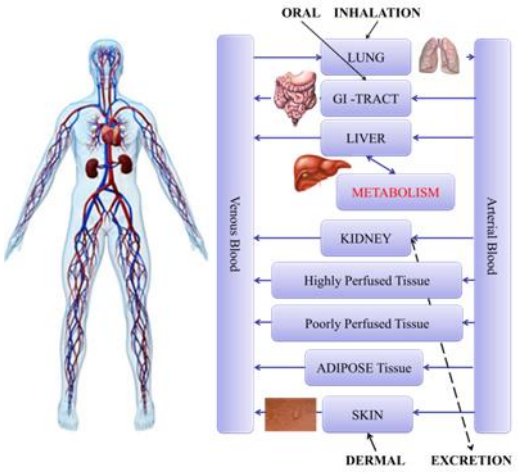
European Commission

JRC CONFERENCE AND WORKSHOP REPORTS

# EURL ECVAM WORKSHOP ON NEW GENERATION OF PHYSIOLOGICALLY-BASED KINETIC MODELS IN RISK ASSESSMENT

Paini A, Joossens E, Bessems J, Desalegn A, Dorne JL, Gosling JP, Heringa MB, Klaric M, Kramer N, Loizou G, Louisse J, Lumen A, Madden JC, Patterson EA, Proença S, Punt A, Setzer RW, Suciu N, Troutman J, Yoon M, Worth A, Tan YM.




2017

Joint Research Centre

EUR 28794 EN

**Contact information**
Name: Alicia Paini
Address: E. Fermi 2749, I-21027 ISPRA (VA)
Email: alicia.paini@ec.europa.eu
Tel.: +39 0332783986

Name: Andrew Worth
Address: E. Fermi 2749, I-21027 ISPRA (VA)
Email: andrew.worth@ec.europa.eu
Tel.: +39 0332789963

**JRC Science Hub**
https://ec.europa.eu/jrc

JRC108231

EUR 28794 EN

| | | | |
|---|---|---|---|
| PDF | ISBN 978-92-79-73848-7 | ISSN 1831-9424 | doi:10.2760/619902 |
| Print | ISBN 978-92-79-73849-4 | ISSN 1018-5593 | doi:10.2760/035371 |

Luxembourg: Publications Office of the European Union, 2017



How to cite this report: Paini et al., EURL ECVAM WORKSHOP ON NEW GENERATION OF PHYSIOLOGICALLY-BASED KINETIC MODELS IN RISK ASSESSMENT, EUR 28794 EN, Publications Office of the European Union, Luxembourg ,2017, ISBN ISBN 978-92-79-73849-4, doi 10.2760/035371, PUBSY No. JRC108231

# Contents

## Acknowledgements

We would like to thank the great job that Eva AHS and Georgia TOSIOU did in the logistic organization of the workshop. Many thanks to Adelaide Dura for the remake of figures in HQ.




*Participants to the* EURL ECVAM two-day workshop entitled "PHYSIOLOGICALLY-BASED KINETIC MODELLING IN RISK ASSESSMENT – REACHING A WHOLE NEW LEVEL IN REGULATORY DECISION-MAKING" held on the 16$^{th}$ -17$^{th}$ of November 2016 at the Joint Research Centre (JRC) in Ispra, Italy.

Bessems J, Corvi, R., Desalegn A, *Dorne JL,* Gosling JP, Heringa MB, Joossens E, Klaric M, Kramer N, Loizou G, Louisse J, *Lumen A,* Madden JC, Patterson EA, Paini A, Proença S, Punt A, Richarz A, Setzer RW*, Suciu N*, Tan C, Troutman J, Yoon M, Worth A.

**Authors**

Alicia Paini[1], Elisabeth Joossens[1], Jos Bessems[1], Anteneh Desalegn[1], Jean-Lou Dorne[3], John Paul Gosling[4], Minne B. Heringa[5], Martina Klaric[6], Nynke Kramer [7], George Loizou[8], Jochem Louisse[9], Annie Lumen[10], Judith C. Madden[11], Eann A. Patterson[12], Susana Proença[1], Ans Punt[13], R. Woodrow Setzer[2], Nicoleta Suciu[14], John Troutman[15], Miyoung Yoon[16], Andrew Worth[1], Yu-Mei Tan[2]

Affiliations

[1] European Commission Joint Research Centre, Directorate Health, Consumers and Reference Materials, European Commission, Joint Research Centre, Ispra, Italy.

[2] U.S. Environmental Protection Agency, National Exposure Research Laboratory, 109 TW Alexander Dr, Research Triangle Park, NC 27709, USA.

[3] European Food Safety Authority, 1a, Via Carlo Magno, 1A, 43126 Parma PR, Italy.

[4] School of Mathematics, University of Leeds, Leeds, U.K.

[5] RIVM - The National Institute for Public Health and the Environment, Bilthoven, The Netherlands.

[6] Cosmetics Europe, Brussels, Belgium

[7] Institute for Risk Assessment Sciences, Utrecht University, P.O. Box 80177, 3508TD Utrecht, The Netherlands.

[8] Health and Safety Laboratory, Buxton, UK

[9] Division of Toxicology, Wageningen University, Tuinlaan 5, 6703 HE Wageningen, The Netherlands.

[10] Division of Biochemical Toxicology, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR 72079, USA.

[11] School of Pharmacy and Bimolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK.

[12] School of Engineering, University of Liverpool, Liverpool, L69 3GH, UK.

[13.] Wageningen University and Research Centre, RIKILT Institute of Food Safety, Akkermaalsbos 2, 6708 WB, Wageningen, The Netherlands.

[14.] Istituto di Chimica Agraria ed Ambientale, Università Cattolica del Sacro Cuore, Piacenza, Italy.

[15.] Central Product Safety, The Procter & Gamble Company, Cincinnati, OH, USA.

[16] ScitoVation, 6 Davis Drive, PO Box 110566, Research Triangle Park, NC 27709, United States

# Abstract

The European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) Strategy Document on Toxicokinetics (TK) outlines strategies to enable prediction of systemic toxicity by applying new approach methodologies (NAM). The central feature of the strategy focuses on using physiologically-based kinetic (PBK) modelling to integrate data generated by *in vitro* and *in silico* methods for absorption, distribution, metabolism, and excretion (ADME) in humans for predicting whole-body TK behaviour, for environmental chemicals, drugs, nano-materials, and mixtures. In order to facilitate acceptance and use of this new generation of PBK models, which do not rely on animal/human *in vivo* data in the regulatory domain, experts were invited by EURL ECVAM to (i) identify current challenges in the application of PBK modelling to support regulatory decision making; (ii) discuss challenges in constructing models with no *in vivo* kinetic data and opportunities for estimating parameter values using *in vitro* and *in silico* methods; (iii) present the challenges in assessing model credibility relying on non-animal data and address strengths, uncertainties and limitations in such an approach; (iv) establish a *good kinetic modelling practice* workflow to serve as the foundation for guidance on the generation and use of *in vitro* and *in silico* data to construct PBK models designed to support regulatory decision making.

To gauge the current state of PBK applications, experts were asked upfront of the workshop to fill a short survey. In the workshop, using presentations and discussions, the experts elaborated on the importance of being transparent about the model construct, assumptions, and applications to support assessment of model credibility. The experts offered several recommendations to address commonly perceived limitations of parameterization and evaluation of PBK models developed using non-animal data and its use in risk assessment, these include: (i) develop a decision tree for model construction; (ii) set up a task force for independent model peer review; (iii) establish a scoring system for model evaluation; (iv) attract additional funding to develop accessible modelling software.; (v) improve and facilitate communication between scientists (model developers, data provider) and risk assessors/regulators; and (vi) organise specific training for end users. The experts also acknowledged the critical need for developing a guidance document on building, characterising, reporting and documenting PBK models using non-animal data. This document would also need to include guidance on interpreting the model analysis for various risk assessment purposes, such as incorporating PBK models in integrated strategy approaches and integrating them with *in vitro* toxicity testing and adverse outcome pathways. This proposed guidance document will promote the development of PBK models using in vitro and silico data and facilitate the regulatory acceptance of PBK models for assessing safety of chemicals.

# 1 Introduction

The European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) Strategy Document on Toxicokinetics (TK)[1] outlines the strategies to enable prediction of systemic toxicity by applying new approach methodologies (NAM) that include toxicokinetic (TK) considerations. The central feature of the strategy focuses on using physiologically-based kinetic (PBK)[2] modelling to integrate data generated by *in vitro* and *in silico* methods for absorption, distribution, metabolism and excretion (ADME) in humans for predicting whole-body TK behaviour. In PBK models, the body is represented as interconnected compartments linked via blood flow. PBK models use differential equations to describe the ADME processes that govern the fate of the chemical within the body. These models are able to simulate concentration-time curves, in target organs and in blood. The proper use of PBK models can reduce uncertainties and identify gaps that currently exist in risk assessments that use default extrapolation factors (e.g., 10x for inter-species extrapolation) to estimate human health risks based on *in vivo* animal toxicity studies. PBK models provide the scientific basis on physiology and TK for extrapolations across species, life-stages, routes of exposure, and exposure scenarios. PBK models also provide the means for estimating health risks on both the individual and population levels. In the past, quantitative knowledge of the *in vivo* tissue/blood concentration-time relationship was a prerequisite for calibrating and evaluating the predictive capability of a PBK model. Today, new generations of PBK models are increasingly being developed, using non-animal data, as the field of risk assessment evolves towards the goal of reducing, and eventually replacing, the use of animals for predicting toxicity in humans.

In order to facilitate acceptance and use of this new generation of PBK models in the regulatory domain, experts were invited by the EURL ECVAM to attend a two-day workshop entitled "PHYSIOLOGICALLY-BASED KINETIC MODELLING IN RISK ASSESSMENT – REACHING A WHOLE NEW LEVEL IN REGULATORY DECISION-MAKING" on the 16$^{th}$ -17$^{th}$ of November 2016 at the Joint Research Centre (JRC) in Ispra, Italy. The main objectives of this workshop were to (i) identify current challenges in the application of PBK modelling to support regulatory decision making; (ii) discuss challenges in constructing models with no *in vivo* kinetic and dynamic data and opportunities for estimating parameter values using *in vitro* and *in silico* methods; (iii) present the challenges in assessing credibility of models relying on non-animal data and address strengths, uncertainties and limitations; (iv) establish a *good kinetic modelling practice* workflow to serve as the foundation for guidance on the generation and use of *in vitro* and *in silico* data to construct PBK models designed to support regulatory decision making.

## 1.1 Historical Background

The 2016 JRC workshop was organized with the aim of highlighting the construction, validation, and promotion of PBK models that rely only on non-animal measurements and predictions. This workshop did not reiterate the issues that have been tackled in previous PBK modelling-related workshops. The prior workshops (Table 1), as well as the available guidance documents, dedicated to PBK modelling in the last two decades are briefly summarized below.

The US Environmental Protection Agency (EPA), in 2006, published the document entitled "Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment"[3], which serves as a learning tool for scientists and risk assessors, and addresses the application and evaluation of PBK models for risk assessment purposes. In addition, this document can be informative to

[1] http://publications.jrc.ec.europa.eu/repository/bitstream/JRC96418/eurl%20ecvam%20toxicokinetics%20strategy.pdf

[2] PBK: is synonyms of PBPK, PBBK, PBTK and is used in this report to define physiologically based kinetic models.

[3] https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=157668&CFID=68657522&CFTOKEN=85518773.

PBK model developers because it provides an overview of the types of data and models that EPA requires for consideration of a model for use in risk assessment. The EPA document aimed to address the following three questions: Why are risk assessors interested in using PBK models? How are PBK models evaluated for use in a risk assessment? What are the questions or data gaps in a risk assessment that can be addressed by PBK models?

The World Health Organization (WHO), in 2010, published a guidance document on "Principles of Characterizing and Applying PBK Models in Risk Assessment"[4] to promote best practice in PBK modelling, including transparency, to facilitate understanding and sharing of these models in risk assessment reports. In addition, Meek et al., (2013) reported several case studies illustrating the approach established by the WHO International Programme of Chemical Safety (IPCS) on characterization and application of PBK models in risk assessment.

The European Food Safety Authority published, in 2014, a scientific opinion on good modelling practice in the context of mechanistic effect models for risk assessment of plant protection products[5]. The opinion identified several critical steps for using environmental models in risk assessment, such as problem formulation, model domain of applicability, selection of environmental scenario for pesticides, toxicokinetic characteristics, and species selection.

The European Committee for Standardization (CEN) (2015) organized a workshop on "Standard documentation of large chemical exposure models" in 2015. The resulting CEN workshop agreement (CWA) was expected to facilitate a more rigorous formulation of exposure models description and the understanding by users. The main outcome was a CEN CWA document which establishes terms and definitions for exposure models and their elements, specifies minimum requirements for the amount and type of information to be documented, and proposes a structure for communicating the documentation to different users (Ciffroy et al., 2016).

In the pharmaceutical field, the European Medicines Agency (EMA) has an on-going effort to harmonize the utilization of PBK model platforms in drug submission to the EMA. The EMA published, in July 2016, a "Guideline on the qualification and reporting of PBK modelling and simulation". The US Food and Drug Administration (FDA) also published, in December 2016, guidance on "PBK Analyses-Format and Content - Guidance for industry". Both documents are currently undergoing a public comment period. In addition to these agency documents, Certara has recently summarized recent advances in development and application of PBK models to support regulatory decision making in the pharmaceutical field (Jamei, 2016; Zhuang et al., 2016).

The 2016 JRC workshop had a different theme from these previous workshops and guidance documents, even though several fundamental requirements for PBK modelling identified in those previous efforts are still valid. In the 2016 JRC workshop, requirements for developing, documenting and evaluating in PBK models without the use of *in vivo* pharmacokinetic data were discussed. The outcome of this workshop was a workflow that can support the future development of a Good Kinetic Modelling Practices on the generation and use of *in vitro* and *in silico* data to construct PBK models for supporting regulatory decision-making. Figure 1 describes the steps of a workflow already established in the EPA guidance, as well as in the WHO/IPCS. The workshop outputs will be added to extend this existing workflow. The extended workflow should not only help in the PBK model development/documentation of this new generation of PBK models built with only alternative data, but also aid risk assessors in better understanding/interpreting the PBK model analysis.

