are in equilibrium, in

order to ensure coherence between the targeted chronic exposure reference value and the respective estimated BE. However, real life exposure is rarely constant or periodically repeated, and this requires some additional need for caution in the interpretation of the biomonitoring data. Requirements for using the above methods involve ensuring the specificity and sensitivity of the biomarker, quantitative risk evaluation, estimation of the proper uncertainty factors for translating the external dose that corresponds to the Point of Departure (PoD) in an animal to a human biomarker concentration (Angerer et al., 2011), as well as some knowledge of the toxicokinetic behavior of the respective biomarker. Use of reliable PBTK models is the most convenient way to translate external exposure reference values into BEs. The biomonitoring equivalents derived for several compounds are presented in Table 8.

Table 8. Biomonitoring Equivalent values for several compounds

Environmental Chemical	Matrix	Analyte	BE value	Reference value	Reference
DDT/DDE/DDD	Blood	(DDT only) (RDDT/DDE/DDD)	30,000 ng/g lipid 40,000 ng/g lipid	FAO/WHO (10 µg/kg/day)	(Kirman et al., 2011)
Hexachlorobenzene	Blood	Hexachlorobenzene	16 ng/g lipid	Health Canada (0.05 µg/kg/day)	(Aylward et al., 2010)
Dioxin TEQ	Blood	Dioxin TEQ	15 ng/g lipid	ATSDR LOAEL (0.12 ng/kg/day)	(Aylward et al., 2008)
Hexabromocyclododecane	Blood, breast milk	Hexabromocyclododecane	190,000 ng/g lipid	EU Draft (POD) (2 mg/kg/day)	(Aylward and Hays, 2011)
Deltamethrin	Blood Urine	Deltamethrin Dimethylcyclopropane carboxylic acid	20 µg/L[1] and 2µg/L[2] 50 µg/L[1] and 7µg/L[2]	EC (10 µg/kg/day)	(Aylward et al., 2011)
PBDE 99	Blood	PBDE 99	520 ng/g lipid	US EPA (0.1 µg/kg/day)	(Krishnan et al., 2011)
Cyfluthrin	Urine	4-fluoro-3-phenoxybenzoic acid	400 µg/L	FAO/WHO ADI (10 µg/kg/day)	(Hays et al., 2009)
Triclosan	Urine	total triclosan (free plus conjugates)	2600 µg/L	EC (120 µg/kg/day)	(Krishnan et al., 2010b)
Bisphenol A	Urine	BPA-glu	2000 µg/L	EFSA (50 µg/kg/day)	(Krishnan et al., 2010a)
Di-2(ethylhexyl) phthalate - DEHP	Urine	MEHP, MEHHP, and MEOHP MEHP, MEHHP, MEOHP, and 5cx-MEPP MEHP, MEHHP, MEOHP, 5cx-MEPP, and 2cx-MMHP	660 µg/L 1000 µg/L 1100 µg/L	EFSA (50 µg/kg/day)	(Aylward et al., 2009b)
Diisononyl phthalate - DiNP	Urine	Oxidative (OH-, oxo-, and carboxy-MINP) metabolites MiNP	15 µg/L[3] 10.7 µg/L[4] 12.7 µg/L[5] 10.6 µg/L[6] 0.7 µg/L[3] 0.5 µg/L[4] 0.6 µg/L[5] 0.5 µg/L[6]	EFSA (150 µg/kg/day)	(Hays et al., 2011)
di-n-butyl phthalate - DBP	Urine	MBP	0.2 µg/L	EFSA (10 µg/kg/day)	(Aylward et al., 2009a)

Environmental Chemical	Matrix	Analyte	BE value	Reference value	Reference
benzylbutyl phthalate - BzBP	Urine	MBzP	12 µg/L	EFSA (500 µg/kg/day)	(Aylward et al., 2009a)
diethyl phthalate - DEP	Urine	MEP	18 µg/L	EPA (800 µg/kg/day)	(Aylward et al., 2009a)
Benzene (for chronic non-cancer exposure)	Blood Urine Blood Urine Blood Urine Blood	benzene Unmetabolized benzene benzene Unmetabolized benzene benzene Unmetabolized benzene benzene	0.15 µg/L 0.16 µg/L 1.29 µg/L 1.42 µg/L 0.29 µg/L 0.33 µg/L 0.04 µg/L	USEPA Chronic RfC TCEQ ReV CA REL ATSDR chronic inh. MRL	
Benzene cancer risk-specific exposure levels	Urine Blood Urine Blood Urine	Unmetabolized benzene benzene Unmetabolized benzene benzene Unmetabolized benzene	0.05 µg/L 0.058–0.204 µg/L 0.125–0.286 µg/L 0.058–0.204 ng/L Not calculated 0.15 µg/L 0.16 µg/L	USEPA, risk-specific concentrations (1E-04 risk - 13.0–45.0 µg/m^3) USEPA, risk-specific concentrations (1E-06 risk - 0.13–0.45 µg/m^3)	(Hays et al., 2012)
	Blood Urine Blood Urine	benzene Unmetabolized benzene benzene Unmetabolized benzene	0.204 µg/L 0.286 µg/L 0.204 ng/L Not calculated	TCEQ, ESL cancer (1E-04 risk - 44.6 µg/m^3) TCEQ, ESL cancer (1E-04 risk – 0.446 µg/m^3)	
Toluene	Blood	Toluene	50 µg/L 40 µg/L 3 µg/L 3 µg/L 30 µg/L	USEPA chronic RfC (128 mg/m^3) Health Canada chronic inhalation TDI (150 mg/m^3) WHO air quality guideline (332 mg/m^3) ATSDR chronic inhalation MRL (132 mg/m^3) ATSDR acute MRL (150 mg/m^3)	
Cadmium	Urine	Cadmium	1.2 µg/L	FAO/WHO (10 µg/kg/day)	(Hays et al., 2008)
Arsenic, inorganic	Urine	Inorganic arsenic, monomethylated arsenic, and dimethylated arsenic	6.4 µg/L	ATSDR (0.3 µg/kg/day)	(Hays et al., 2010)

