Cross-Mediterranean Environment and Health Network

CROME-LIFE

Deliverable B.4.2

Technical report detailing reverse dosimetry methodology for exposure reconstruction

LIFE ENVIRONMENT PROGRAMME

LIFE12 ENV/GR/001040

Action: B.4

TASK: 4.4

Report Date: 31/03/2016

http://www.crome-life.eu

Bibliographical Information

Project: Cross-Mediterranean Environment and Health Network – CROME-LIFE

Subject: Technical report detailing reverse dosimetry methodology for exposure reconstruction

LIFE ENVIRONMENT PROGRAMME
Contract No. **LIFE12 ENV/GR/001040**
Duration of Contract: 01/07/2013 - 31/12/2016
ACTION: B.4- Exposure assessment
TASK: 4.4 - Reverse dosimetry
Editing Partner: AUTH
Other Partners:
Report Date:31/03/2016

Pages:18 (including figures, tables, attachments)

Key Words: exposure reconstruction, reverse dosimetry, Bayesian Markov Chain Monte Carlo, human biomonitoring

Contact Person Editing Partner
Name: Prof. Dimosthenis Sarigiannis
Phone: +30 2310994562
Fax:
e-mail: sarigiannis@auth.gr

Authors Editing Partner
Name:
Phone:
Fax:
e-mail:

CROME-LIFE web site: http://www.crome-life.eu

Table of Content

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 1).

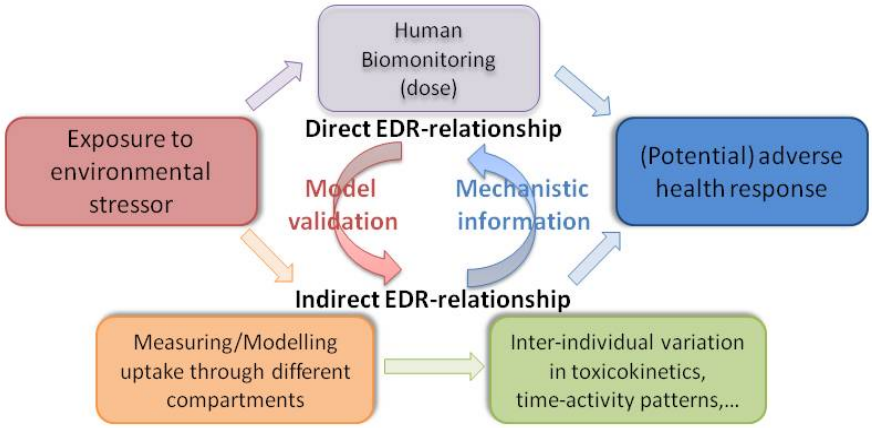

Figure 1. The Exposure-Dose-Response Triad to evaluate the potential adverse health effects of exposure to environmental agents (adapted from Smolders and Schoeters (2007))

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 2) for linking biomonitoring data to health outcomes: direct comparison to toxicity values, forward and reverse dosimetry.