[4] http://www.who.int/ipcs/methods/harmonization/areas/pbpk/en/
[5] https://www.efsa.europa.eu/it/efsajournal/pub/3589

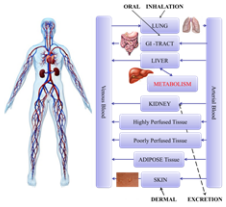
Table 1. Twenty years of PBK models workshops 1995-2016. Abbreviations- European Partnership for Alternative Approaches to Animal Testing (EPAA); National Institute of Environmental Health Sciences (NIEHS); Institut national de l'environnement industriel et des risques (INERIS); Good Modelling Practice (GMP).

| ***Workshop Title Date, Location*** | ***Workshop Host/organizer*** | ***Notes conclusions*** |
|---|---|---|
| Application of Physiologically-based Pharmacokinetic (PBPK) Modelling to Support Dose Selection. 10 March 2014. White Oak Campus, Silver Spring, MD, USA | US FDA | The workshop endeavoured to (i) assess the current state of knowledge in the application of PBK in regulatory decision-making, and (ii) share and discuss best practices in the use of PBK modelling to inform dose selection in specific patient populations (Wagner et al., 2015). |
| ''Potential for further integration of toxicokinetic modelling into the prediction of *in vivo* dose–response curves without animal experiments''. 13-14 October, 2011. Joint Research Centre, Italy | EPAA & EURL ECVAM | The aim of the workshop was to critically appraise PBK modelling software platforms as well as a more detailed state-of-the-art overview of non-animal based PBK parameterisation tools. Such as: 1) Identification of gaps in non-animal test methodology for the assessment of ADME. 2) Addressing user-friendly PBK software tools and free-to-use web applications. 3) Understanding the requirements for wider and increased take up and use of PBK modelling by regulators, risk assessors and toxicologists in general. 4) Tackling the aspect of obtaining *in vivo* human toxicokinetic reference data via micro-dosing following the increased interest by the research community, regulators and politicians (Bessems et al., 2014). |
| The International Workshop on the Development of GMP for PBPK models. 26-28 April 2007, Crete, Greece | The Mediterranean Agronomic Institute of Chaina | Clear descriptions of good practices for (1) model development i.e., research and analysis activities, (2) model characterization i.e., methods to describe how consistent the model is with biology and the strengths and limitations of available models and data, such as sensitivity analyses, (3) model documentation, and (4) model evaluation i.e., independent review that will assist risk assessors in their decisions of whether and how to use the models, and also for model developers to understand expectations of various model purposes e.g., research versus application in risk assessment (Loizou et al in 2008). |
| Uncertainty and Variability in PBPK Models. 31$^{st}$ October - 2$^{nd}$ November 2006, RTP, NC, USA | EPA/NIEHS/CIIT/ INERIS | Better Statistical Models and Methods; Better Databases for physiological properties and their variation; Explore a wide range of Chemical Space; Training, Documentation and Software. The outcome of this workshop has been summarized by Barton et al. (2007). |

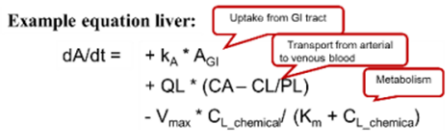
| | | |
|---|---|---|
| Physiologically Based Kinetic (PBK) modelling: Meeting the 3Rs Agendas, October 10-12, 2005, Ispra, Italy | ECVAM | To better define the potential role of PBK modelling as a set of techniques capable of contributing to the reduction, refinement and replacement of the use of laboratory animals in the risk assessment process of potentially toxic chemicals; discuss the need for technical improvements and applications; to identify the need to increase understanding and acceptance by regulatory authorities of the capabilities and limitations of these models in toxicological risk assessment. The recommendations were categorised into i) quality of PBK modelling; ii) availability of reference data and models; and iii) development of testing strategy (Bouvier d'Yvoire et al., 2007). |
| The use of biokinetics and *in vitro* methods in toxicological risk evaluation, 1995, Utrecht, The Netherlands | ECVAM | Reports fifteen (15) recommendations to encourage and guide future work in the PBK model field. 1. Explore possibilities to integrate in vitro data into the models; 2. Models are built on a case by case basis; 3. Establish documentation to illustrate what is needed experimentally; 4. Availability of data required for constructing models; 5. Establish databases; 6. Refine the partition coefficient; 7. Penetration rate should be incorporated into PBK models (barriers information); 8. Biotransformation CYP P450 reactions and information should be included into the model; 9. Emphasis on species comparison (rodent vs human); 10. Target organs and metabolism; 11. In vitro systems should be a reliable representation of in vivo; 12. PBK models should include dynamics; 13.Validation of PBK models should be done with independent data set; 14. Evaluation of the different software are; 15. Sensitivity analysis employed to identify potential source of errors (Blaauboer et al., 1996). |

## 0. Hypothesis

## 1. Definition of conceptual model




## 2. Translation to math equation




## 3. Define parameters

- Physiological and anatomical: tissue volumes, blood flow rates
- Physicochemical: Partition coefficients
- Biochemical uptake constant, metabolic parameters

[Literature, in vivo, in silico prediction QSARs]

## 4. Solving the equation

R packages; Berkeley Madonna; Matlab.

## 5. Evaluation of model performance

- In vivo data
- Human in vivo data
- Sensitivity analysis

## 6. Model predictions

- Applicability (repeated or single exposure)
- Exposure scenario set up

## 7. Model reporting and dissemination

**Figure 1.** Proposed workflow as basis for writing the guidance document (adapted from Rietjens et al., 2011).

## 1.2 Internal Survey

The experts that were invited to the EURL ECVAM PBK Model Workshop were asked beforehand to fill in a brief questionnaire on PBK modelling. This survey was conducted to get a general perspective of the participant's views on the topic (chapter 2).

## 1.3 In vitro to in vivo extrapolation

*In vitro* (within the glass) refers to the technique of performing a given procedure in a controlled environment outside of a living organism. *In vivo* (with in the living) refers in toxicology research to experimentation using a whole. Living organism as opposed to a partial or dead organism. Several works and workshop reports on the in vitro to in vivo extrapolation (IVIVE) approach are available in the literature; available selected reference on this topic are: Blaauboer (2010), Coeke (2013), Wetmore (2013, 2015), Groothuis et al., (2015), Yoon, (2014; 2015), Wilk-Zasadna (2015), Chang (2015), and most recently Bell et al. (2018).

Currently, *In vitro* to *in vivo* extrapolation refers to two different approaches/models:

1. To scale up *in vitro* measured metabolic parameters for use in the PBK model to estimate the metabolic clearance in a real *in vivo* situation (e.g., scaling up intrinsic clearance values determined in microsomes to whole liver);

2. To quantitatively translate a nominal concentration (e.g., a point of departure concentration) used *in vitro* assays to a corresponding exposure dose in vivo (reverse dosimetry using PBK models). This process is referred as quantitative *in vitro* to *in vivo* extrapolation (QIVIVE) and essentially indicates an extrapolation of dose response relationships rom *in vitro* to in vivo. More is discussed in chapter 4.3.2.

## 1.4 Workshop Charge Questions

The workshop was organized into three sessions: (i) Identifying regulatory needs (summarized in Chapter 3); (ii) Constructing a PBK model without the use of *in vivo* data (summarized in Chapter 4); and (iii) Assessing model credibility (summarized in Chapter 5). Each session started with thought-provoking presentations by participating experts. Following the presentations, the discussions took place in smaller break-out groups. Consensus recommendations among the experts are summarized in Chapter 6.

Charge questions that were discussed during the three sessions are summarized below:

- Which types of *in silico* models and high throughput *in vitro* measurements/data can be used to support PBK model development and how can we evaluate them?
- What are the strengths, uncertainties and limitations in using such *in vitro* measurements and *in silico* models?
- How to validate PBK models in the absence of supporting *in vivo* data?
- What are the needs and challenges in building an animal free PBK model?
- How to identify sources of uncertainty in PBK modelling?
- What are the critical needs for a longer term strategy to incrementally refine and deploy PBK modelling in parallel with an appropriate evolution of regulatory practice?

Additional documentation on questions for discussion can be found in pre-conference material shared between the experts (Annex 1).

# 2 Internal Survey on application of PBK models

To gather a general perspective of the participant's views on the application of PBK models, an internal survey was performed prior to the workshop. The experts were invited to fill in, prior to the JRC workshop, a brief questionnaire on PBK models and ADME/TK properties. The questionnaire contained 12 questions of which the first 3 requested information on the use of PBK models. Questions 4 to 8 were directed more on gaining information on computational implementation, model parameterization, and model evaluation. Question 9 was an attempt to identify existing gaps in data for model parameterization and evaluation. Question 10 referred to extrapolation from *in vitro* to *in vivo*. Finally, questions 11 and 12 focussed on regulatory acceptance and good modelling practices. The results of this survey were presented at the workshop, and they were also an integral part of the discussion throughout the workshop.

Out of the 22 participants invited to the workshop, 14 took the survey. The results are summarized below.

## 2.1 Survey Results

**Q1. Do you use PBK/PBPK/PBTK models in your current or past daily work?**

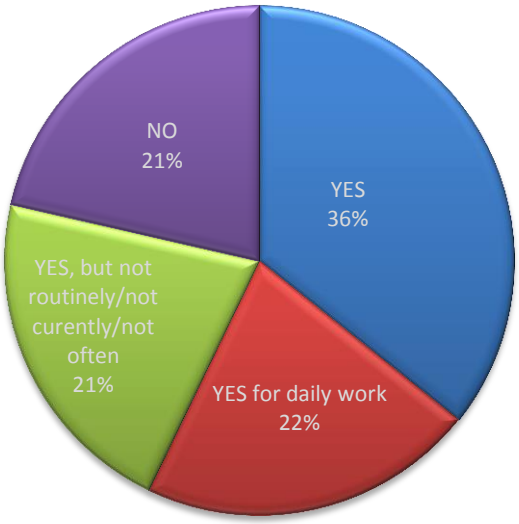



**Figure 2.** Pie chart showing participants results in percentage on utilization of PBK.

Out of the 14 participants who took the survey, 21% replied "NO" to this question, the remaining 79% replied that they have or are currently applying PBK models in daily work. It must be kept in mind that the survey currently was filled out only by experts within the field.

**Q2. For which application(s) do you use PBK modelling (e.g., human or ecological health risk assessment, experimental design)?**

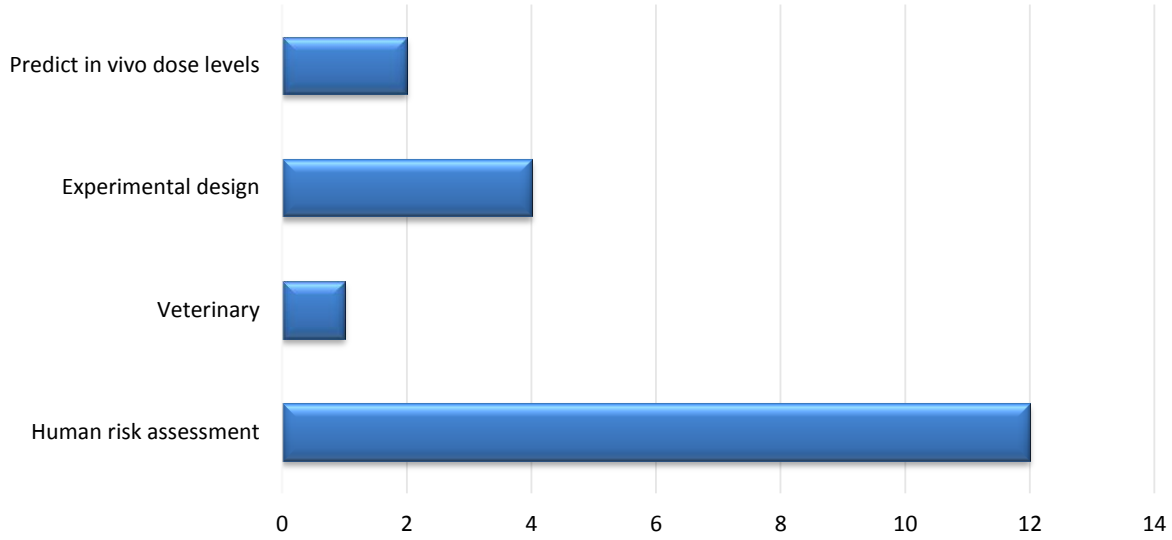



**Figure 3.** Pie chart showing participants replies for which application they had used PBK models.

Figure 3 shows PBK models are mainly used for human risk assessment by the workshop participants, followed by use for experimental design.

**Q3. In which field (e.g., medicine, food safety, REACH)?**

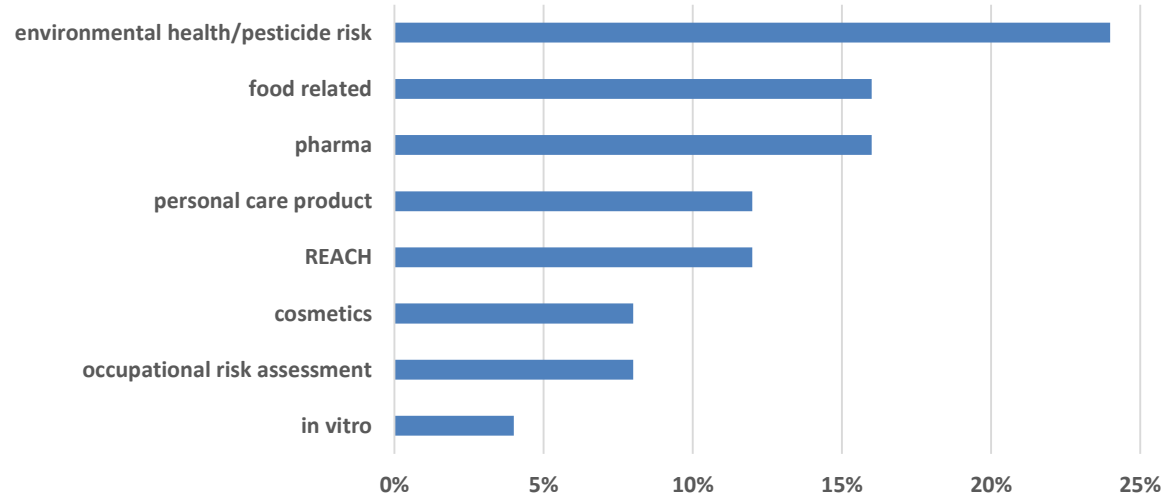



**Figure 4.** Participants responses in percentage on the field in which PBK models are applied.

As it can be seen, experts identified a variety of fields in which PBK models are being used. The fields that received more than 10% of responses include pharma, food related, personal care product, and Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Additionally to the most common fields listed a small percentage responded that PBK models are used in analysis and integration of *in vitro* data and processes.

**Q4. What specific software platform do you use to build PBK models (e.g., simCYP, GastroPlus, MATLAB, R, Berkley Madonna)?**

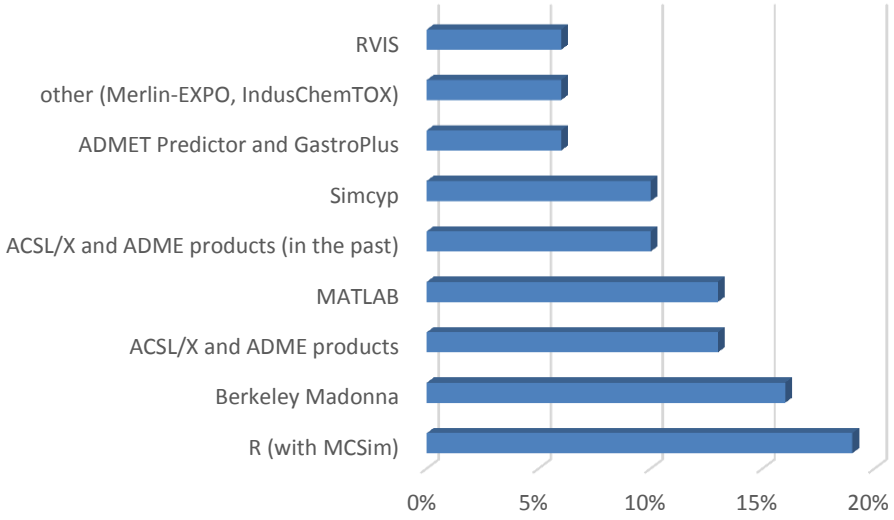



**Figure 5.** Summary of the most used language/program to write PBK models

The software language R (with MCSim) and Berkely Madonna are the most used platforms for PBK modelling among the field experts, followed by ACLSX, MATLAB, SimCyp, Gastroplus, ADMET, Rvis, and others. It must be kept in mind that RVis is currently under further development.