[1] adults
[2] children
[3] children 6-11 years
[4] adolescents 11-16 years
[5] men >16 years
[6] women >16 years

7 Conclusions

Biomonitoring, which is the measurement of chemicals in human tissues or fluids such as blood or urine, is being widely used. As a result, concern is growing about the implications of the presence of these chemicals in the body. A major challenge is that methodologies and data typically used for linking biomonitoring data to specific health outcomes do not exist for most of the chemicals being measured. Without sensitive, reliable, and accurate methods for interpreting biomonitoring data, worst case assumptions about health risk will continue to be applied, especially by the public.

Additionally, with the emergence of improved analytical techniques that are capable of detecting an increasing number of chemicals at ever lower concentrations in the human body, the number of chemicals that are being identified in humans and the sheer volume of biomonitoring data are growing. The need for methods to interpret and use these data in a human health risk-based context is becoming critical to responding to public and regulatory concerns about chemical safety.

Several approaches are emerging to effectively interpret biomonitoring data. Among them reverse dosimetry represent one of most promising and feasible approaches from both a scientific and financial perspective as it can be performed to estimate the external exposure that is consistent with the measured biomonitoring data through the backward application of PBTK models.

With this approach, biomonitoring data can be effectively linked to an exposure scenario and compared to regulatory exposure guidance values. This facilitates the consideration of available biomonitoring data for screening for potential risks of chemicals. Relationships to the probabilistic distribution of exposure are equally intriguing. For example, typically those receiving the highest exposure are of greatest interest. The probabilistic modeling allows investigators to incorporate the inter-personal differences in human biochemistry and the uncertainty surrounding the duration and intensity of chemical exposures, resulting in a likely distribution of exposure. In this perspective knowledge of distribution permits a more in-depth evaluation of their exposure and/or dose, which can then be used to mitigate risk. Reverse dosimetry can serve as a useful model for how to place expanding biomonitoring information into a science-based framework that generates meaningful information for risk assessments and related decisions by regulators, the chemical industry, and the public.

From the methodological point of view, using an integrated exposure framework for linking mechanistically external and internal exposure, provides a comprehensive overview on how realistic exposure scenarios are translated into internal dose to humans, accounting for the age-dependent and route specific bioavailability differences. To this aim a generic PBTK modelling framework that captures lifetime internal exposure is a valuable tool with many applications in the modern risk assessment arena, by exploiting the continuously growing

wealth of in vitro testing and biomonitoring data. In addition, it provides the ground for the direct association among different types of environmental, exposure, biomonitoring as well as toxicity data, all of them essential for developing the individual exposome

The developed generic PBTK model includes i) lifespan evolution in physiology, from the moment of conception till 80 years of life-time and it is differentiated by gender; ii) a detailed description of pregnancy (mother-foetus interaction) and lactation (toxicants concentration in milk) periods, iii) a detailed compartmental description of human anatomy and receptor binding; iv) a detailed description of various exposure rotes (i.e. inhalation, dermal and oral). Moreover, the generic character of the model is ensured by the capability of assessing new chemicals or chemicals with limited information. To this aim, the model is linked to quantitative structure-activity relationships (QSARs), so as to calculate chemical-specific input parameters of PBTK models (partition coefficients and metabolic parameters such as the maximal velocity (Vmax) and Michaelis affinity constant (Km) or the intrinsic clearance (Vmax/Km).

Assessing exposure at multiple scales across the source-to-dose continuum, needs to take into account the actual complexity of the environmental and biological/physiological processes that are critical to the proper description of the phenomena involved. This results in targeted interventions and consequently more cost efficient risk management. In addition, a comprehensive integrated exposure framework estimating tissue dosimetry for the various relevant exposure scenarios, could be of great use in exploiting the in vitro HTS results rapidly produced by ToxCast21, advancing thus both exposure science and toxicology towards serving the needs of risk assessment in the 21st century.

Starting from the application of PBTK model and /or biomonitoring data collected, next step will be the quantitative estimation of the effect on human health due to exposure to selected chemicals.

To this aim internal doses will be coupled to health impacts on the local population through advanced statistical methods to derive the dose – response functions which account for differences in exposure patterns, susceptibility differences and inter-individual variation (due to lifestyle, age, sex or physiological status) in health response. To estimate the health impact we will use a statistical approach based on survey-weighted logistic multivariate regression adjusted for different covariates (age, sex, socio-economic status (SES) etc.) linking internal doses with health effect considering the interdependence of the covariates (using as metric an analogy of the “linkage disequilibrium” metric used in genome-wide association studies).