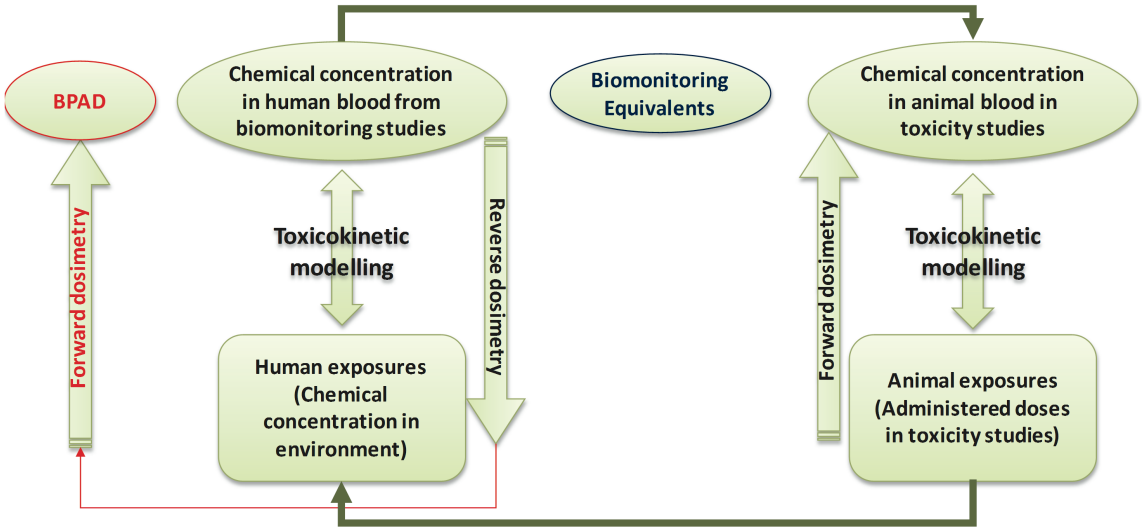

Figure 2. Interpretation of biomonitoring data

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 PBPK 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, Houck et al. 2010, Judson, Kavlock 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.

2. Exposure reconstruction modelling framework

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 PBTK model can be subdivided into Bayesian and non-Bayesian approaches (Georgopoulos, Balakrishnan et al. 2008). Moreover, the computational inversion techniques (and exposure reconstruction techniques as well), can be classified as deterministic or stochastic (Moles, Mendes 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 pharmacokinetic parameters. Moreover, a gradient method has been used by Isukapalli et al. (2000) calculating the uncertainty in PBTKs. A maximum likelihood method has been carried out for short and long term for exposure reconstruction using a PBTK for chloroform (Roy, Weisel 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,

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

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

The posterior distribution of the biomarker measurements will be

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

and

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

Hence taken into account,

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

Therefore,

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

And

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

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

Therefore,

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

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

2.1.1 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 , Spiegelhalter 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 PBTK models for inverse modeling (Lyons, Yang et al. 2008, Georgopoulos, Sasso et al. 2009, Chen, Shih et al. 2010, McNally, Cotton et al. 2012, McNally, Cotton et al. 2014)

2.1.2 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.

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.

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 $(X_0,...,X_n, \theta_0,..., \theta_n)$ 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, Saksman et al. 2001).

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.

2.1.3 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 1 presents the available MCMC algorithms based on Metropolis Hasting sampling that can been used.

Table 1. 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, Laine et al. 2006)
Adaptive Metropolis(AM)	Multivariate	(Haario, Saksman et al. 2001)
Componentwise Metropolis (CHM)	Componentwise	(Haario, Saksman 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, Laine 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, Saksman 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, Rosenbluth 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, Roberts 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, Gelman et al. 1997).

2.2 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).

2.2.1 DEMC algorithms

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

Table 2. 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 0. Also the snooker update fraction (Gilks, Roberts 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.