**Q5. Sources of chemical-specific parameters (e.g., partition coefficients, metabolism, skin/oral absorption, protein binding) for the model (e.g., in vitro system, in silico predictions, database)?**

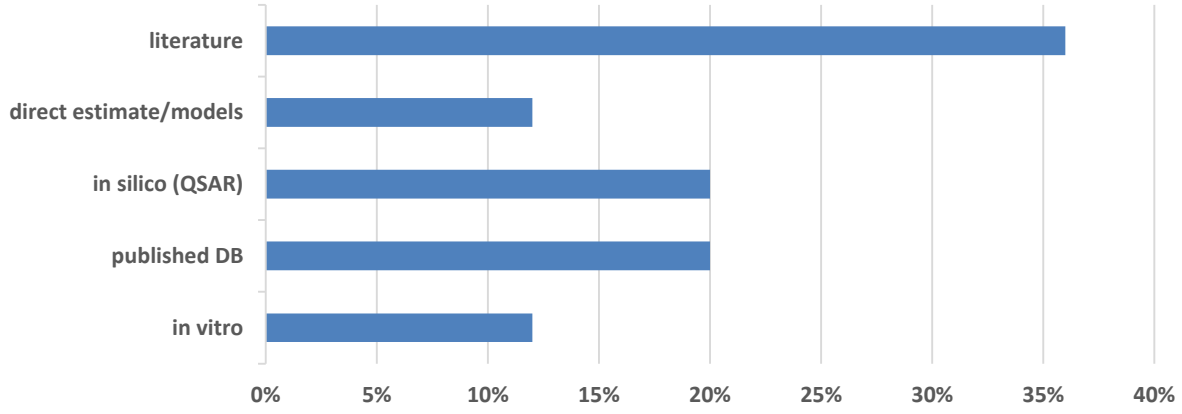



**Figure 6.** Chart showing the distribution of responses in percentage for the question: where do the input parameters come from? (Categories provided: literature, in silico (QSAR prediction), direct model estimates, published databases (DB), in vitro data).

The most used sources for obtaining values for chemical specific parameters are literature, *in silico* such as Quantitative Structure-Activity Relationship (QSAR) predictions and published databases (DB). Ad hoc *in vitro* data is only used by 12 % of the participants.

**Q6. How do you evaluate model performance?**

This question was an open question, but the answers can be summarized by the following types:

- ✓ Model predictions vs experimental data for a set of known compounds;
- ✓ Increased stability by evaluation using range of predicted values;
- ✓ Sensitivity analysis;
- ✓ Independent Expert evaluation.

**Q7. What experimental data for ADME are critical to build PBK models?**

The experimental data needed depends of course on the PBK model built, but in general, the following set of data was reported to be needed:

- ✓ Metabolism ($V_{max}$, $K_m$)
- ✓ Plasma protein binding
- ✓ Chemical absorption, bioavailability
- ✓ Saturable or linear ADME descriptors
- ✓ Blood:tissue partitioning (Log P/log D)
- ✓ Data on renal excretion
- ✓ Physicochemical properties
- ✓ Intrinsic clearance values

It was stressed by one respondent that the experimental data needed depends on the purpose of the study and the chemical of interest. For some purposes or chemicals, only one kinetic parameter may be of importance. Some examples are listed below:

1) What is a safe daily oral intake for humans of a chemical present in food, knowing the dose-response in rats? For modelling such chronic oral exposure, it is usually sufficient to have % absorbed, $V_{max}$ and $K_m$ of disappearance of test substance (hepatic clearance) experimentally, for both rats and humans. Partition coefficients can be estimated very well (QSARs), however, this is only true for certain classes of chemicals. The Poulin et al (2001; 2000) and Schmitt et al (2008) family of models work well for most organic molecules, but do not handle say metals or perfluorinated compounds. For renal excretion, the glomerular filtration rate (GFR) can be used.
2) What is a safe daily oral intake for humans of a chemical present in foodstuffs, knowing the dose-response in rats and seeing the effects found are in the foetuses? In this case, the transfer to the placenta may need to be included.
3) What is a safe daily oral intake for humans of a chemical present in foodstuffs, knowing the dose-response in rats and knowing the substance is metabolized by an enzyme which has very variable activity in the human population? In this case, you might want to include population variability for the metabolism.

**Q8. What experimental data for ADME are critical to support extrapolation for risk assessment?**

Given the diversity of the answers to this question, the provided answers are summarized here in random order:

- ✓ Ideally, measured data for the dose metric that is being used to assess potential adverse and non-adverse effect levels in preclinical tox species and humans. Measured data would be required, but these are often not available.
- ✓ If predictions based on the PBK model are e.g. for the blood concentration, in vivo data on blood levels (present upon oral dosing) are needed to assess how well the model predicts. If specific tissue concentrations are predicted, in vivo data on tissue levels would be required, but these are often not available.
- ✓ Time course and dose response data for model calibration and evaluation. Even if the goal is to reduce in vivo testing, it would be critical to have these in vivo data on some "representative chemicals".
- ✓ It depends on the details of the model and the kind of extrapolation required. In general, anything that differs between the domains (extrapolation from and extrapolation to).
- ✓ Data from systemic toxicity studies regarding potential target organs, identification of metabolites and their potential fate in the body.
- ✓ Metabolic rate constants and intrinsic clearance values, human exposure data, human biological monitoring data and in vitro concentration-response relationships.
- ✓ More often it is absorption rates and metabolic elimination that are key. Distribution can be estimated as a first tier using QSAR or biological prediction models. Urinary excretion can be approached as a first tier using glomerular filtration rate in combination with protein binding. But in all cases, any assumption should be clearly stated in any discussion/conclusion based on using the PBK model at stake.
- ✓ Chemical-specific:
    - Data to verify if fraction of xenobiotic unbound can be extrapolated.
    - Experimental data in the right dose range that is relevant to risk assessment.
    - Good specification of ADME that drives target-site concentrations and data to extrapolate those driving factors.

  Physiological:
    - Ontogeny information/data on transporters and metabolic enzymes (across life-stages and across species)."
- ✓ Data needed for the construction of the PBK model– giving you a conservative estimate of risk.
- ✓ Depends on the extrapolation step(s) needed to bridge the extrapolation from the hazard Point of Departure (PoD) to the human exposure scenario under consideration. If predominantly based on *in vitro* information, Quantitative IVIVE (QIVIVE) will have to be performed which requires a relatively complex model with many input parameters.

**Q9. What ADME property should be addressed experimentally, that we are currently missing for both modelling and risk assessment (e.g. generate more data for membranes or bioavailability)? What is your priority?**

The following ADME property as answers were given (in brackets are the numbers of respondent selections per property): Metabolism (6), protein binding (3), renal excretion (3), transport mediated uptake/efflux (3), absorption skin, gastrointestinal (GI) and respiratory tract (2), intrinsic clearance (2), and foetal disposition, placenta barrier (1).

**Q10. How do you apply in vitro to in vivo extrapolation?**

- ✓ Graphically link the *in vivo* to *in vitro* dose using the blood/plasma C,t-curve based
- ✓ In vitro ADME measurements (e.g., *in vitro* metabolism rates, skin absorption rates) being scaled up to *in vivo* ADME rates.

- ✓ *In vitro* toxicity assays being linked to *in vivo* effects using AOPs. For example, binding of thyroid peroxidase enzyme with chemicals in vitro may be linked to T4 inhibition (a key event) and subsequent thyroid function disruption.
- ✓ In vitro dose response data being linked to in vivo blood or target tissue concentrations using PBK models.

**Q11. Do you have any challenges or experience to share in gaining regulatory acceptance of PBK models?**

Seven participants answered "No", and seven answered "Yes", with more specific answers were given in the following list:

- ✓ Lack of understanding and expertise
- ✓ Discussion on correct model type to use, model complexity
- ✓ Lack of libraries containing parameters to use within user friendly models
- ✓ Need for higher level of confidence

**Q12. Define "GOOD MODELING PRACTICE" (in 3 sentences)**

The following answers were provided by respondents:

1. Good modelling practice would include a transparent, clear and explicit documentation of the model structure, equations and model assumptions. In addition, the model structure and parameter inputs need to be biologically plausible and the performance of the model to describe supporting data, in terms of fit to measured data, and an assessment of sensitivity and uncertainty in dose metric predictions that are relevant to toxicity and risk assessment should be evaluated.

2. Clear description on the goal of the model.

3. Description of assumptions and requirements of the model, determination of a realistic human exposure scenario (including time duration, frequency, level).

4. Make/use a model that is as simple as possible for your question (minimal number of compartments)

5. Model concept provided (scheme with blocks and arrows)

6. Model representation by mathematical equations adequately describes the biological system, as well as the structure/properties of the chemical.

7. Clear indication on the derivation of the model parameters, including the uncertainty around each of the parameters.

8. Model implementation in any software language accurately reflects the mathematical equations.

9. Model performance is evaluated using existing PK data or the model of a similar chemical has been evaluated with existing PK data.

10. The final model needs to be transparent and replicable.

11. Finally a global sensitivity analysis of model structure (Van Hoey et al., 2014), including their distribution in a probabilistic way, should be performed.

## 2.2 Survey discussion and follow up

From this internal survey, several key findings were identified among the experts, which represented however only a small portion of the scientists and risk assessors involved in ADME/TK and PBK modelling and applications. Nevertheless, one of the findings is that the challenges in gaining regulatory acceptance of PBK models are a lack of understanding in model construction, interpretation of model results, and application of model results in a risk assessment context. This challenge is partly contributed by the lack of expertise in the use of these models among end users (such as regulators). Without the necessary expertise, it is difficult for end users to properly review or apply these complex models. Another challenge is the lack of libraries and databases containing parameter values that have been thoroughly vetted to use by the PBK modelling community. Also, most PBK models are not coded in user friendly platforms, and thus, non-programmers may not be able to review model code. All of these challenges stressed the need for better communication between modellers and end users.

Another important finding is that the 14 workshop participants prioritized ADME properties that should be addressed experimentally (Q9) in the following order: Metabolism > > absorption skin, GI and respiratory tract, renal excretion, transport mediated uptake/efflux, protein binding > intrinsic clearance > foetal disposition, placenta barrier. Also, the participants' responses to the question of which experimental data are critical when constructing a PBK model (Q7) were similar to the list reported by Zhuang et al., 2016: MW, log P, pKa, basic, acidic, neutral, pH dependent solubility, PPB, blood:plasma ratio; apparent permeability (Caco 2 or MDCK), intrinsic clearance, protein concentration in *in vitro* test, *in vitro* test matrix binding, Vmax, Km, % enzyme contribution to metabolism, reversible inhibition, CYP inhibition / induction.

The basis for Good Modelling Practice includes a clear and explicit documentation of the model structure, equations and model assumptions, according to the 14 workshop participants. In addition, the model structure and parameter inputs need to be physiologically plausible and the performance of the model needs to describe supporting data. Also, an assessment of sensitivity and uncertainty in dose metric predictions that are relevant to toxicity and risk assessment should be conducted. Model documentation should clearly state all the assumptions underpinning the model and justify the choices of model structure and values of parameters. Furthermore, model code should be carefully annotated.

After the workshop, a new survey was compiled and released on the 11$^{th}$ of January 2017 via the EU survey tool (EURL ECVAM PBK model survey 2017) to understand the use and application of PBK models in broader scientific and regulatory communities. The results by country are available at http://bolegweb.geof.unizg.hr/questionaire/pbk/, (When accessing the link, double click on the selected country to retrieve results by country). The results from this international survey are summarized and analysed in Paini et al., (2017).

# 3 Identify regulatory needs

## 3.1 Identify Regulatory Needs - summary of presentations

**Title:** Evolving PBPK applications in regulatory risk assessment: current situations and future goals

**Dr. Cecilia Tan (EPA)** opened the session with an introduction on applying PBK modelling approach to support regulatory risk assessment. As stated in the 2006 U.S. Environmental Protection Agency's (EPA) *Approaches for the application of physiologically based pharmacokinetic (PBPK) models and supporting data in risk assessment*, "PBK model analysis is accepted as a scientifically sound approach to estimate the internal dose of a chemical at a target site and as a mean to evaluate and describe the uncertainty in risk assessment". PBK modelling has the capability to predict internal dose metrics under new and inaccessible conditions, and thus, it has been used to support extrapolations from high to low doses, across species and life stages, to different exposure scenarios (e.g., route, frequency, and duration). PBK modelling can also quantify uncertainty and variability in physiology and pharmacokinetic properties, and examine their impacts on the overall uncertainty and variability on the predicted dose metric. PBK models that have been used to support regulatory decision making at the EPA are those that have been calibrated and evaluated with *in vivo* data (e.g., time course of blood or tissue concentration). The two main challenges, however, faced by the risk assessors are difficulties in identifying: (1) independent peer reviewers with knowledge in various disciplines (e.g., pharmacokinetic concepts, risk assessment application, mathematics, statistics, physiology, chemistry, biochemistry, computer programming) to properly review a PBK model; (2) a user-friendly platform for submitting PBK models to reviewers who are not programmers. As the paradigm shifts to developing PBK models using non-animal data, new challenges arise for regulatory agencies to evaluate the predictive capability and proper applications of these models. For example, the construction of these models can no longer include parameters that are required to empirically fit to *in vivo* data, and thus more thoughts are required on key parameters to be included in the model. Among the various model parameters, the most challenging one is likely to be the identification of metabolites that are likely to be generated *in vivo*. Case studies with chemicals that have existing *in vivo* data are valuable in examining some of these new *in vitro* and *in silico* approaches for estimating key parameters to appropriately capture the pharmacokinetic profiles. Also, existing PBK models in the literature can potentially provide insights on unique pharmacokinetic properties for chemicals with specific structure/properties (e.g., perfluorinated chemicals).

**Title:** Global Tools Connecting Exposure, Toxicokinetics and Toxicity in Food Safety

Global tools connecting exposure, toxicokinetics and toxicity in food safety: the contribution of the European Food Safety Authority(EFSA) was presented by **Dr. Jean-Lou Dorne (EFSA).** Chemical risk assessment in the food safety area involves the classic steps bringing hazard and exposure together for risk characterisation. In the food safety area, sound hazard identification and hazard characterisation requires an understanding of both toxicokinetic (TK) and toxicodynamic (TD) processes for compounds entering the human body via the oral route. This enables the translation of external dose (exposure) into internal dose TK processes incorporating absorption, distribution, metabolism, excretion of chemicals (ADME) and toxicity for sound dose response modelling. Since its creation in 2002, EFSA has published over 2000 risk assessments for over 4000 substances in the human health, animal health and the ecological areas. Openfoodtox, EFSA's open source database which provides summary

hazard data for individual chemicals, has been designed using the Organization for Economic Cooperation and Development (OECD) harmonised templates (publicly available as of December 2016). In addition, the development of open source TK tools and models to further integrate exposure, TK processes and toxicity in the human health, animal health and ecological areas are outlined. These include PBK models as well as dynamic energy budget models in ecotoxicology. Another challenge is the harmonisation of human and ecological risk assessment of combined exposure to multiple chemicals ("chemical mixtures"). The need to take into account international developments (e.g. OECD, WHO, the EPA, the three non-food committees of the European Commission) is highlighted as critical for the harmonisation of methodologies. Future developments of global risk assessment tools in the food safety area are discussed in the context of mechanistic alternatives to animal testing such as *in silico* and *in vitro* tools and tiered weight of evidence approaches tailored to support risk assessors in a practical way. International cooperation between national and international scientific advisory bodies and academic institutions concludes as the corner stone for the translation of 21st century toxicological research into harmonised methodologies and tools and for the training of the next generation of risk assessors.