Overall, our comprehensive modelling framework supports the association of a variety of environmental, exposure and biomonitoring data, as well as the incorporation of recent advances of in vitro toxicology using high-throughput systems in the risk assessment process enhancing thus significantly the artillery of environmental health science and chemical safety

regulators . The necessity for associating these types of data is increased by the world-wide interest for “exposure based” risk assessment.

A current limitation for further introducing the PBBK models in the risk assessment arena is the lack of generic character of these models. By introducing a multi-route multi-compartmental PBBK model with many metabolites, eventually capturing perfusion limited as well as membrane limited blood flow, the limiting step of describing the ADME process of a large chemical space is the proper parameterization.

Under this scope, the use of QSARs for predicting the essential parameters for a broad chemical space is greatly facilitating this effort. Within HEALS, an originally new QSARs methodology was developed, coupling Abraham’s solvation equation to Artificial Neural Networks, yielding significantly improved results compared to the existing methodologies. Predictive capacity of the Abraham’s model for new chemical entities is influenced by chemical nature of the training set molecules used for development of the model. The test set molecules predicted very well when are structurally similar to the training set molecules. However, the methodology, seemed to perform adequately, also when compounds with different structures were treated as “data poor” chemicals. In any case, enrichment of databases with experimental data on the required parameters, will allow the better training of the model, increasing its predicting capability for data poor chemicals, even to the ones that do not share chemical structure similarities to the ones used in the training set database. The improved performance of the proposed QSARs methodology, relies on:

- The ability of Abrahams solvation equation to capture the main properties of the chemical structure that potentially affect the interactions to biological systems
- The “flexibility” of ANNs to represent mathematically the complex interactions of biochemical micro-processes.

There is also a sufficient number of HBM dataset collected in the frame of currently ongoing HBM programmes both in Europe and outside Europe which can be used for HEALS across a broad chemical space. Also for other validation exercises, including the reverse dosimetry approach, a sufficiently large database can be constructed from existing individual anonymized biomarker measurements.

8 References

Abraham, K., Mielke, H., Huisinga, W., Gundert-Remy, U., 2005. Elevated internal exposure of children in simulated acute inhalation of volatile organic compounds: Effects of concentration and duration. Archives of Toxicology. 79**,** 63-73.

Abraham, M. H., 1993. Application of solvation equations to chemical and biochemical processes. Pure and Applied Chemistry. 65**,** 2503-2512.

Angerer, J., Aylward, L. L., Hays, S. M., Heinzow, B., Wilhelm, M., 2011. Human biomonitoring assessment values: Approaches and data requirements. International Journal of Hygiene and Environmental Health.

Aylward, L. L., Hays, S. M., 2011. Biomonitoring-based risk assessment for hexabromocyclododecane (HBCD). Int J Hyg Environ Health. 214**,** 179-87.

Aylward, L. L., Hays, S. M., Gagné, M., Krishnan, K., 2009a. Derivation of Biomonitoring Equivalents for di-n-butyl phthalate (DBP), benzylbutyl phthalate (BzBP), and diethyl phthalate (DEP). Regulatory Toxicology and Pharmacology. 55**,** 259-267.

Aylward, L. L., Hays, S. M., Gagné, M., Krishnan, K., 2009b. Derivation of Biomonitoring Equivalents for di(2-ethylhexyl)phthalate (CAS No. 117-81-7). Regulatory Toxicology and Pharmacology. 55**,** 249-258.

Aylward, L. L., Hays, S. M., Gagne, M., Nong, A., Krishnan, K., 2010. Biomonitoring equivalents for hexachlorobenzene. Regul Toxicol Pharmacol. 58**,** 25-32.

Aylward, L. L., Krishnan, K., Kirman, C. R., Nong, A., Hays, S. M., 2011. Biomonitoring equivalents for deltamethrin. Regul Toxicol Pharmacol. 60**,** 189-99.

Aylward, L. L., Lakind, J. S., Hays, S. M., 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 Environ Health A. 71**,** 1499-508.

Bailin, P. S., Byrne, M., Lewis, S., Liroff, R., Public awareness drives market for safer alternatives: bisphenol A market analysis report., 2008.

Baláž, S., Lukáčová, V., 1999. A model-based dependence of the human tissue/blood partition coefficients of chemicals on lipophilicity and tissue composition. Quantitative Structure-Activity Relationships. 18**,** 361-368.

Ballard, P., Leahy, D. E., Rowland, M., 2003. Prediction of in vivo tissue distribution from in vitro data. 3. Correlation between in vitro and in vivo tissue distribution of a homologous series of nine 5-n-alkyl-5-ethyl barbituric acids. Pharmaceutical Research. 20**,** 864-872.

Bartels, M., Rick, D., Lowe, E., Loizou, G., Price, P., Spendiff, M., Arnold, S., Cocker, J., Ball, N., 2012. Development of PK- and PBPK-based modeling tools for derivation of biomonitoring guidance values. Computer Methods and Programs in Biomedicine. 108**,** 773-788.