2.3 Methodology in CROME-LIFE and selected algorithm

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

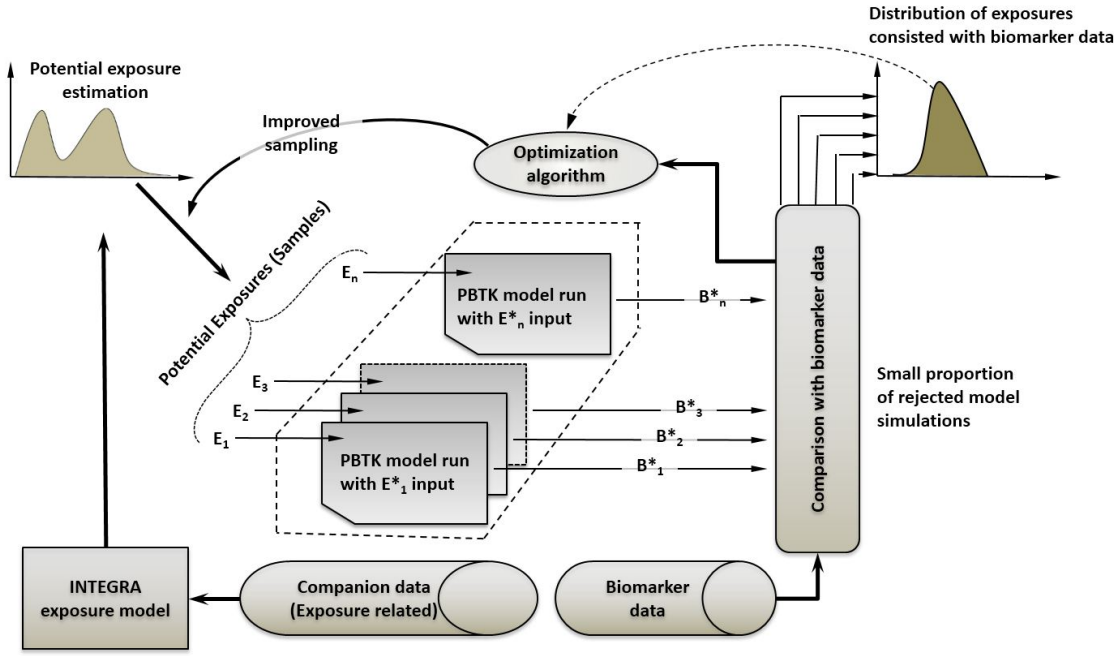

Figure 3. Optimization-aided exposure reconstruction based on HBM data using time-evolving PBTK models (figure adapted from Georgopoulos et al. (2009)).

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 3 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 4.

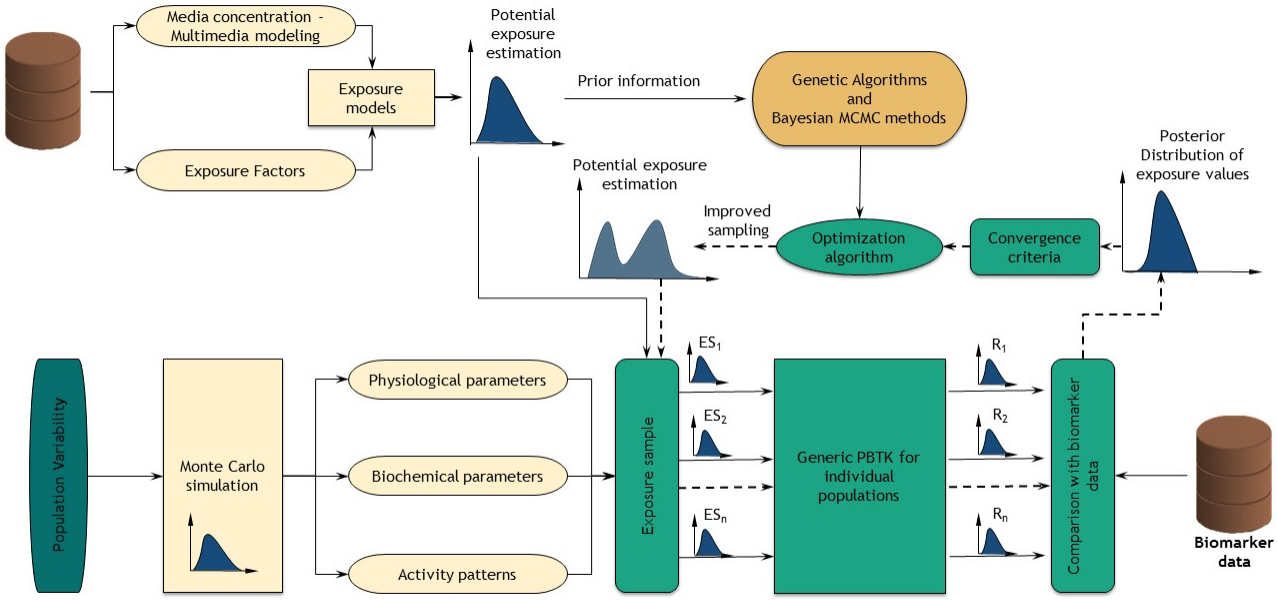

Figure 4. Exposure reconstruction flowchart procedure

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.

3. References

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

Chen, C.-C., et al. (2010). "Exposure estimation using repeated blood concentration measurements." Stochastic Environmental Research and Risk Assessment **24**(3): 445-454.

Gelman, A., et al. (1996). "Efficient metropolis jumping hules." Bayesian statistics **5**: 599-608.

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

Gelman, A. and D. B. Rubin (1996). "Markov chain Monte Carlo methods in biostatistics." Statistical methods in medical research **5**(4): 339-355.

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

Georgopoulos, P. G., et al. (2008). A Comparison of Maximum Likelihood Estimation Methods for Inverse Problem Solutions Employing PBPK Modeling with Biomarker Data: Application to Tetrachloroethylene, CCL.

Georgopoulos, P. G., et al. (2009). "Reconstructing population exposures to environmental chemicals from biomarkers: Challenges and opportunities." Journal of Exposure Science and Environmental Epidemiology **19**(2): 149-171.