**Title:** What does the regulator need?

**Dr. Minne Heringa (RIVM)** gave a perspective from an EU regulator, coming from the Netherlands National Institute for Public Health and the Environment (RIVM). Her presentation was prepared together with Dr. Peter Bos and Dr. Marco Zeilmaker. She showed how RIVM and other groups are working on the development of a new human risk assessment paradigm that enables an adequate prediction of human health risks based on alternatives to animal testing. The various published concepts all contain a tiered approach (e.g., Embry et al., 2014; Bos et al., 2015). Starting with very basic information in a first tier, more and better information is supposed to be collected to refine the risk assessment in every subsequent tier while identifying and focusing on target toxicity endpoints. This is somewhat similar to the Cooper Stage-Gate® model used by many industries for their innovation process of e.g. new chemicals. Starting with just an idea, more information on the market potential and safety is collected in each subsequent stage, with a go – no go decision at the gate following each stage (e.g. Edgett, 2015). These concepts show that in an animal-free risk assessment, performed within industries as required in the EU, PBK models will finally need to fit into different tiers, with increasing requirements to be met by the model in each subsequent tier. In addition, the pieces of information generated in each tier and the required input of the PBK model in the same tier need to match.

The experience at RIVM with the use of PBK models for e.g. food safety issues is that most models in literature are too complex. They cannot be used for the questions faced, because too many input parameters are required, for which data are not available. PBK models should therefore be kept as simple as possible in accordance with the requirements of each tier. In conclusion, it was stressed that PBK models need to be kept as simple as possible with the preconditions that they are fit for the purpose of each tier and, in each tier, need to be compatible with the input data from e.g. *in vitro* tests.

## 3.2 Discussion of regulatory needs

*The experts were divided into three groups of 7 experts each, to address questions regarding identifying regulatory needs. Smaller breakout groups allowed better interaction and discussion, and tackled the issue by different point of views. This paragraph reports the points discussed by the members. Training for regulators (risk managers) and risk assessors was a first key point discussed; followed by model evaluation, guidance, and harmonized terminologies. Groups were then brought together to elaborate and share their discussion outcomes.*

With the current evolution of science and technologies, risk assessors and risk managers should keep up with development of NAMs. The modern methodologies highlighted from this breakout discussion on which regulators would need training include: -omics, TK & TD, organ on a chip, high-throughput screening methods, read across, Adverse Outcome Pathways (AOPs), IVIVE, and Integrated Assessment and Testing Strategies (IATA).

It is not necessary for regulators to have detailed training on all diverse aspects of PBK models; rather, it may be sufficient to provide tailored training focusing on specific needs of each regulatory sector. For example, some risk assessors may need to run the model, so they will need to have the software and expertise to review and run model code. Other risk assessors may rely on a model peer review system to confirm the accuracy of model code. In this case, they may just need to be trained to interpret the data and put those into context. Risk assessors can also put together technical committees that consist of members with a range of expertise to review model code and interpret model results.

In addition to the content of training, the format of training should also be tailored to achieve maximum effectiveness in understanding the use and application of these models. In addition to the traditional classroom setting, other formats could be used, such as webinars, ad hoc short courses, and more refined MSc course. Offering training online could potentially generate a bigger audience. For regulators who have more confidence in *in vivo* data, a way forward would be to make these courses more easily accessible.

While training is essential, another critical need is to establish guidance and GMP on how to apply PBK models for the intended regulatory purposes. The GMP should include clear documentation to report model scope and purposes, details of model development and evaluation, interpretation of the results, and risk assessment applications. In addition, the individual responsible for a specific step in the process should be clearly identified, and thus, end users can easily identify individuals to request targeted training, if needed, on specific topic in the process. Listing individuals who are responsible for each step in the process also increases the transparency. The first thing that needs to be documented is the context in which the model is to be used, since a reliable model may be misused when results of the simulations are applied for the wrong purpose. For example, a first tier screening level model may not be appropriate for supporting the establishment of a regulatory guidance value. Currently, guidance for documenting PBK models is lacking from both modelling and regulatory communities on how to properly report and evaluate PBK models, or interpret and apply the model outputs.

The Scientific Committee for Consumer Safety considers all available scientific data for the safety evaluation of cosmetic substances, including PBK modelling. In the most recent Notes of Guidance for the Testing of Cosmetic Ingredients and their Safety Evaluation (SCCS/1564/15), they define the conditions for the use of PBK models submitted for risk assessment purposes. PBK modelling has already been accepted as a

tool for risk assessment or as supportive information in some of the chemical specific dossiers evaluated by the SCCS. These conditions can be a starting point for the new, general guidance document.

While training and guidance are both essential, their maximum benefits cannot be achieved without frequent dialogue between regulators, modellers and proposers. The frequent dialogue not only allows the proposers to better understand the needs of the regulators, but also allows the regulators to provide feedback along the model development, evaluation, and application processes. For example, regulators already indicated at the workshop they prefer to use the simplest model possible (although some by necessity are complex models), as finding the input data is otherwise impossible. The dialogue can also help regulators identify their needs for specific training, and help proposers understand the criteria for regulatory acceptance. At the same time, there are also some challenges with the public's perception of bias towards this process of reviewing being the primary concern. In this case, the model can be reviewed by independent peer reviewers or technical committees to minimize the concern. Another challenge is that when a model is submitted to multiple regulatory groups, the frequent dialogue between proposers and regulators may become time-consuming. A potential solution to this problem is to set up a harmonised template for model evaluation, while still allowing the template to be flexible for the specific regulation and country. To set up such a template, we may also need a harmonized and defined "ontology".

When a PBK model is developed for supporting regulatory risk assessment, the modeller may want to identify who will use the models (e.g., REACH, pesticides, food and feed) for which regulatory support. An overview of EU Legal framework and relevant regulation and directive can be found in Table 2. Identifying the end users early on can allow for better design of the study. For example, if a read across approach is likely to be applied by the end users, TK data for different chemicals may be important supporting materials that should be included in the submission package. In another example, the "throughput" (i.e. the number of models/chemicals) of the study can be determined when the modellers are aware of the specific regulatory endpoints. Once the end users are identified, the modellers will need to communicate with the users to understand the intended purpose of using PBK models. The required level of confidence in model outcomes would therefore be evaluated based on the relevance and goodness of in vitro data used for model development for the given purpose of model application. As the majority of toxicity and safety decisions are expected to be based on in vitro and cell-based assay results, weight of evidence from human in vivo data may be used to evaluate model performance by regulators. In such cases, regulators may need to consider allowing data to be generated from human trials, e.g., micro-dosing, as appropriate. Finally, case studies are always a good way to convince regulators either by showing how absurd the old way is (using default Uncertainties Factors (UF)) with animal toxicity study data) or how the new way is more science-based. When developing case studies, the modellers should consider sufficient coverage of chemical space.

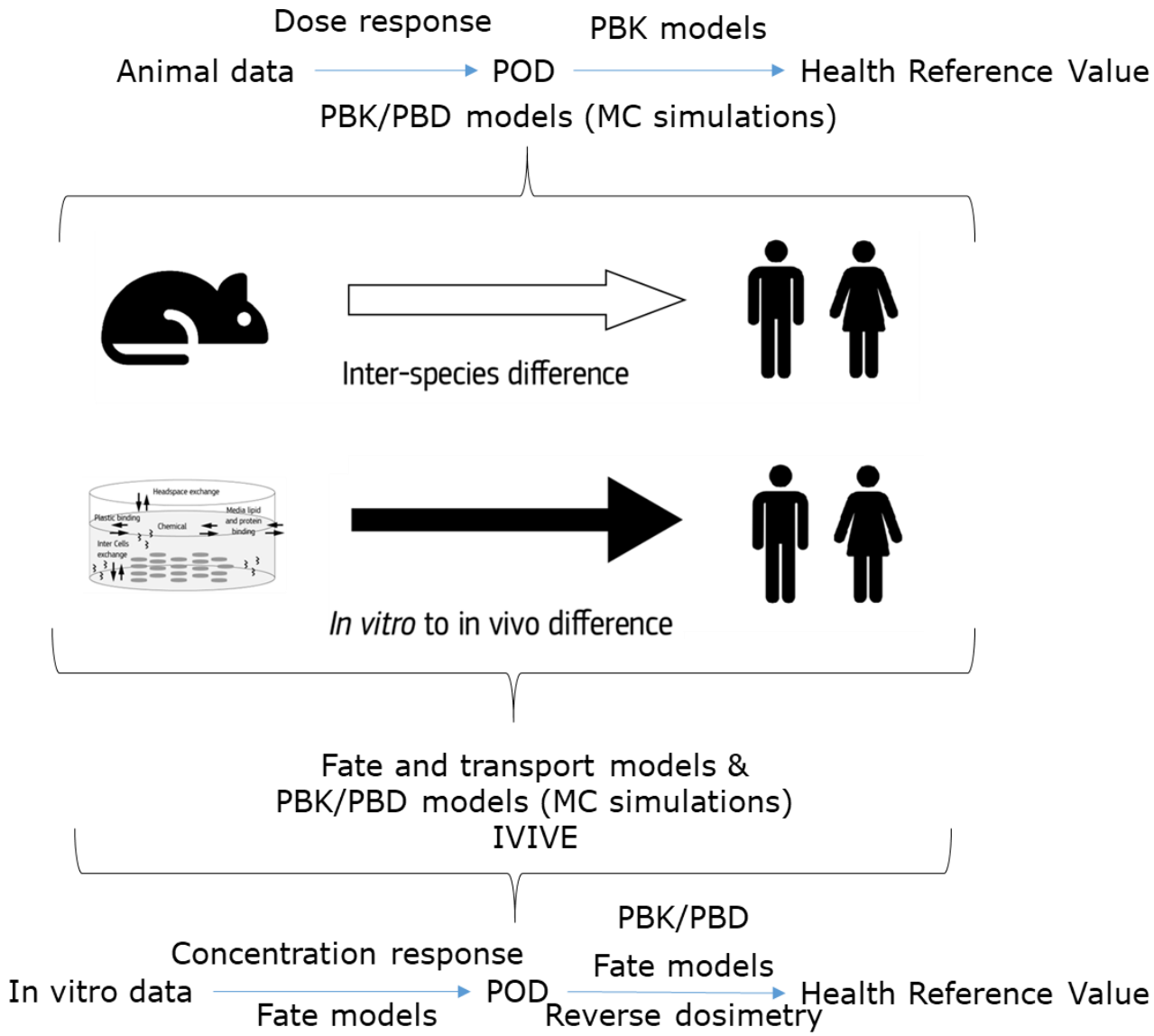



**Figure 7.** PBK/PBD models applied in risk assessment can strengthen the characterization and better define the dose- and species dependent influences on bioactivation, detoxification and possible adverse effects of chemicals thereby providing a basis for more reliable extrapolation from in vitro to in vivo or from animal experimental data to the human situation.

In a situation where safety assessment is conducted for a new chemical on the market and no ADME data are available, the following criteria may be used to facilitate regulatory acceptance:

1. A model needs to be transparent, with usable code;
2. Model uncertainty needs to consider biological plausibility, and be clearly described and quantified when possible;
3. Uncertainty in exposure scenarios needs to be characterized, because this uncertainty will propagate to PBK model results;
4. Consider user-friendly platform;
5. A model needs to be fit for purpose and no more complex than that, and all required parameters should be measurable.

In summary the key discussion points raised during the regulatory needs session were:

- Published PBK models are often too complicated for risk assessment practice
- Keep models as simple as possible.
- Models need to be fit for purpose.
- Most regulators are reluctant to deal with mathematical models/codes
- Modellers will need to understand the regulatory need or purpose for which the model is to be applied and regulators will need to accept the shift to alternative methods.
- Target the training – use case studies to show the use of models to incorporate TK into risk assessment. The more challenging training would be programming, which requires regulators to have certain technical skills
- Establish criteria to educate and train reviewers for PBK models.
- Dialogue between regulators and developers is also very important!
- Facilitate communication on these tools
- Clear documentation/reporting.
- Different type of users will require different levels of information.
- More efforts should be on analysis and evaluation of results and prediction rather than model evaluation.

## *3.3* Supporting information – Relevant Regulations & Directives

The experts agreed that the following EU regulations were identified that could benefit from the application of PBK models (Table 2).

Table 2. EU Legal framework and relevant regulation and directive.

| | **Legal frameworks** | **Regulation or Directive** |
|---|---|---|
| 1 | REACH | Regulation (EC) 1907/2006 |
| 2 | Biocides | Regulation (EC) 528/2012 |
| 3 | Plant protection products | Regulation (EC) 1107/2009<br>Regulation (EC) 283/2013 and Regulation (EC) 284/2013 |
| 4 | Novel foods | Regulation (EC) 258/97, under revision |
| 5 | Food improvement agents:<br>a. Enzymes<br>b. Additives<br>c. Flavourings | Regulation (EC) 1331/2008<br>Regulation (EC) 1332/2008<br>Regulation (EC) 1333/2008<br>Regulation (EC) 1334/2008 |
| 6 | Food contact materials | Regulation (EC) 1935/2004 |
| 7 | Feed additives | Regulation (EC) 1831/2003<br>Regulation (EC) 767/2009 |
| 8 | Veterinary medicinal products | Directive 2001/82/EC, as amended by Directive 2004/28/EC and Directive 2009/9/EC |
| 9 | Cosmetics | Regulation (EC) 1223/2009 |
| 10 | Detergents | Regulation (EC) 648/2004 |
| 11 | Classification, labelling and packaging (CLP) of substances and mixtures | Regulation (EC) 1272/2008 |
| 12 | Adequate risk management in scenarios of emergency response planning (ERP) and for the purpose of land use planning | LUP; as required *e.g.*, within the SEVESO II Directive 96/82/EC |
| 13 | (Inter)national frameworks for OEL derivation, such as performed by national committees or the EC Scientific Committee on Occupational Exposure Limits . | |

## *3.4* Supporting information – Case studies

Example of case studies capturing successful and ongoing efforts where PBK modelling is considered in a regulatory context.

Perchlorate is both a naturally occurring and a man-made chemical that we are ubiquitously exposed to via food and drinking water. Perchlorate can inhibit the thyroidal uptake of iodide leading to thyroid hormone deficiencies. Thyroid hormones play a crucial role in the neurodevelopment of the foetus and infants. Pregnant women, their foetuses, and infants are particularly sensitive to thyroid perturbations with possible adverse neurodevelopmental effects due to perchlorate exposure. The most recent Federal Register Notice issued on June 6, 2016 reads "EPA has begun the process for developing a National-Primary-Drinking-Water Regulation for Perchlorate. EPA, with contributions from FDA scientists, developed a draft Biologically Based Dose-Response (BBDR) model to determine under what conditions of iodine nutrition and exposure to perchlorate across sensitive life stages low serum free and total thyroxine would result" (Federal Register Notice, 2016). One of the driving works for this collaborative PBK modelling approach with regulatory application was a PBK model developed in Lumen et al. (2013) that allowed for the prediction of perchlorate exposure scenarios accounting for the iodine nutritional status at which a late-gestation pregnant woman and her foetus are at risk for alterations in thyroid hormone levels to a level of concern. This model culminated from several model component developments over the years by experts in the field to study perchlorate kinetics and its dose-response. Lumen et al. (2013) model was developed using only available literature data and applying across species and life-stage extrapolation.