Beaudouin, R., Micallef, S., Brochot, C., 2010. A stochastic whole-body physiologically based pharmacokinetic model to assess the impact of inter-individual variability on tissue dosimetry over the human lifespan. Regulatory Toxicology and Pharmacology. 57**,** 103-116.

Becker, K., Seiwert, M., Angerer, J., Heger, W., Koch, H. M., Nagorka, R., Roßkamp, E., Schlüter, C., Seifert, B., Ullrich, D., 2004. DEHP metabolites in urine of children and DEHP in house dust. International journal of hygiene and environmental health. 207**,** 409-417.

Béliveau, M., Tardif, R., Krishnan, K., 2003. Quantitative structure-property relationships for physiologically based pharmacokinetic modeling of volatile organic chemicals in rats. Toxicology and Applied Pharmacology. 189**,** 221-232.

Biedermann, S., Tschudin, P., Grob, K., 2010. Transfer of bisphenol A from thermal printer paper to the skin. Analytical and Bioanalytical Chemistry. 398**,** 571-576.

Bois, F. Y., Jamei, M., Clewell, H. J., 2010. PBPK modelling of inter-individual variability in the pharmacokinetics of environmental chemicals. Toxicology. 278**,** 256-267.

Boogaard, P. J., Hays, S. M., Aylward, L. L., 2011. Human biomonitoring as a pragmatic tool to support health risk management of chemicals - Examples under the EU REACH programme. Regulatory Toxicology and Pharmacology. 59**,** 125-132.

Bronzino, J. D., 2000. The Biomedical Engineering HandBook. CRC Press LLC -IEEE PRESS.

Bunge, A. L., Cleek, R. L., 1995. A new method for estimating dermal absorption from chemical exposure: 2. Effect of molecular weight and octanol-water partitioning. Pharmaceutical Research. 12**,** 88-95.

Cahill, T. M., Cousins, I., Mackay, D., 2003. Development and application of a generalized physiologically based pharmacokinetic model for multiple environmental contaminants. Environmental Toxicology and Chemistry. 22**,** 26-34.

CDC, Fourth National Report on Human Exposure to Environmental Chemicals. Department of Health and Human Services Centers for Disease Control and Prevention., Atlanta, GA, 2009.

Chen, C.-C., Shih, M.-C., Wu, K.-Y., 2010. Exposure estimation using repeated blood concentration measurements. Stochastic Environmental Research and Risk Assessment. 24**,** 445-454.

Cheng, S., Bois, F. Y., 2011. A mechanistic modeling framework for predicting metabolic interactions in complex mixtures. Environmental Health Perspectives. 119**,** 1712-1718.

Cleek, R. L., Bunge, A. L., 1993. A new method for estimating dermal absorption from chemical exposure. 1. General approach. Pharmaceutical Research. 10**,** 497-506.

Dennison, J. E., Andersen, M. E., Dobrev, I. D., Mumtaz, M. M., Yang, R. S., 2004. PBPK modeling of complex hydrocarbon mixtures: gasoline. Environ Toxicol Pharmacol. 16**,** 107-19.

Dudek, A. Z., Arodz, T., Gálvez, J., 2006. Computational methods in developing quantitative structure-activity relationships (QSAR): A review. Combinatorial Chemistry and High Throughput Screening. 9**,** 213-228.

Edginton, A. N., Ritter, L., 2009. Predicting plasma concentrations of bisphenol A in children younger than 2 years of age after typical feeding schedules, using a physiologically based toxicokinetic model. Environmental Health Perspectives. 117**,** 645-652.

Eissing, T., Kuepfer, L., Becker, C., Block, M., Coboeken, K., Gaub, T., Goerlitz, L., Jaeger, J., Loosen, R., Ludewig, B., Meyer, M., Niederalt, C., Sevestre, M., Siegmund, H. U., Solodenko, J., Thelen, K., Telle, U., Weiss, W., Wendl, T., Willmann, S., Lippert, J., 2011. A computational systems biology software platform for multiscale modeling and simulation: Integrating whole-body physiology, disease biology, and molecular reaction networks. Frontiers in Physiology. FEB.

EPA, Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures.EPA/630/R-00/002. Risk Assessment Forum Technical Panel., Washington DC, USA, 2000.

Erickson, B. E., 2008. Bisphenol a under scrutiny. Chemical and Engineering News. 86**,** 36-39.

Franco, A., Prevedouros, K., Alli, R., Cousins, I. T., 2007. Comparison and analysis of different approaches for estimating the human exposure to phthalate esters. Environment international. 33**,** 283-291.

Frasch, H. F., Barbero, A. M., 2003. Steady-State Flux and Lag Time in the Stratum Corneum Lipid Pathway: Results from Finite Element Models. Journal of Pharmaceutical Sciences. 92**,** 2196-2207.

Fromme, H., Bolte, G., Koch, H. M., Angerer, J., Boehmer, S., Drexler, H., Mayer, R., Liebl, B., 2007. Occurrence and daily variation of phthalate metabolites in the urine of an adult population. International journal of hygiene and environmental health. 210**,** 21-33.

Gelman, A., Roberts, G., Gilks, W., 1996. Efficient metropolis jumping hules. Bayesian statistics. 5**,** 599-608.

Gelman, A., Rubin, D. B., 1992. Inference from iterative simulation using multiple sequences. Statistical science. 457-472.