Gilks , W., et al. (1996). Markov Chain Monte Carlo in Practice. Boca Raton, Chapman and Hall/CRC Press.

Gilks, W. R. and G. O. Roberts (1996). Strategies for improving MCMC. Markov chain Monte Carlo in practice, Springer**:** 89-114.

Gilks, W. R., et al. (1994). "Adaptive direction sampling." The statistician: 179-189.

Haario, H., et al. (2006). "DRAM: efficient adaptive MCMC." Statistics and Computing **16**(4): 339-354.

Haario, H., et al. (2001). "An adaptive Metropolis algorithm." Bernoulli: 223-242.

Haario, H., et al. (2005). "Componentwise adaptation for high dimensional MCMC." Computational Statistics **20**(2): 265-273.

Isukapalli, S., et al. (2000). "Efficient sensitivity/uncertainty analysis using the combined stochastic response surface method and automated differentiation: Application to environmental and biological systems." Risk Analysis **20**(5): 591-602.

Judson, R. S., et al. (2010). "In vitro screening of environmental chemicals for targeted testing prioritization: The toxcast project." Environmental Health Perspectives **118**(4): 485-492.

Judson, R. S., et al. (2011). "Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment." Chemical Research in Toxicology **24**(4): 451-462.

L. Holmes, R. W., JD Galambos, DJ Strickler, S (2000). "A method for optimization of pharmacokinetic models." Toxicology Mechanisms and Methods **10**(1): 41-53.

Laskey, K. B. and J. W. Myers (2003). "Population markov chain monte carlo." Machine Learning **50**(1-2): 175-196.

Liang, F. (2002). "Dynamically weighted importance sampling in Monte Carlo computation." Journal of the American Statistical Association **97**(459): 807-821.

Liang, F. and W. H. Wong (2001). "Real-parameter evolutionary Monte Carlo with applications to Bayesian mixture models." Journal of the American Statistical Association **96**(454): 653-666.

Lyons, M. A., et al. (2008). "Computational toxicology of chloroform: reverse dosimetry using Bayesian inference, Markov chain Monte Carlo simulation, and human biomonitoring data." Environ Health Perspect **116**(8): 1040-1046.

McNally, K., et al. (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., et al. (2014). "PopGen: A virtual human population generator." Toxicology **315**(1): 70-85.

Metropolis, N., et al. (1953). "Equation of state calculations by fast computing machines." The journal of chemical physics **21**(6): 1087-1092.

Mira, A. (2001). "On Metropolis-Hastings algorithms with delayed rejection." Metron **59**(3-4): 231-241.

Moles, C. G., et al. (2003). "Parameter estimation in biochemical pathways: a comparison of global optimization methods." Genome research **13**(11): 2467-2474.

Muzic Jr, R. F. and B. T. Christian (2006). "Evaluation of objective functions for estimation of kinetic parameters." Medical physics **33**(2): 342-353.

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

Robert, C. and G. Casella (2010). Metropolis–Hastings Algorithms. Introducing Monte Carlo Methods with R, Springer New York**:** 167-197.

Roberts, G. O., et al. (1997). "Weak convergence and optimal scaling of random walk Metropolis algorithms." The annals of applied probability **7**(1): 110-120.

Roberts, G. O. and J. S. Rosenthal (2009). "Examples of adaptive MCMC." Journal of Computational and Graphical Statistics **18**(2): 349-367.

Roy, A., et al. (1996). "Studies of multiroute exposure/dose reconstruction using physiologically based pharmacokinetic models." Journal of Clean Technology, Environmental Toxicology and Occupational Medicine **5**(4): 285-295.

Smolders, R. and G. Schoeters (2007). "Identifying opportunities and gaps for establishing an integrated EDR-triad at a European level." International Journal of Hygiene and Environmental Health **210**(3–4): 253-257.

Sparacino, G., et al. (2002). "WINSTODEC: a stochastic deconvolution interactive program for physiological and pharmacokinetic systems." Computer Methods and Programs in Biomedicine **67**(1): 67-77.

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

Storn, R. and K. Price (1997). "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces." Journal of global optimization **11**(4): 341-359.

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**(3): 239-249.

ter Braak, C. J. and J. A. Vrugt (2008). "Differential evolution Markov chain with snooker updater and fewer chains." Statistics and Computing **18**(4): 435-446.

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