MERLIN-Expo is a library of environmental multimedia fate models that was developed in the frame of the FP7 EU project 4FUN to provide an integrated assessment tool for state-of-the-art exposure assessment for environment, biota and humans, allowing the detection of scientific uncertainties at each step of the exposure process. MERLIN-Expo is composed of fate models dedicated to non-biological receptor media (surface waters, soils, outdoor air), biological media of concern for humans (several cultivated crops, mammals, milk, fish), as well as wildlife biota (primary producers in rivers, invertebrates, fish) and humans by applying PBK modelling. These multimedia models together with PBK models can be linked together to create flexible scenarios relevant for both human and wildlife biota exposure. Standardized documentation for each model and training material were prepared to support an accurate use of the tool by end-users. Furthermore, one of the objectives of the 4FUN project was to increase the confidence in the applicability of the MERLIN-Expo tool through targeted demonstration activities based on complex realistic case studies. In particular, the 4FUN consortium researchers aimed at demonstrating the feasibility of building complex realistic exposure scenarios satisfying the needs of stakeholders, the accuracy of the modelling predictions through a comparison with actual measurements, and how uncertainty margins can improve risk governance. The case studies can be seen as reference cases that provide guidance to future users on how to apply the tool in different situations and how to interpret the results from the assessments with the tool taking into account relevant regulatory frameworks.

Simcyp is one of the generic PBK modelling and simulation software for drug compounds. The availability of the generic PBK modelling tools like Simcyp is one the main reasons behind for the recent rise of PBK modelling in drug development (Jamei et al 2009, Bouzom et al., 2012). Simcyp provides a user-friendly platform to utilize the ADME data that are collected as part of drug development in a way to support both developers and regulators for several purposes, among which the most popular application has been the prediction of drug-drug interactions (Yoshida et al., 2017). In addition, applications for

predicting specific populations such as paediatrics have been increasing (Leong et al., 2012). Advances in in vitro and in silico tools and technologies are behind the success of generic PBK tools like Simcyp in the pharmaceutical field. These models can reproduce the clinical observations and more importantly to simulate, i.e., predict, the untested clinical outcomes, allowing the evaluation of the effects of intrinsic (e.g., organ dysfunction, age, genetics etc.) and extrinsic (e.g., drug-drug interactions) factors, alone or in combinations, on drug exposure. There are several areas that are considered as current challenges in accepting model-informed drug development, which can provide insights into what the acceptance criteria should be for the PBK model-based drug development. Among those criteria, two of them are noteworthy including the adequacy of submitted PBK models is to be based on their intended purposes at different stages of drug development, i.e., determination whether a model is fit-for purpose and the need to identify and transparently communicate the knowledge gaps. Use of generic and user-friendly PBK software like Simcyp has certainly been able to support meeting those challenges.

At RIVM, generic PBK models were not used until recently, and custom-made models were applied for each case. These were kept as simple as possible. One example, with many illustrations of what a risk assessor encounters then, is the kinetic (not PBK) model used for the human risk assessment of orally consumed $TiO_2$ nanoparticles (Heringa et al., 2016). The model was based on the findings of an in vivo rat study where organ and blood levels were determined at different time points after a repeated dose exposure had ended (Geraets et al., 2014). These data showed liver and spleen were the main organs where $TiO_2$ was taken up, there was no elimination from the body, there was some elimination from the liver, but spleen levels only kept rising. A simple model with only a gut, a liver, a spleen, and a "rest" compartment was therefore built, with elimination only from the liver to spleen. The latter had no biological explanation, but was essential to let the model fit the data. Later, it became apparent from the hazard data that the testes and ovaries could be target organs, for which it was subsequently of interest to know the internal concentration. Therefore, an additional compartment was made for the gonads. However, fitting all parameters of the extended model to the data was problematic, as the values for the gonads differed by several orders of magnitude to those for the other organs. Therefore, the kinetic constants for the gonads were determined by fitting the extended model to the data, with the parameters for the other compartments fixed (determined in the earlier fit with the simpler model). This had no consequences for the rest of the model, as the transfer rate to ovaries or testes is negligible compared to those to liver, spleen and rest compartment. By applying this model to perform a risk assessment based on organ concentrations, a higher risk was found than when performing a risk assessment based on orally ingested doses (as is common). An explanation for this difference in this case may be the fact that $TiO_2$ nanoparticles accumulate in the body, which makes extrapolation in time essential. A rat study cannot last longer than two years, while humans live much longer, and can thus accumulate much longer. For accumulating substances, kinetic models are thus essential to extrapolate to the longer exposure durations in humans.

Within recent EFSA evaluations on food-relevant chemicals, PBK models have been used in two evaluations to assess species differences (Punt et al., 2017). Firstly, in case of bisphenol A, interspecies differences were assessed based on in vivo kinetic data from different species in combination with PBK modelling. Results revealed particularly differences between mice and humans, with mice having 14.7-fold lower plasma levels of bisphenol A compared with humans at a similar oral exposure, suggesting a higher sensitivity of humans. This difference was taken into account in setting the TDI (EFSA, 2015a). Secondly, in the case of acrylamide, humans were found to have 1.4-2-fold

lower blood levels of the reactive metabolite glycidamide, suggesting relatively lower sensitivity of humans (EFSA, 2015b). Nonetheless, in the latter case the default safety margin of 10,000 for genotoxic carcinogens (covering a factor 4 for species differences in kinetics) was not reduced based on these data (EFSA, 2015b).

# 4 Constructing a PBK model without in vivo data

## *4.1* Constructing a PBK model without in vivo data - summary of presentations

**Title:** QIVIVE: PBK modelling-based reverse dosimetry of in vitro toxicity data

The second session on construction of PBK models without in vivo data started with a PBK modelling-based translation of in vitro toxicity data to the in vivo situation **by Dr. Jochem Louisse (WUR).** The implementation of in vitro methods in toxicological risk assessment is slow. One possible reason is that in vitro methods provide concentration-response data, whereas dose-response data are required to set a PoD to derive safe exposure levels for chemicals. PBK modelling provides a means to translate in vitro concentration-response data to in vivo dose-response data. These predicted dose-response data may be used to derive a PoD that can be used in the risk assessment of chemicals. A few proof-of-principle studies are available that show that in vivo toxic dose levels can be predicted without using animals by translating in vitro concentration-response data to the in vivo situation with help of PBK modelling. However, the approach needs to be optimized before it can be applied in toxicological risk assessment. For example, more insight in the uncertainties in the predicted toxicity dose levels is needed.

**Title:** In vitro-based parameterization of PBK models

Followed by a presentation on in vitro-based parameterization of PBK models by Dr. **Miyoung Yoon (ScitoVation**). The physiological, mechanistic basis of the PBK models is both their strength (in a sense that it provides the exceptional predictive power) and their weakness (as the development of PBK models can be expensive and time-consuming). Obtaining chemical specific parameters, metabolism parameters in particular, has been the biggest challenge in expanding the use of PBK models to a wide range of chemicals as well as in gaining acceptance by regulatory agencies. Currently, it would be necessary to perform in vitro assays of the dose-response (capacity and affinity) for metabolic clearance. These assays are generally more expensive than the dynamic (toxicity) assays, since they necessarily involve the development of an analytical method for quantifying the concentration of the parent compound and its metabolite(s) in each tissue of interest over time. Quantification of the concentration of compound in the dynamic assays should also be performed or at least estimated as the in vitro kinetics is also critical in accurately determining in vitro dose-response relationship (Groothuis et al. 2015; Teeguarden and Barton 2004). Thanks to the advances of in vitro technologies that have occurred in the past few years for determining chemical metabolism and their variability in humans in vitro, the development of 'generic' or 'ready-to-use' PBK modelling platforms has been possible. These generic platforms will contribute to increasing the application of PBK modelling and eventually their acceptance by regulatory bodies by supporting risk-based decisions in different tier of safety assessment. The complexity of the PBK models would depend on the purpose (Wambaugh et al. 2015) of the intended use of the model along with the compound physico-chemical and biochemical characteristics and consequently deriving the types and complexities of in vitro assays for model development. The validity of the in vitro and in silico-based parameterization strategies has been shown with a number of environmental chemicals (reviewed in Yoon et al., 2012) and have been applied to build generic PBK modelling platforms for chemicals. In fact, there has been a rise in the development of generic PBK models for chemicals in recent years such as Population

Lifecourse Exposure-To-Health-Effects Modeling Suite (PLETHEM[6]) and httk package (Pearce et al., 2017).

**Title:** High Throughput PBTK: Open-Source Data and Tools for Dosimetry and Exposure Reconstruction

High Throughput PBTK: Open-Source Data and Tools for Dosimetry and Exposure Reconstruction **by Dr. R. Woodrow Setzer.** High throughput assays serve an increasingly important role in evaluating chemical safety. The ability to screen thousands of chemicals for bioactivity in hundreds of assays goes a long way towards addressing the problem of the very large number of chemicals in commerce with little or no toxicological information available. Concentrations of test chemicals in bioassays need to be converted to dose levels to be able to relate the bioactivity measure in the assays to potential effects in exposed people. With the inclusion of information about population variability and even crude exposure estimates, chemicals can be roughly classified or ordered in terms of potential concern. High-throughput toxicokinetics uses in vitro estimates of plasma protein binding and metabolic clearance to parameterize simple pharmacokinetic models, and adds in silico predictions of partition coefficients to parameterize more general physiologically-based toxicokinetic models. While the predictions of these models are usually less precise than those of chemical-specific models, we can characterize their uncertainty by comparing their predictions to in vivo data, and, to an extent predict when they will fail. We have improved prediction error by improving some of the computational details about partition coefficients. All of this is encapsulated in the R package httk, publicly available from CRAN. It includes relevant data for over 500 chemicals to run both one-compartment and general PBK models. The current version (1.5) incorporates demographic-specific information on population variability of the US population derived from the National Health and Nutrition Examination Survey (NHANES), and provides tools for using Monte Carlo methods to quantify variability of pharmacokinetic predictions.

## 4.2 Discussion on challenges in constructing models

*The second breakout session was planned as the first one, smaller groups met to address challenges in constructing models with no in vivo data, rather, by applying NAMs, QSARs, and in vitro TK and TD data. The objective is to identify key elements that are required in PBK models that are designed to support regulatory risk assessment.*

When there are no data to inform model structure, a minimum PBK model should comprise of the following organs: liver, slowly and richly perfused tissues. Depending on the exposure route, a compartment representing the skin, intestine or lung should be added to the minimum model. If a compound is highly lipophilic, a fat compartment is required, and it may also be necessary for the model to describe the uptake into the lymphatic system. Finally, depending on the hazard data available more compartments, such as target organs, and biological processes can be added to the PBK model. After the model structure is decided, a decision tree could be useful to guide modellers to construct a PBK model without using *in vivo* data for calibration. Some suggested elements for such a decision tree are listed as following:

[6] http://www.scitovation.com/plethem.html

1. Start with the generic model structure and refine for specific chemicals based on chemical structure, physico-chemical properties, and biological similarities to other data-rich chemicals;
2. Examine the mode of action (MoA) of a chemical analogue *in vitro* and/or *in vivo* to determine the *in vitro* studies that are most predictive of the fate in the organism for the chemical of interest. Applicability domain of chemicals should be well defined but quite difficult to do;
3. Draw a realistic case scenario followed by sensitivity analysis.

There is a high value in developing and using one compartment models parameterized with only protein binding and clearance data (Rotroff et al., 2010; Wetmore et al., 2012, 2013, 2014; Tonnelier et al., 2012). The httk R package contains information for such models for over 400 chemicals (see chapter 4.3.2). While predictions from such models are inherently more uncertain than more chemical-specific models, they may still be fit-for-purpose for risk-assessment related applications (Wambaugh et al. 2015).

Steps for developing PBK models built solely using NAMs, *in vitro*, and *in silico* data are similar to those built based on *in vivo* data (adapted from figure 1): (i) problem formulation and identification of relevant exposure scenarios, including dose range and routes; (ii) construction of mathematical equations; (iii) search for existing models for chemical analogues to identify unique features that need to be considered; (iv) identification of values for model parameters from *in vitro* experiments or *in silico* methods carried out ad hoc, via the literature or in databases; (v) integration with other models if needed (*in vitro* fate and transport models); (vi) code implementation; (vii) model evaluation using local or global sensitivity analysis and model validation (check mass balance) by applying read across or comparison to similar models; (viii) analysis and interpretation.

There is the need to build a framework describing how and when to use the different tiers of PBK models, staying within their applicability domain. As more information is gained and models develop there is an increase confidence in their utilization. The strategy would be to start with an aggregated model then move to more specific models, knowing where the decision points are for further testing. In using a more fundamental model, it is possible that a key pathway or MoA, specific to a given target chemical, will not be taken into consideration. In this respect the model may not be the most conservative in terms of risk assessment. One way to address this would be to build a chemically agnostic resource, for example a database of all known ADME/TK processes. The potential relevance of any of these processes for a given chemical can then be considered prior to determining which model would be the most appropriate to represent the system. As much information as possible concerning the MoA of the chemical should be gained. Important is to capture the knowledge to help in defining the equations that will describe our biological process/es captured by the PBK models.

When using a PBK model to convert an *in vitro* point of departure (PoD) to external dose, it is important to evaluate which *in vitro* concentration should be taken as PoD (e.g., area under the curve or peak concentration), and how it corresponds to an *in vivo* situation. For example, the common practice assumes an *in vitro* PoD to be equivalent to a blood or plasma concentration, but is this assumption always valid (Rotroff, (2010), Tonnelier (2012), and Wetmore (2012, 2013, 2014, 2015)? The role of *in vitro* biokinetic study is crucial to translate a nominal concentration used in *in vitro* systems to the actual level of free concentration the cells experience and produce the effect. We can apply several methodologies to address this such as in vitro fate and transport models recently developed by several research teams (Kramer 2010a, 2010b; Armitage et al., 2014; Zaldivar Comenges et al., 2017). A multimedia model approach can be used, however, the further you go in the tiers, the more certainty the exposure and effect dose estimates should obtain, to be able to rely on the margin of exposure. Additionally, is

worth mentioning the US FDA practice of ranking chemicals based upon Cmax/AC50 which is described in Fallahi-Sichani et al. (2013).