Gelman, A., Rubin, D. B., 1996. Markov chain Monte Carlo methods in biostatistics. Statistical methods in medical research. 5**,** 339-355.

Genz, A., Bretz, F., 2009. Computation of multivariate normal and t probabilities. Springer.

Georgopoulos, P. G., Balakrishnan, S., Roy, A., Isukapalli, S., Sasso, A., Chien, Y.-C., Weisel, C. P., A Comparison of Maximum Likelihood Estimation Methods for Inverse Problem Solutions Employing PBPK Modeling with Biomarker Data: Application to Tetrachloroethylene. CCL, 2008a.

Georgopoulos, P. G., Sasso, A. F., Isukapalli, S. S., Lioy, P. J., Vallero, D. A., Okino, M., Reiter, L., 2009. Reconstructing population exposures to environmental chemicals from biomarkers: Challenges and opportunities. Journal of Exposure Science and Environmental Epidemiology. 19**,** 149-171.

Georgopoulos, P. G., Wang, S. W., Yang, Y. C., Xue, J., Zartarian, V. G., McCurdy, T., Ozkaynak, H., 2008b. Biologically based modeling of multimedia, multipathway, multiroute population exposures to arsenic. Journal of Exposure Science and Environmental Epidemiology. 18**,** 462-476.

Gilks , W., Spiegelhalter , D., Richardson , S., 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall/CRC Press, Boca Raton.

Gilks, W. R., Roberts, G. O., Strategies for improving MCMC. Markov chain Monte Carlo in practice. Springer, 1996, pp. 89-114.

Gilks, W. R., Roberts, G. O., George, E. I., 1994. Adaptive direction sampling. The statistician. 179-189.

Haario, H., Laine, M., Mira, A., Saksman, E., 2006. DRAM: efficient adaptive MCMC. Statistics and Computing. 16**,** 339-354.

Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive Metropolis algorithm. Bernoulli. 223-242.

Haario, H., Saksman, E., Tamminen, J., 2005. Componentwise adaptation for high dimensional MCMC. Computational Statistics. 20**,** 265-273.

Haddad, S., Charest-Tardif, G., Krishnan, K., 2000. Physiologically based modeling of the maximal effect of metabolic interactions on the kinetics of components of complex chemical mixtures. Journal of Toxicology and Environmental Health - Part A. 61**,** 209-223.

Hays, S. M., Aylward, L. L., 2009. Using Biomonitoring Equivalents to interpret human biomonitoring data in a public health risk context. Journal of Applied Toxicology. 29**,** 275-288.

Hays, S. M., Aylward, L. L., Gagne, M., Krishnan, K., 2009. Derivation of Biomonitoring Equivalents for cyfluthrin. Regul Toxicol Pharmacol. 55**,** 268-75.

Hays, S. M., Aylward, L. L., Gagne, M., Nong, A., Krishnan, K., 2010. Biomonitoring equivalents for inorganic arsenic. Regul Toxicol Pharmacol. 58**,** 1-9.

Hays, S. M., Aylward, L. L., Kirman, C. R., Krishnan, K., Nong, A., 2011. Biomonitoring Equivalents for di-isononyl phthalate (DINP). Regulatory Toxicology and Pharmacology. 60**,** 181-188.

Hays, S. M., Nordberg, M., Yager, J. W., Aylward, L. L., 2008. Biomonitoring Equivalents (BE) dossier for cadmium (Cd) (CAS No. 7440-43-9). Regulatory Toxicology and Pharmacology. 51**,** S49-S56.

Hays, S. M., Pyatt, D. W., Kirman, C. R., Aylward, L. L., 2012. Biomonitoring Equivalents for benzene. Regulatory Toxicology and Pharmacology. 62**,** 62-73.

Hertzberg, R. C., Rice, G., Teuschler, L. K., Methods for health risk assessment of combustion mixtures. In: Hazardous waste incineration: evaluating the human health and environmental risks. In:

S. Roberts, et al., Eds.), Hazardous waste incineration: evaluating the human health and environmental risks. CRC Press LLC, 1999, pp. 105-148.

ICPR, Basic anatomical and physiological data for use in radiological protection: reference values. In: J. Valentin, (Ed.), The International Commission on Radiological Protection, 2002.

Isukapalli, S., Roy, A., Georgopoulos, P., 2000. Efficient sensitivity/uncertainty analysis using the combined stochastic response surface method and automated differentiation: Application to environmental and biological systems. Risk Analysis. 20**,** 591-602.

Johnson, M. E., Blankschtein, D., Langer, R., 1997. Evaluation of solute permeation through the stratum corneum: Lateral bilayer diffusion as the primary transport mechanism. Journal of Pharmaceutical Sciences. 86**,** 1162-1172.

Jongeneelen, F. J., Berge, W. F. T., 2011. A generic, cross-chemical predictive PBTK model with multiple entry routes running as application in MS Excel; design of the model and comparison of predictions with experimental results. Annals of Occupational Hygiene. 55**,** 841-864.

Judson, R. S., Houck, K. A., Kavlock, R. J., Knudsen, T. B., Martin, M. T., Mortensen, H. M., Reif, D. M., Rotroff, D. M., Shah, I., Richard, A. M., Dix, D. J., 2010. In vitro screening of environmental chemicals for targeted testing prioritization: The toxcast project. Environmental Health Perspectives. 118**,** 485-492.

Judson, R. S., Kavlock, R. J., Setzer, R. W., Cohen Hubal, E. A., Martin, M. T., Knudsen, T. B., Houck, K. A., Thomas, R. S., Wetmore, B. A., Dix, D. J., 2011. Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. Chemical Research in Toxicology. 24**,** 451-462.

Keys, D. A., Wallace, D. G., Kepler, T. B., Conolly, R. B., 1999. Quantitative evaluation of alternative mechanisms of blood and testes disposition of Di(2-ethylhexyl) phthalate and mono(2-ethylhexyl) phthalate in rats. Toxicological Sciences. 49**,** 172-185.

Kirman, C. R., Aylward, L. L., Hays, S. M., Krishnan, K., Nong, A., 2011. Biomonitoring Equivalents for DDT/DDE. Regulatory Toxicology and Pharmacology. 60**,** 172-180.

Krauss, M., Schaller, S., Borchers, S., Findeisen, R., Lippert, J., Kuepfer, L., 2012. Integrating Cellular Metabolism into a Multiscale Whole-Body Model. PLoS Computational Biology. 8.

Krishnan, K., Adamou, T., Aylward, L. L., Hays, S. M., Kirman, C. R., Nong, A., 2011. Biomonitoring equivalents for 2,2',4,4',5-pentabromodiphenylether (PBDE-99). Regul Toxicol Pharmacol. 60**,** 165-71.

Krishnan, K., Gagné, M., Nong, A., Aylward, L. L., Hays, S. M., 2010a. Biomonitoring Equivalents for bisphenol A (BPA). Regulatory Toxicology and Pharmacology. 58**,** 18-24.

Krishnan, K., Gagné, M., Nong, A., Aylward, L. L., Hays, S. M., 2010b. Biomonitoring Equivalents for triclosan. Regulatory Toxicology and Pharmacology. 58**,** 10-17.

L. Holmes, R. W., JD Galambos, DJ Strickler, S, 2000. A method for optimization of pharmacokinetic models. Toxicology Mechanisms and Methods. 10**,** 41-53.

Laskey, K. B., Myers, J. W., 2003. Population markov chain monte carlo. Machine Learning. 50**,** 175-196.

Liang, F., 2002. Dynamically weighted importance sampling in Monte Carlo computation. Journal of the American Statistical Association. 97**,** 807-821.

Liang, F., Wong, W. H., 2001. Real-parameter evolutionary Monte Carlo with applications to Bayesian mixture models. Journal of the American Statistical Association. 96**,** 653-666.

Liao, C., Kannan, K., 2011. High levels of bisphenol A in paper currencies from several countries, and implications for dermal exposure. Environmental Science and Technology. 45**,** 6761-6768.

Lorber, M., Angerer, J., Koch, H. M., 2010. A simple pharmacokinetic model to characterize exposure of Americans to Di-2-ethylhexyl phthalate. Journal of Exposure Science and Environmental Epidemiology. 20**,** 38-53.

Lorz, P. M., Towae, F. K., Enke, W., Jäckh, R., Bhargava, N., Hillesheim, W., 2007. Phthalic acid and derivatives. Ullmann's Encyclopedia of Industrial Chemistry.

Lyons, M. A., Yang, R. S., Mayeno, A. N., Reisfeld, B., 2008. Computational toxicology of chloroform: reverse dosimetry using Bayesian inference, Markov chain Monte Carlo simulation, and human biomonitoring data. Environmental health perspectives. 116**,** 1040-1046.

Matthews, J. B., Twomey, K., Zacharewski, T. R., 2001. In vitro and in vivo interactions of bisphenol A and its metabolite, bisphenol A glucuronide, with estrogen receptors α and β. Chemical Research in Toxicology. 14**,** 149-157.

McNally, K., Cotton, R., Cocker, J., Jones, K., Bartels, M., Rick, D., Price, P., Loizou, G., 2012. Reconstruction of exposure to m-xylene from human biomonitoring data using PBPK modelling, bayesian inference, and Markov chain Monte Carlo simulation. Journal of Toxicology. 2012.

McNally, K., Cotton, R., Hogg, A., Loizou, G., 2014. PopGen: A virtual human population generator. Toxicology. 315**,** 70-85.

Mendum, T., Stoler, E., Vanbenschoten, H., Warner, J. C., 2011. Concentration of bisphenol a in thermal paper. Green Chemistry Letters and Reviews. 4**,** 81-86.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., Teller, E., 1953. Equation of state calculations by fast computing machines. The journal of chemical physics. 21**,** 1087-1092.

Mira, A., 2001. On Metropolis-Hastings algorithms with delayed rejection. Metron. 59**,** 231-241.

Mitragotri, S., Anissimov, Y. G., Bunge, A. L., Frasch, H. F., Guy, R. H., Hadgraft, J., Kasting, G. B., Lane, M. E., Roberts, M. S., 2011. Mathematical models of skin permeability: An overview. International Journal of Pharmaceutics. 418**,** 115-129.

Moles, C. G., Mendes, P., Banga, J. R., 2003. Parameter estimation in biochemical pathways: a comparison of global optimization methods. Genome research. 13**,** 2467-2474.