Values of some PBK model parameters can be directly measured (e.g., organ volumes, blood flows), but values of other parameters, such as clearance, are inferred from other studies. As pointed out by the results of the internal survey, metabolism is an important feature to be included in the model, especially when metabolites are the possible toxic moiety. The PBK model can be constructed based on different degrees of data availability, and metabolism could be included in the higher tier models. In addition to metabolism, transporters are another challenging piece to address. It is important that the data are produced according to the new OECD good in vitro method practice (GIVIMP)[7]. The GIVIMP document is meant to serve as a technical guidance on good scientific and quality practices to support the regulatory human safety assessment of chemicals using in vitro methods. Within the literature a vast number of *in silico* predictive models for ADME properties have been published, including models for skin and gastro-intestinal uptake, volume of distribution, tissue partitioning (particularly to brain), plasma protein binding, renal and hepatic clearance. Mostrag-Szlichtyng et al (2010) provide an extensive review of in silico tools (QSAR models and Software) for prediction of such properties which are relevant to PBK model building. Prediction of metabolism (rate, extent, nature of metabolites and potential for inhibition) are of particular importance and software used for predicting various aspects of metabolism has been reviewed by Kirchmair et al (2015). A common criticism of software for predicting metabolites is over-prediction i.e. theoretically possible metabolites are not differentiated from those that occur experimentally. In order to reduce over-prediction within the Meteor Nexus software (Lhasa Ltd, Leeds) Marchant et al (2017) describe a process whereby k-nearest neighbour analysis is combined with expert knowledge of biotransformation to reduce metabolite over-prediction). For example, if metabolism is very slow, it may not be detected in short term assay. Another example, Phase III efflux of metabolites cannot be picked up *in silico*. If the parent compound is metabolised, then a model including elimination pathways is needed. To do this, the first step is to determine which methods of elimination are relevant to the target chemical. For example if the chemical is known to be predominantly excreted unchanged in urine then investigation of metabolism is less relevant. Where a chemical is known to be metabolised or to undergo biliary excretion, predictive models representing these individual components of elimination may be required. In silico and in vitro models have been developed for predicting different processes involved in elimination. These include in silico models for total clearance (Lombardo et al 2014) and metabolism (Pirovano et al 2015) and in vitro models for biliary excretion (Ghibellini et al (2006). However, more work is required in developing models for elimination and the applicability domain for existing models needs to be carefully considered before application to a range of chemicals. As more information becomes available and models are further refined they can be used with increasing confidence. One strategy would be to start with an aggregated model then move to more specific models, knowing where the decision points are for further testing.

PBK models can be applied to relate the external exposure dose with internal concentrations that reached the organs can exert a dynamic effect at a cellular level. Integration of kinetics and dynamics information into a PBK/PBD model will help to determine better dose – concentration - time – response relationships (and dependence profile of cellular response) of the delivered dose.

How can we trust a PBK model prediction if there are no in vivo data to evaluate the simulation (more in this topic can be found in chapter 5)? As a solution to this point, a

[7] http://www.oecd.org/env/ehs/testing/OECD_Draft_GIVIMP_in_Human_Safety_Assessment.pdf

read across approach could be used. For instance, referring back to those cases for which you do have data. i.e. use input parameters for "similar" known compounds for a PK read across. Re-parameterise existing model with inputs, which can be obtained, for the target chemical; maybe supplement any known *in vivo* data for surrogate compounds with *in vitro* studies for target. Use a model based on similarity, such as biological, behavioural similarity, and the influence of each parameter (e.g. log P, specific functional group known) should be checked by sensitivity analysis. Additionally, if the individual ADME properties could be predicted these data can be read across, since is not possible to read-across a Ct or an AUC curve. Or the read across approach could be done for similar chemical class for which PBK model already exists. Additionally to read across as a NAM another ascending technology is organ on a chip/human on a chip that can be applied to evaluate and gain trust in these new PBK models.

The Threshold of Toxicological Concern (TTC) approach could be applied in predicting an internal concentration of concern / no concern (Cramer 1978; Munro 1996; Kroes eet al., 2007).

In terms of the most conservative estimation for risk assessment the worst case scenario should be assumed. For example in determining the potential concentration to which internal organs may be exposed absorption from the site of administration can be assumed to be 100% with metabolism being assumed to be 0%. In cases where the chemical is known to form a toxic metabolite the most conservative model would be one where metabolism is assumed to be 100% conversion to the metabolite of concern. In a similar manner, the extraction ratio (i.e. relative amount entering an organ of interest via the blood flow compared to the amount leaving) can be set to 0 or 1 depending on which is more appropriate to give the most conservative estimate for toxicity. For example, if the chemical is potentially toxic to the bladder, calculations assuming an extraction ratio of 1 for kidney may be more protective.

Then the discussion shifted to the need for modelling platforms, such as MEGEN-RVis, PLETHEM, MERLIN-EXPO, which are open source and can be used by individuals with different degrees of knowledge about PBK modelling. These tools allow non-programmers to run the model and learn about the behaviours of the model. However, the biggest concern of these open source modelling platforms is that funding sources for further development and maintenance are not stable. Most of these platforms were initiated by a research grant. But when the project terminates, the developers often cannot find other funding sources to continue the project. One suggestion offered by the experts to address this challenge is to develop specific features for end users for a charge. Such consortium may increase the confidence of the users on these platforms that have specific features designed for their needs.

For model sustainability, it is essential to have access to model equations, as these can be easily coded later. There should be the possibility that when a model is changed the new model is updated and the changes recorded for end-users. The development of an open source library where all models developed could be placed, after a peer review process, was considered important.

Discussion on extrapolation from in vitro to in vivo took place. QIVIVE is an essential process in linking an *in vitro* measured biological (adverse) readout to a potential in vivo outcome as it provides a means to consider exposure and dosimetry and enable the use of *in vitro* data for risk-based evaluations beyond hazard identifications. Multiscale modelling and models describing chemicals fate *in vitro* and *in vivo* contribute to the integrated decision-making process. The challenges faced when applying (Q)IVIVE to risk assessment are, i) the fact that exposure in vitro has different elements than exposure in vivo; ii) the identification when metabolism plays a role in chemical mode of action, e.g., metabolic activation; iii) how to compare in vitro prediction to in vivo. To increase confidence in this end, we must be sure that the in vitro system is not lacking

metabolic competence when relevant, e.g., by ensuring the use of metabolically competent cell models to predict in vivo metabolic clearance. In the long run, prediction of human metabolism in silico needs to be achieved. In the short term, continued improvement of in vitro metabolism assays are recommended to reproduce metabolic rates and metabolite profiles comparable to in vivo. Concerning challenge iii), it is not always straightforward to relate an effect concentration or dose-response curve for an initial event, which may be eventually leading to the disturbance of cellular homeostasis in vitro, to a relevant in vivo exposure situation. Predicting the effect of in-life repeated exposure based on in vitro presents an additional challenge in using in vitro-based PBK models for risk assessment. Solutions to overcome these challenges are i) the evaluation and improvement of in vitro models to address these challenges; ii) identification of the assumptions and uncertainties in using in vitro models for QIVIVE and IVIVE approaches themselves; iii) integration with other approaches such as AOP modelling to appropriately consider modes of action. The most important first step in this direction is to consider what the data requirements should be to ensure the relevance of the given in vitro models for the purpose in risk assessment, to develop appropriate PBK models.

The main discussion points on the construction of a PBK model with no animal data are summarized below:

- Use read across approaches for estimating TK properties.
- Establish UF for IVIVE
- Develop open access modelling platform to facilitate regulatory acceptance, although maintenance of such platforms may be challenging.
- Highlight main fate/processes in organism [create a Knowledge Base of ADME / TK properties].
- Develop a decision tree (Wambaugh et al., 2015) to guide the construction of PBK models using only NAMs, *in vitro* and *in silico* methods.

## 4.3 Supporting information:

### 4.3.1 ADME/TK Databases

In 2008 ECVAM (former EURL ECVAM) commissioned from RIVM (the Netherlands) the development of a pilot database, ECVAM KinParDB (ECVAM Kinetic Parameters DataBase), with kinetic parameters of compounds used as reference substances in various in vitro toxicity tests. Briefly, the kinetic properties of chemicals can provide valuable information in human risk assessment. In vivo as well as in vitro, biological targets are exposed to concentrations of the compounds or their metabolites. Concentrations and their time course, mostly determined in blood or plasma, provide the most direct link between the observed or predicted in vivo effects and the effects observed in vitro. Accurate quantitative knowledge of the in vivo concentration-time relationship is therefore a prerequisite for the correct interpretation of in vitro toxicity test results.

Classical compartmental modelling parameters were chosen to describe the in vivo kinetic properties as they fulfill the needs for prediction of in vivo concentration time profiles under linear conditions. Typical classical compartmental modeling parameters are systemic bioavailability (F), absorption rate constant (ka), volume of distribution (Vd) and elimination rate constant (ke). Protein binding parameters were added to facilitate calculation such as unbound substance concentrations. The database is filled with human and rat kinetic parameters (mainly based on intravenous and oral administration) for 100 substances following assessment of their reliability. Beside an input module (storage template) for the database, a retrieval template was developed to facilitate further use of kinetic data. Additionally a Kinetics Calculation Tool (ECVAM KinCalTool) was developed; this is a self-explaining calculation tool for the construction of a C,t-curve, using a 1- or 2-compartment kinetic model and the kinetic parameters as present in KinParDB. The KinParDB and the KinCalTool are currently publicly available, via the EURL ECVAM website (https://eurl-ecvam.jrc.ec.europa.eu/validation-regulatory-acceptance/toxicokinetics) for use and we invite the scientific and toxicological community to make use of this application.

Another publicly available resource is the online chemical database with modeling environment (oCHEM database) located at https://ochem.eu/home/show.do. This contains 500 ADME and toxicity relevant parameters for a range of chemicals; for some endpoints there are an extensive number of data points for others data are sparser. For example there are > 46, 300 log P values, however for specific tissue:plasma partition coefficients (e.g. heart, kidney, bone etc) there are fewer than ten chemicals. Parameters relevant to PBK model building include volume of distribution (1555 values), pKa (1589 values), Caco2 permeability (462 values), plasma protein binding 3857 values) etc. Note that although the site is moderated and annotations may be made, it is possible for other users to upload information to the system therefore data quality checking is essential prior to using the data for model building. Similarity searching is also possible to identify similar compounds or those containing a given substructure.

Przybylak et al (2017) reviewed 140 datasets of ADME parameters assessing their suitability for modelling purposes based on factors such as availability, size of dataset, format of data and nature of information provided. From this analysis, 31 "benchmark datasets" were identified for a range of ADME parameters such as extent of plasma protein binding, absorption, clearance, bioavailability etc. These datasets, predominantly based on data for drugs, have been made available in Excel format to assist other model developers.

The BRENDA enzyme database (located at www.brenda-enzymes.org) is maintained and developed by the Institute of Biochemistry and Bioinformatics at the Technical University

of Braunschweig. Enzyme function data is extracted from the literature and quality checked by biology or chemistry graduates; it is available free of charge to academic users on-line and as an in-house database for commercial users. Data are available for 83,000 enzymes from 137, 000 references including >135,000 KM values, >38,000 Ki values, >62,445 Kcat and 49,000 IC50 values. Data availability and values vary greatly for individual enzymes; however, this is potentially a useful resource from which to develop predictive models for enzyme activity. Pirovano et al (2015) demonstrated the possibility of using literature data to develop predictive models for $K_m$ and $V_{max}$.

Sources of data for PBK modelling identified either during the workshop or subsequently have been collated in table 3.

Table 3. List of available databases available online that can provide valuable piece of information to build PBK models.

| **Chemical Specific Parameters** | |
|---|---|
| **Name** | **Link** |
| US FDA drug database - drugs@fda | https://www.fda.gov/drugs/informationondrugs/ucm135821.htm |
| DIDB from U-Washington (Seattle, WA, USA) | https://www.druginteractioninfo.org/ |
| Drugbank | https://www.drugbank.ca/ |
| Pharmapendium | https://www.pharmapendium.com/#/login |
| Merck Index | https://www.rsc.org/merck-index |
| GastroPlus ADMET predictor | http://www.simulations-plus.com/software/gastroplus/ http://www.simulations-plus.com/software/admet-property-prediction-qsar/ |
| Simcyp | https://www.certara.com/software/pkpd-modeling-and-simulation/physiologically-based-pharmacokinetic-modeling-and-simulation/ |
| Simcyp free ADME calculator app | https://play.google.com/store/apps/details?id=air.android.com.simcyp.calculators&hl=it |
| PopGen | http://xnet.hsl.gov.uk/popgen/ |
| US EPA iCSS dashboard | https://actor.epa.gov/dashboard/ |
| PubChem Compound | https://pubchem.ncbi.nlm.nih.gov/search/ |
| The Interspecies database – | https://www.interspeciesinfo.com/ |
| ChemSpider | http://www.chemspider.com/ |
| EDETOX database for dermal penetration data | https://apps.ncl.ac.uk/edetox/ |
| US EPAs ECOTOX database | https://cfpub.epa.gov/ecotox/ |
| ToxCast and Tox21 datasets | https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data |
| Httk | https://cran.r-project.org/web/packages/httk/index.html |
| on-line chemical modelling environment - oCHEM | https://ochem.eu/home/show.do |
| KinParDB | https://eurl-ecvam.jrc.ec.europa.eu/validation-regulatory-acceptance/toxicokinetics |
| Brenda | www.brenda-enzymes.org |
| Przybylak | DOI: 10.1080/17425255.2017.1316449 |
| OECD toolbox | http://www.oecd.org/chemicalsafety/risk-assessment/oecd-qsar-toolbox.htm |
| Episuite | https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface |
| **Physiological Parameters** | |
| **Name** | **Link/Reference** |
| Embedded in Simcyp | https://www.certara.com/software/pkpd-modeling-and-simulation/physiologically-based-pharmacokinetic-modeling-and-simulation/ |

| | |
|---|---|
| Implemented in PkSim | http://www.systems-biology.com/products/PK-Sim.html |
| Built in Gastroplus ADMET predictor | http://www.simulations-plus.com/software/gastroplus/<br>http://www.simulations-plus.com/software/admet-property-prediction-qsar/ |
| UK Census | https://www.ukcensusonline.com/ |
| Child growth graphs | https://www.cdc.gov/growthcharts/cdc_charts.htm |
| ICRP | http://www.icrp.org/page.asp?id=145 |
| MEGen | http://megen.useconnect.co.uk/ |
| US EPA Physiological Information Database<br>PID database<br>HERO Database | https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=202847&CFID=90333472&CFTOKEN=83385957<br><br>https://hero.epa.gov/hero/index.cfm |
| RIVM Interspecies database | https://www.interspeciesinfo.com/ |
| P3M | Price et al., (2003) Modeling interindividual variation in physiological factors used in PBPK models of humans. Crit Rev Toxicol 33(5):469-503. |
| NHANES | https://www.cdc.gov/nchs/nhanes/ |
| Brown et al, 1997 | Brown et al., 1997 (Toxicol. Indust. Health 13:407-484) |
| PhysioBank | https://www.physionet.org/physiobank/database/ |
| HESS | http://www.nite.go.jp/en/chem/qsar/hess-e.html |

### 4.3.2 PBK modelling Software

The recent discontinuation of a widely used modelling software product (***acslX***) has highlighted the need for software tool resilience. Maintenance of, and access to, corporate knowledge and legacy work conducted with discontinued commercial software is highly problematic. The availability of a robust, free to use, global community-supported application should offer such resilience and help address the issue of confidence in mathematical modelling approaches required by the regulatory community. RVis, described below, is an attempt at providing such resource. A funding scheme to develop more of these open source softwares should be set up. Industries and academia are encouraged to collaborate and build these computational tools. Below, we categorized three types of computational tools: (i) programming software that has the capability to solve differential equations; (ii) open-source programs developed specifically for PBK modelling; and (iii) commercial PBK models platforms.