Mumtaz, M., Hertzberg, R., The status of interactions data in risk assessment of chemical mixtures. In: J. Saxena, (Ed.), Hazard Assessment of Chemicals. Hemisphere puPublishing Corporation, Washington DC, USA, 1993, pp. 47-79.

Muzic Jr, R. F., Christian, B. T., 2006. Evaluation of objective functions for estimation of kinetic parameters. Medical physics. 33**,** 342-353.

Nitsche, J. M., Wang, T. F., Kasting, G. B., 2006. A two-phase analysis of solute partitioning into the stratum corneum. Journal of Pharmaceutical Sciences. 95**,** 649-666.

NRC, Human Biomonitoring for Environmental Chemicals. The National Academies Press, Washington, DC., 2006.

Paixão, P., Aniceto, N., Gouveia, L. F., Morais, J. A. G., 2013. Tissue-to-blood distribution coefficients in the rat: Utility for estimation of the volume of distribution in man. European Journal of Pharmaceutical Sciences. 50**,** 526-543.

Pelekis, M., Krishnan, K., 2004. Magnitude and mechanistic determinants of the interspecies toxicokinetic uncertainty factor for organic chemicals. Regulatory Toxicology and Pharmacology. 40**,** 264-271.

Peyret, T., Krishnan, K., 2011. QSARs for PBPK modelling of environmental contaminants. SAR and QSAR in Environmental Research. 22**,** 129-169.

Peyret, T., Poulin, P., Krishnan, K., 2010. A unified algorithm for predicting partition coefficients for PBPK modeling of drugs and environmental chemicals. Toxicology and Applied Pharmacology. 249**,** 197-207.

Poulin, P., Krishnan, K., 1995. An algorithm for predicting tissue: Blood partition coefficients of organic chemicals from n-octanol: Water partition coefficient data. Journal of Toxicology and Environmental Health. 46**,** 117-129.

Price, K., Krishnan, K., 2011. An integrated QSAR-PBPK modelling approach for predicting the inhalation toxicokinetics of mixtures of volatile organic chemicals in the rat. SAR and QSAR in Environmental Research. 22**,** 107-128.

Program, N. T., 1982. Carcinogenesis Bioassay of Di (2-ethylhexyl) phthalate (CAS No. 117-81-7) in F344 Rats and B6C3F1 Mice (Feed Studies). National Toxicology Program technical report series. 217**,** 1.

Ramsey, J. C., Andersen, M. E., 1984. A physiologically based description of the inhalation pharmacokinetics of styrene in rats and humans. Toxicology and Applied Pharmacology. 73**,** 159-175.

Reddy, M. B., Yang, R. S. H., Clewell III, H. J., Andersen, M. E., 2005. Physiologically Based Pharmacokinetic Modeling: Science and Applications. John Wiley & Sons.

Rim, J. E., Pinsky, P. M., van Osdol, W. W., 2007. Using the method of homogenization to calculate the effective diffusivity of the stratum corneum. Journal of Membrane Science. 293**,** 174-182.

Robert, C., Casella, G., Metropolis–Hastings Algorithms. Introducing Monte Carlo Methods with R. Springer New York, 2010, pp. 167-197.

Roberts, G. O., Gelman, A., Gilks, W. R., 1997. Weak convergence and optimal scaling of random walk Metropolis algorithms. The annals of applied probability. 7**,** 110-120.

Roberts, G. O., Rosenthal, J. S., 2009. Examples of adaptive MCMC. Journal of Computational and Graphical Statistics. 18**,** 349-367.

Rodgers, T., Rowland, M., 2006. Physiologically based pharmacokinetic modelling 2: Predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. Journal of Pharmaceutical Sciences. 95**,** 1238-1257.

Roy, A., Weisel, C. P., Gallo, M., Georgopoulos, P., 1996. Studies of multiroute exposure/dose reconstruction using physiologically based pharmacokinetic models. Journal of Clean Technology, Environmental Toxicology and Occupational Medicine. 5**,** 285-295.

Sarigiannis, D., Gotti, A., Karakitsios, S., 2011. A Computational Framework for Aggregate and Cumulative Exposure Assessment. Epidemiology. 22**,** S96-S97.

Sarigiannis, D., Karakitsios, S., 2011. Perinatal Exposure to Bisphenol A: The Route of Administration Makes the Dose. Epidemiology. 22**,** S172.

Sarigiannis, D. A., Gotti, A., 2008. Biology-based dose-response models for health risk assessment of chemical mixtures. Fresenius Environmental Bulletin. 17**,** 1439-1451.

Sasso, A. F., Isukapalli, S. S., Georgopoulos, P. G., 2010. A generalized physiologically-based toxicokinetic modeling system for chemical mixtures containing metals. Theoretical Biology and Medical Modelling. 7.

Schulz, C., Angerer, J., Ewers, U., Kolossa-Gehring, M., 2007. The German Human Biomonitoring Commission. International Journal of Hygiene and Environmental Health. 210**,** 373-382.

Schulz, C., Wilhelm, M., Heudorf, U., Kolossa-Gehring, M., 2011. Update of the reference and HBM values derived by the German Human Biomonitoring Commission. Int J Hyg Environ Health. 215**,** 26-35.