1. ***Computer Languages/syntaxes – differential equation solvers***

***Berkeley Madonna*** *is arguably the fastest, most convenient, general purpose differential equation solver available today. It is relatively inexpensive and runs on both Windows and Mac OS. Developed on the Berkeley campus under the sponsorship of NSF and NIH, it is currently used by academic and commercial institutions for constructing mathematical models for research and teaching. (https://www.berkeleymadonna.com/)*

*The* ***MATLAB*** *platform is optimized for solving engineering and scientific problems. The matrix-based MATLAB language is the world's most natural way to express computational mathematics. Built-in graphics make it easy to visualize and gain insights from data. A vast library of prebuilt toolboxes lets you get started right away with algorithms essential to your domain. The desktop environment invites experimentation, exploration, and discovery. These MATLAB tools and capabilities are all rigorously tested and designed to work together. (https://nl.mathworks.com/products/matlab.html)*

***R*** *is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To download R, please choose your preferred CRAN mirror. (https://www.r-project.org/)*

***acslX*** *is a modelling, execution, and analysis environment for continuous dynamic systems and processes.*

### 2. Computer programs developed especially for open source PBK modeling

***MEgen/RVis: MEGen*** *is a model equation generator (EG) linked to a parameter database, RVis* is a prototype application for the analysis of structure and performance of physiologically based pharmacokinetic (PBK), and other models, written in the free, open source syntax R; *and are discussed further in Chapter 5.1. (http://megen.useconnect.co.uk/; Loizou and Hogg, 2011)*

***Merlin Expo****: MERLIN-Expo tool contains a set of models for simulating the fate of chemicals in the main environmental systems and in the human body (http://merlin-expo.eu/; Ciffroy et al 2016).*

***COSMOS KNIME Biokinetic workflows****: The models developed within the COSMOS Project (SEURAT-1) have been implemented into flexible, freely available KNIME workflows, which can further be adapted to users' needs (http://www.cosmostox.eu/what/webtutorials/). (Sala Benito et al., 2017)*

High-Throughput Toxicokinetics ***(httk) is an R package for PBK modelling,*** *https://cran.r-project.org/web/packages/httk. One can use the package to calculate steady state blood levels; it contains a one-compartment and a 4-compartment model. For drugs, httk methods predict within order of magnitude of values measured in clinical trials (Wang, 2010). Measured protein binding, clearance, and calculated K's from Schmitz models were applied and 100% bioavailability was assumed. AUCs predicted iv data, reasonably but for oral data, AUCs were over predicted, which might be due to the assumed 100% bioavailability. A limited correlation was found between the predicted Css and the human in vivo Css ($R^2$ of 0.34) Important factors were found to be: $F_{up}$ (fraction unbound), predicted Css (the higher, the worse the prediction), ionization (pKa_donor), and elimination rate. NHANES population variability in physiological parameters (comparable to POPGen) is part of the httk (Ring et al., 2017) Future planned refinements are: revised Ks, human gestational PBTK, and an inhalation exposure route.*

***PopGen*** *is a simulation program designed to clarify various population genetic events. It is meant mainly for teaching purposes. (http://cc.oulu.fi/~jaspi/popgen/popgen.htm)*

***PLETHEM*** *stands for Population Lifecourse Exposure-To-Health-Effects Model Suite. This computational platform is being developed by ScitoVation under a Memorandum of Understanding with EPA's National Exposure Research Laboratory and National Center for Computational Toxicology (Pense et al., 2017). PLETHEM will provide a freely available, open-source, user-friendly platform for rapid modelling across the source-to-outcome continuum using only in silico and in vitro data. (http://www.scitovation.com/plethem.html)*

### 3. Commercial PBK models platforms

*The* ***Simcyp's*** *Population-based Simulator, includes extensive demographic, physiologic and genomic databases which include algorithms which account for patient variability. This enables the user to predict drug behaviour in virtual patient populations instead of a virtual reference man, allowing individuals at extreme risk to be identified. (https://www.certara.com/software/pbpk-modeling-and-simulation/physiologically-based-pharmacokinetic-modeling-and-simulation/)*

***Gastroplus/ADMET/PBPK PLUS*** *is a mechanistically based simulation software package that simulates intravenous, oral, oral cavity, ocular, inhalation, dermal/subcutaneous, and intramuscular absorption, pharmacokinetics, and pharmacodynamics in humans and animals. This smoothly integrated platform combines a user-friendly interface with powerful science to make faster and more informed project decisions! (http://www.simulations-plus.com/software/gastroplus/)*

*The Computational Systems Biology Software Suite **(PKSim)** contains different software tools and has been designed using a modular concept to allow efficient multi-scale modelling and simulation. The overall platform with its various software tools is implemented in a modular way as will be explained in more detail below. The central software tools are PK-Sim® and MoBi®. While PK-Sim® is based on a whole-body concept, the focus of its counterpart, MoBi®, is at the molecular level. However, both tools extend to additional physiological scales. (http://www.systems-biology.com/products/pk-sim.html). Since 2017 the PKsim has become open-source under GPLv2 and is now available on github https://github.com/Open-Systems-Pharmacology/Suite.*

# 5 Assessing model credibility

## 5.1 Assessing model credibility - summary of presentations

***Title:*** *Challenges in assessing model credibility*

An introductory presentation to the session was given by **Dr. Elisabeth Joossens (JRC)**. Model performance is a key factor when assessing its credibility. A common and good start to do this is by comparing model results against experimental data. If possible, for a whole set of reference data, or at least for some points. As a result, one gets an idea on how well the model managed to reproduce the set of experimental data but credibility can be increased by showing that these results are stable to any variation. So the comparison should be complemented by an analysis of the uncertainty and variability of the model. Variability refers to real differences over time, space, or members of a population and is a property of the system being modelled. It can arise from inherently stochastic processes, but also as a result of explainable and sometimes controllable differences among members of a population (inter-individual variability). While they are important to be reported, variability cannot be reduced. At the same time, every model has several uncertainties covering the lack of knowledge of the true system, the true value of a quantity or real relationships among quantities. They can be classified as scenario uncertainty, model uncertainty, and input (or data) or parameter uncertainty. When possible, they should be reported in a quantitative way but some can only be described in a qualitative way. These reducible uncertainties should be analysed via sensitivity. Sensitivity analysis shows how the uncertainty of the output of a model can be apportioned to the different sources of uncertainty in the model input.

***Title:*** *Challenges of assessing model credibility in the light of uncertainty*

**Dr. John Paul Gosling'**s talk on the "Challenges of assessing model credibility in the light of uncertainty" started with a discussion of the separation of the verification and validation of a model. With the former being an exercise in checking implementation (does the model do what I think it is doing?) and the latter being the more important task of checking that the model is fit for purpose (is model an adequate representation of reality for our purposes?). A simple mathematical framework was proposed that could capture the uncertainties in using the model including both input uncertainties and the gap between model and reality. A brief discussion then followed of how various datasets from *in vitro* and historic *in vivo* sources could be used in such a framework. The conclusions from the talk were that predictive performance is not everything and that transparency and acceptance of the gap to reality are important in gaining acceptance of the modelling approach.

*Title: Uncertain credibility or credible uncertainty?*

The following talk was entitled "Uncertain credibility or credible uncertainty", by **Prof. Eann Patterson**, (University of Liverpool). Model credibility was defined as 'the willingness of others to use model predictions to inform decisions' (Schruben, 1980) and hence is in gift of the decision-maker and not the modeller. A parallel was drawn to the nuclear industry where simulation reviews consist of three steps (Kaizer et al., 2015): determining the level of trustworthiness of the predictions from a simulation, identifying the level of trustworthiness required for the intended purpose and, using these two pieces of information, making a decision about whether to trust the specific simulation for the intended purpose. The role of model validation in providing evidence of trustworthiness was discussed and the accepted definition of model validation employed in engineering was highlighted (Asme, 2006), i.e. 'the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model'. Figure 8 was used to explain the relationship between the level of knowledge about the biology of a system, the extent to which measurements

of the system behaviour are approaches to validation were described, namely: quantitative validation based on codified procedures in engineering that can be used when measurement data is available; rational-empirical validation based on the principles of rationalism and empiricism that can be used when limited or sparse measurement data is available; and epistemic validation, based on simplicity, consistency and explanatory power, for use when there is little or no measurement data available. The presentation concluded with a short discussion of probabilistic approaches to handle approximate knowledge of non-linear dynamic systems and the potential support available from merging validation experience in digital twins.

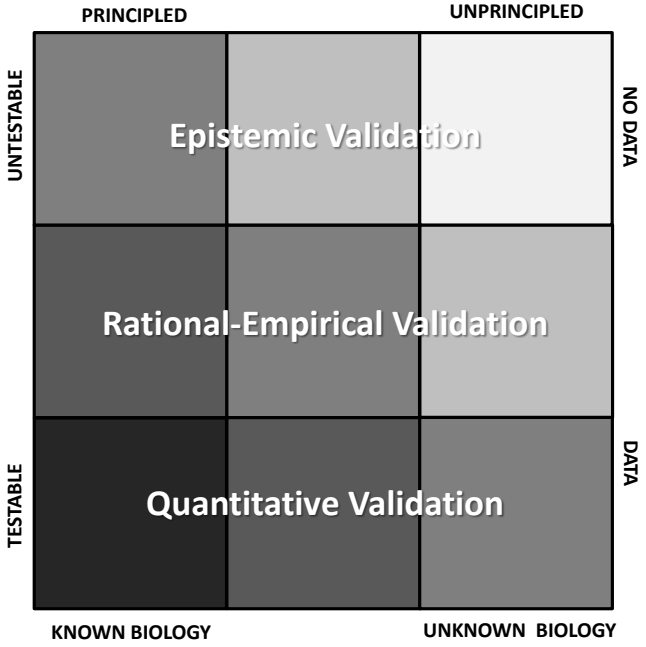



**Figure 8.** Schematic diagram illustrating the relationship between testable and untestable models that are based on known (i.e. principled) or unknown (i.e. unprincipled) biology together with approaches to performing a validation and the likely resultant level of credibility indicated by the greyscale [from Patterson & Whelan, 2017].

**Title:** RVis: A freely available application for the analysis of structure and performance of models written in R

The session concluded with a presentation by **Dr. George Loizou** who gave a brief overview of the development of RVis, a prototype application for the analysis of structure and performance of PBK, and other models, written in the free, open source syntax R.

The widespread adoption and application of PBK modelling in product development and safety assessment has been hampered by criticism that these models are data hungry, resource intensive, complex and require high levels of mathematical expertise and programming skills. Most criticisms can be addressed, as has been demonstrated, with the development of prototype, proof-of-principle, user-friendly web-based tools such as MEGen[8] (Loizou and Hogg, 2011), for the rapid generation of PBK model code, and PopGen[9] (McNally *et al.*, 2014), a virtual human population generator. Both applications shift the emphasis away from the need for high levels of mathematical expertise and

[8] http://megen.useconnect.co.uk/
[9] http://xnet.hsl.gov.uk/Popgen/

programming skills to the understanding of the biology of toxicity and disease that should underpin chemical safety and risk assessment. Further development of such tools would continue to mitigate existing concerns and make this powerful approach more readily accessible to safety toxicologists and risk assessors.

However, the greatest obstacle to the more widespread adoption of PBK modelling is most likely the availability of a common, transparent and independently auditable, free-to-use platform for running models and analysing model structure and output. In response to this need the European Partnership for Alternative Approaches to Animal testing (EPAA) and the Health and Safety Executive (HSE) funded HSL (HSE's Health and Safety Laboratory) to develop a user-friendly in vitro and in vivo exposure predictor. The motivation for this tool is the ultimate replacement of animal testing which requires the ability to predict equivalent human oral, dermal or inhalation exposures that are consistent with measured in vitro target tissue concentrations; an issue which can only be achieved using PBK modelling approaches. The output of this project was RVis, a prototype, proof-of-concept application for the analysis of structure and performance of PBK, and other models, written in the free, open source syntax, R. The first phase of the project was launched in June 2014 and finished in February 2016. In response to the data security concerns of EPAA partners, RVis was designed to be installed on a user's Windows- based PC thereby obviating the need for the uploading of (proprietary) data to a web-based application. In June 2016 the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) convened a meeting of experts to garner feedback from evaluators of RVis from industry and regulatory scientists from the EU, USA and Japan representing the pharmaceuticals, chemicals, cosmetics and agrochemical sectors. Technical improvements needed in the next phase of RVis development were agreed and a second phase of development, funded by CEFIC-LRI began in January 2017.

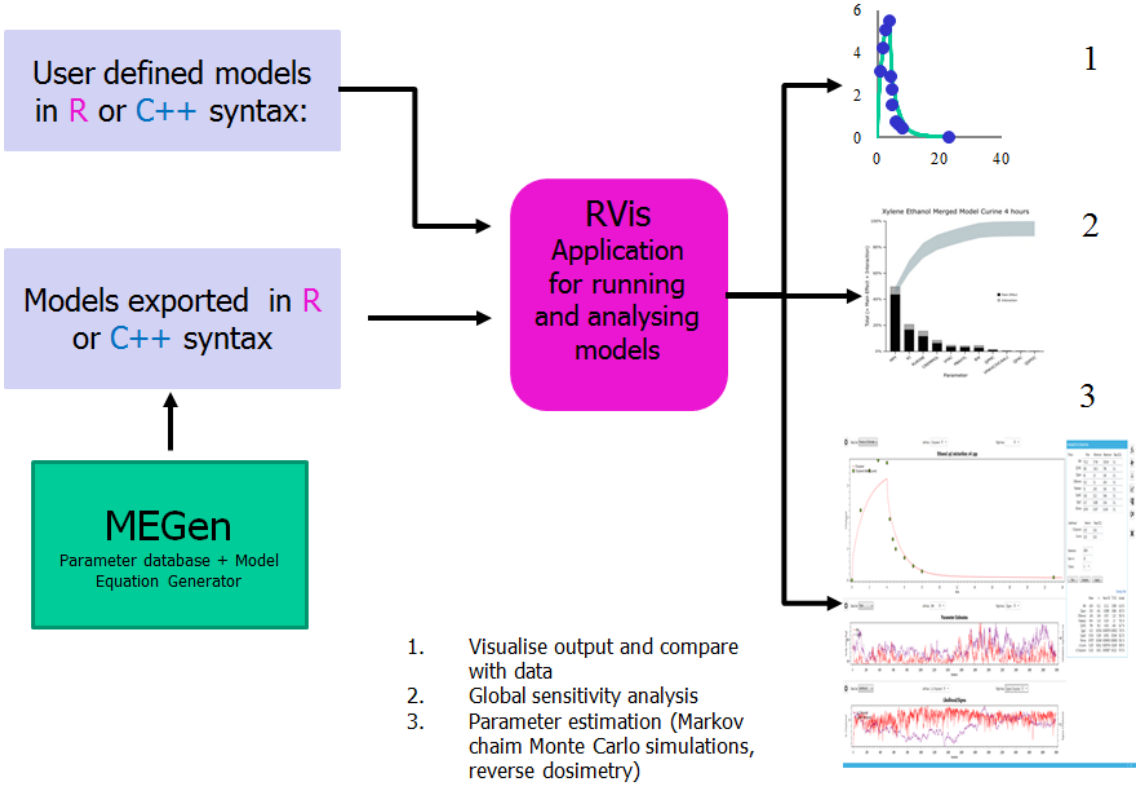



**Figure 9**. RVis: a general purpose modelling platform. Models in R syntax can be run, visualised and graphical output displayed. Model structure may be analysed using parameter elementary

effects screening and global sensitivity analysis (GSA) and parameter estimation using Markov Chain Monte Carlo simulation and Bayesian inference.