Smolders, R., Schoeters, G., 2007. Identifying opportunities and gaps for establishing an integrated EDR-triad at a European level. International Journal of Hygiene and Environmental Health. 210**,** 253-257.

Sparacino, G., Pillonetto, G., Capello, M., De Nicolao, G., Cobelli, C., 2002. WINSTODEC: a stochastic deconvolution interactive program for physiological and pharmacokinetic systems. Computer methods and programs in biomedicine. 67**,** 67-77.

Storn, R., Price, K., 1995. Differential evolution-a simple and efficient adaptive scheme for global optimization over continuous spaces. ICSI Berkeley.

Storn, R., Price, K., 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of global optimization. 11**,** 341-359.

Talreja, P. S., Kasting, G. B., Kleene, N. K., Pickens, W. L., Wang, T. F., 2001. Visualization of the lipid barrier and measurement of lipid pathlength in human stratum corneum. AAPS Journal. 3**,** XVII-XVIII.

Tan, Y. M., Liao, K., Conolly, R., Blount, B., Mason, A., Clewell, H., 2006. Use of a physiologically based pharmacokinetic model to identify exposures consistent with human biomonitoring data for chloroform. Journal of Toxicology and Environmental Health - Part A: Current Issues. 69**,** 1727-1756.

Tanaka, C., Kawai, R., Rowland, M., 2000. Dose-dependent pharmacokinetics of cyclosporin a in rats: Events in tissues. Drug Metabolism and Disposition. 28**,** 582-589.

Ter Braak, C. J., 2006. A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. Statistics and Computing. 16**,** 239-249.

ter Braak, C. J., Vrugt, J. A., 2008. Differential evolution Markov chain with snooker updater and fewer chains. Statistics and Computing. 18**,** 435-446.

Tickner, J. A., Schettler, T., Guidotti, T., McCally, M., Rossi, M., 2001. Health risks posed by use of Di ‐ 2 ‐ ethylhexyl phthalate (DEHP) in PVC medical devices: A critical review. American journal of industrial medicine. 39**,** 100-111.

Tierney, L., 1994. Markov chains for exploring posterior distributions. the Annals of Statistics. 1701-1728.

Touitou, E., 2002. Drug delivery across the skin. Expert Opinion on Biological Therapy. 2**,** 723-733.

USEPA, ToxCast™ Data, http://www.epa.gov/ncct/toxcast/data.html. Vol. 2013.

Valcke, M., Krishnan, K., 2011. Evaluation of the impact of the exposure route on the human kinetic adjustment factor. Regulatory Toxicology and Pharmacology. 59**,** 258-269.

Valentin, J., 2002. Basic anatomical and physiological data for use in radiological protection: reference values: ICRP Publication 89. Annals of the ICRP. 32**,** 1-277.

Vandenberg, L. N., Maffini, M. V., Sonnenschein, C., Rubin, B. S., Soto, A. M., 2009. Bisphenol-a and the great divide: A review of controversies in the field of endocrine disruption. Endocrine Reviews. 30**,** 75-95.

Verner, M. A., Charbonneau, M., Lopez-Carrillo, L., Haddad, S., 2008. Physiologically based pharmacokinetic modeling of persistent organic pollutants for lifetime exposure assessment: A new tool in breast cancer epidemiologic studies. Environmental Health Perspectives. 116**,** 886-892.

Viñas, P., López-García, I., Campillo, N., Rivas, R. E., Hernández-Córdoba, M., 2012. Ultrasound-assisted emulsification microextraction coupled with gas chromatography-mass spectrometry using the Taguchi design method for bisphenol migration studies from thermal printer paper, toys and baby utensils. Analytical and Bioanalytical Chemistry. 404**,** 671-678.

Völkel, W., Colnot, T., Csanády, G. A., Filser, J. G., Dekant, W., 2002. Metabolism and kinetics of bisphenol a in humans at low doses following oral administration. Chemical Research in Toxicology. 15**,** 1281-1287.

Wang, T. F., Kasting, G. B., Nitsche, J. M., 2006. A multiphase microscopic diffusion model for stratum corneum permeability. I. Formulation, solution, and illustrative results for representative compounds. Journal of Pharmaceutical Sciences. 95**,** 620-648.

Willmann, S., Lippert, J., Sevestre, M., Solodenko, J., Fois, F., Schmitt, W., 2003. PK-Sim®: A physiologically based pharmacokinetic 'whole-body' model. Drug Discovery Today: BIOSILICO. 1**,** 121-124.

Wormuth, M., Scheringer, M., Vollenweider, M., Hungerbühler, K., 2006. What are the sources of exposure to eight frequently used phthalic acid esters in Europeans? Risk Analysis. 26**,** 803-824.

Ya-Xian, Z., Suetake, T., Tagami, H., 1999. Number of cell layers of the stratum corneum in normal skin relationship to the anatomical location an the body, age, sex and physical parameters. Archives of Dermatological Research. 291**,** 555-559.

Yang, Y., Xu, X., Georgopoulos, P. G., 2010. A Bayesian population PBPK model for multiroute chloroform exposure. Journal of Exposure Science and Environmental Epidemiology. 20**,** 326-341.

Zhang, H., 2004. A new nonlinear equation for the tissue/blood partition coefficients of neutral compounds. Journal of Pharmaceutical Sciences. 93**,** 1595-1604.