RVis is, in fact, a general purpose modelling platform, not just an in vitro and in vivo exposure predictor (Figure 9). RVis features include the ability to load, run, visualise and plot graphical outputs from models. Model structure may be analysed using parameter elementary effects screening (Morris Test) and global sensitivity analysis (GSA) (extended Fourier Transform Sensitivity Test, eFAST) (McNally *et al.*, 2011) and parameter estimation using Markov Chain Monte Carlo simulation and Bayesian inference (McNally *et al.*, 2012). The parameter estimation feature is used to perform "reverse dosimetry" to reconstruct human dose or exposure concentrations consistent with measured biological monitoring data. An approach for the translation of in vitro concentration-response data to in vivo dose-responses will be implemented in the next version of RVis.

**Further development of RVis**

Two examples of improved features were demonstrated: The Lowry plotter and population generator modules (PopGen UI, user interface Figure 12). Figure 10 is typical of the graphical output from a global sensitivity analysis using the eFAST. In this example, model output is venous blood concentrations of m-xylene. The lines on the chart represent the time-dependent changes in the Total and Main effect sensitivity indices for each model parameter. The strength of this method is that it provides a quantitative analysis of model output variance. However, the interpretation of data in this format is difficult. Figure 11 is an example of a more intuitive presentation of the same data allowing easier interpretation. In this format each bar has two sections, the lower dark colour (purple) represents the "Main" effect and the upper lighter section (blue) the "Interactions" that parameter has with other model parameters. Together they represent the "Total" effect. The chart automatically ranks the proportional contribution each parameter makes to output variance from highest to lowest at any given time during the simulation. For example, in Figure 11 the Main effect of parameter "Fu" (fraction unbound) contributes about 26%, the "Interactions" about 16% giving a "Total" effect of around 42% of variance of model output. At any given time point the sum total of the Total effects of all parameters equals 100%.

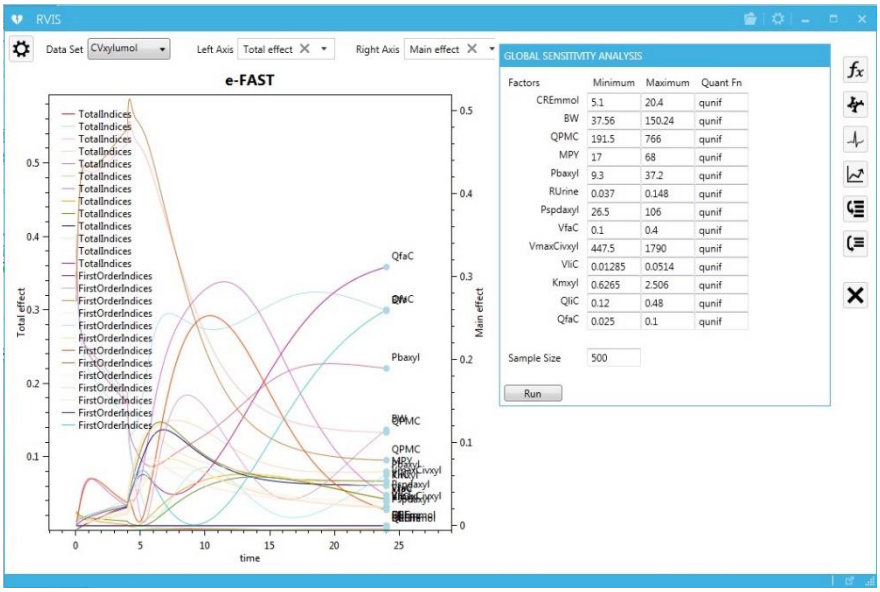



**Figure 10.** Extended Fourier Transform Sensitivity Test (eFAST) global sensitivity analysis output. The lines on the chart represent the time-dependent changes in the Total and Main effect sensitivity indices for each model parameter.

The width of the "ribbon" arising from the parameter with the highest total effect is a measure of the extent and contribution to variance of parameter interactions. The lower bound of the ribbon represents the cumulative total of the Main effects and the upper bound the cumulative total of the Total effects. In reality, due to multiple accounting of interactions associated with each parameter the total eventually exceeds 100%. However, we have imposed a strict limit for the upper bound of the ribbon such that it cannot exceed 100% minus the sum of the Main effects that are not included in the cumulative sum up to that point (top of the ribbon). The user can identify the number of parameters that account for any given proportion of Total variance, e.g., 100% by running a line from the y axis (Total =Main Effect + Interaction) to the ribbon then running a line down to the x-axis. Only those parameters to the left of that line have a significant contribution to Total variance. This information could be used to optimise probabilistic modelling such as Markov Chain Monte Carlo or conventional Monte Carlo sampling by ascribing distributions only to those parameters driving model output variability. This should significantly reduce computational cost.

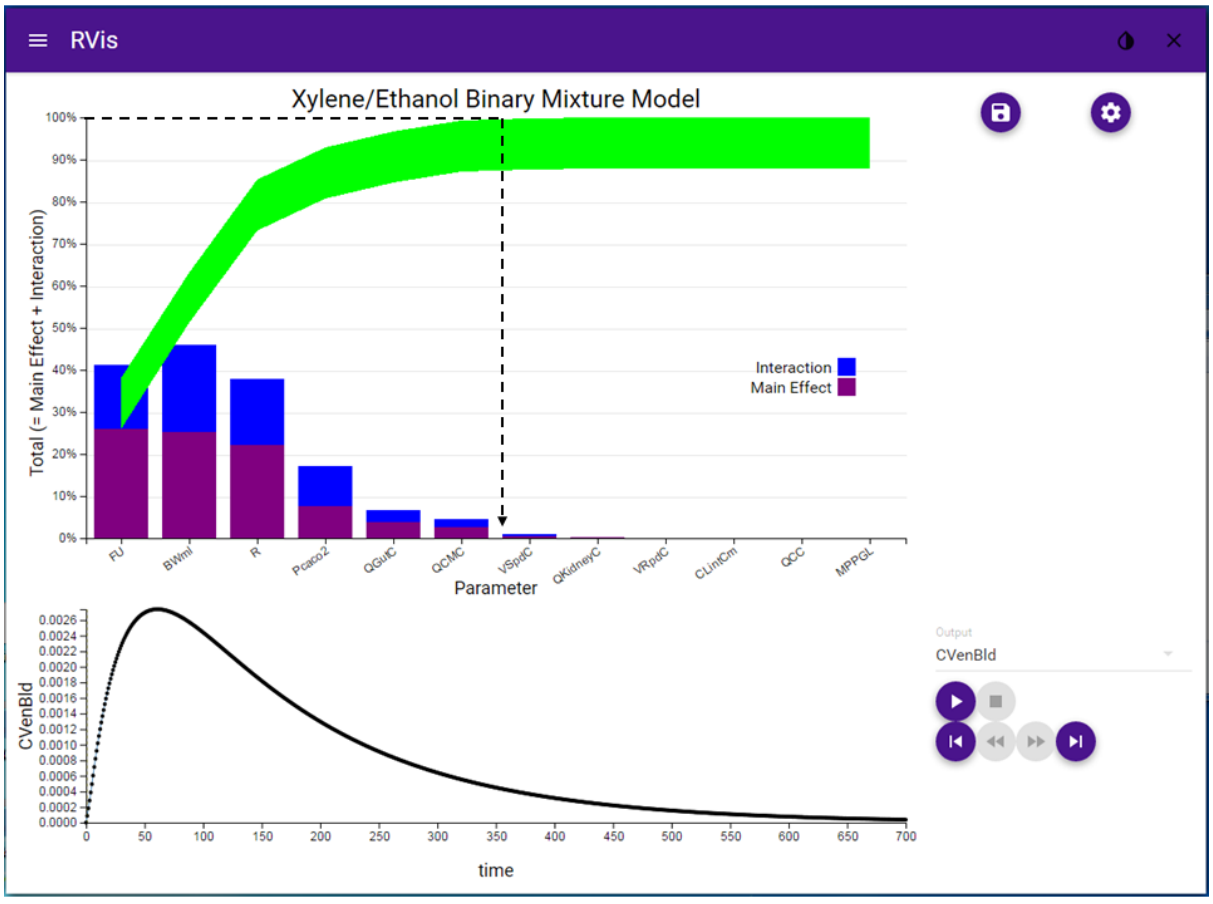



**Figure 11.** The Lowry Plotter: Intuitive interpretation of global sensitivity analysis. The upper panel shows a modified Pareto plot, known as a Lowry Plot, with parameters ranked from left to right according to the magnitude of Total effects at any given time during a simulation. The number of parameters that account for any given proportion of Total variance may be identified as shown with the broken line e.g., 100% by running a line from the y axis (Total (=Main Effect + Interaction) to the ribbon then running a line down to the x-axis. Only those parameters to the left of that line have a significant contribution to Total variance. The lower panel shows the time-dependent model output for which the sensitivity analysis was conducted. Individual Lowry plots can be downloaded for any time point along the simulation by clicking on the line.

The selection of anatomical and physiological parameter distributions required for stochastic and population-based modelling will be made available by incorporating a population generator module into RVis. A prototype, stand-alone working version of this module known as, PopGen UI (PopGen User Interface) will be integrated into RVis. The user will generate virtual healthy human populations via PopGen UI which uses web-services to access PopGen. Figure 12 is an example of the PopGen UI data page. A list recording the date, time and size of user generated virtual human cohorts appears under "My data". A data file is selected by highlighting the file e.g., "18 Oct 10:17 n=10000". A summary of user entries in the PopGen UI homepage such as dataset used, age, BMI and height ranges, ethnicity and parameter units etc., appears on the right hand side. The data can be downloaded as a csv or text file by clicking on the symbols just below the summary of inputs. The user can interrogate the cohort by plotting cohort parameters such as age and body mass against organ masses, blood flow rates and sex. In Figure 11 a plot of bone mass against age for male and females in a mixed cohort of black and Asian people from the UK Health Survey for England[10] dataset is displayed. User defined percentiles for selected parameter distributions e.g., 5$^{th}$ and 95$^{th}$ will be imported into RVis parameter range fields for stochastic modelling.

The further development of RVis would potentially address a number of other areas:

1. Innovating chemical testing. RVis can help reduce chemical testing costs, time and animal use. Standard PBK models rapidly generated using MEGen and exported in R can be exercised and analysed using RVis. The incorporation of in vitro and in silico derived parameters provides the capability to assess potential bioavailability of new chemical entities in people and wildlife. Estimates of bioavailability can be used in tiered exposure assessment and integrated assessment and testing strategies (IATA) which help limit animal numbers and inform the design of specific animal bioassays to define critical dose-response information.
2. Understanding everyday exposures to chemicals. PBK models can be used to predict consumer exposure of new and existing chemicals in commerce e.g., MERLIN-Expo a freely available software platform which integrates a library of environmental multimedia and PBK models (Suciu et al, 2016).
3. Translating research outcomes for product safety. The biological basis of PBK model structure, the estimation of tissue dosimetry and the inclusion of biochemical mechanisms of toxicity provide the basis for data-informed, quantitative chemical safety and risk assessment. Scientifically supported uncertainty factors derived using quantitative, evidence-based models should increase consumer confidence in product safety.

**Possible Regulatory and Policy Impact**

The availability of a resource such as RVis could also have a potentially significant role in three other important areas: the development of internationally recognized good modelling practice (GMP) (Barton *et al.*, 2009; Barton, et al., 2007; Loizou, et al., 2008), rigorous peer-review of PBK models and software resilience.

Regarding GMP, RVis was designed to capture a sensible workflow where a model structure can be quickly and easily analysed using GSA. GSA is the most appropriate form of sensitivity analysis for models that describe non-linear processes such as saturable metabolism and receptor binding (Loizou, et al., 2008; McNally, et al., 2011). The open source, open access, free to use philosophy provides transparency and auditability of model code and performance have been proposed as important elements

[10] http://content.digital.nhs.uk/healthsurveyengland

of GMP. These attributes are also considered to be important features for fostering confidence in mathematical modelling techniques by the regulatory community.

The features that foster GMP could also provide a viable and convenient platform for the peer-review of models. That is, models can easily be exchanged and independently evaluated to provide a more rigorous process for publishing in the peer-reviewed literature.

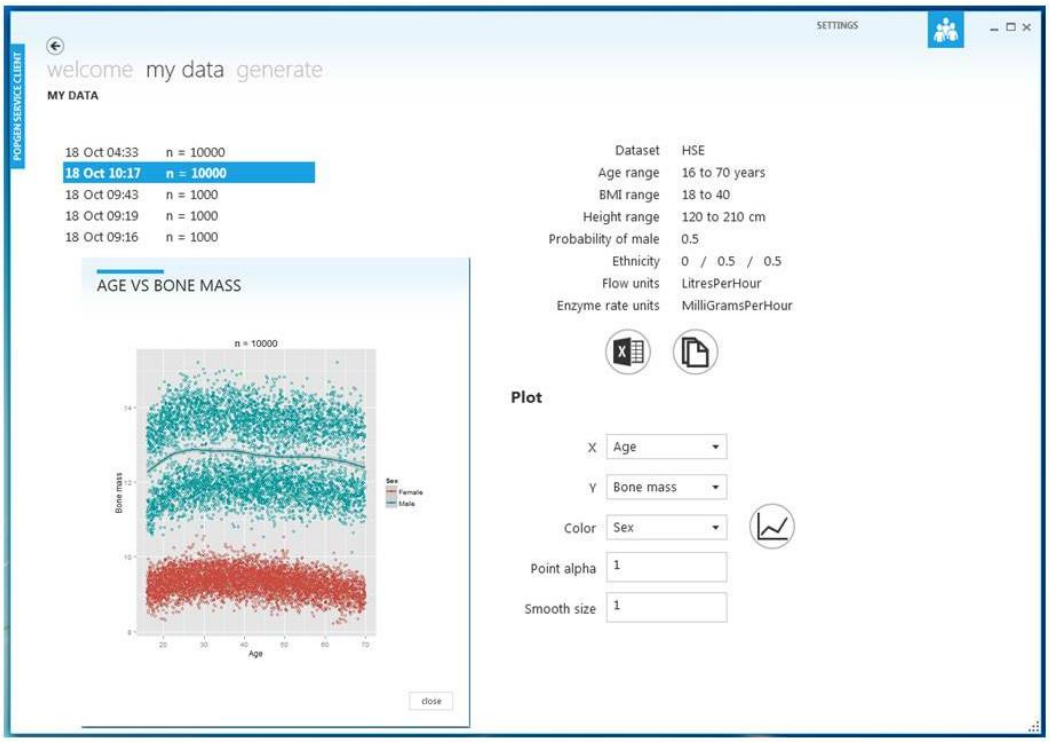



**Figure 12.** PopGen UI. User generated virtual human cohorts appear under "My data". A summary of user inputs are displayed on the right hand side. Data can be downloaded as a csv or text file by clicking on the symbols just below the summary of inputs. Cohort parameters such as age and body mass against organ masses and blood flow rates and sex can be displayed in graphical form. User defined percentiles for any parameter distribution e.g., $5^{th}$ and $95^{th}$ will be imported into RVis parameter range fields for stochastic modelling.

## 5.2 Discussion on model credibility

*The session on assessing model credibility took place on day two of the workshop and participants decided to have the discussion with the entire group rather than smaller group discussion.*

PBK models and other biological models need to be biologically plausible. Sometimes, modellers/mathematicians exclude some biologically important processes because these processes are not mathematically important and models should be kept as simple as possible. However, this must always be done in agreement with the toxicologists, to prevent omission of biologically essential steps. In this case, good documentation on model assumptions is critical for the modellers to demonstrate the credibility of their models to the reviewers and users.

Visualization is a key feature when dealing with communication of these models. A graphical representation of the testability of a model versus knowledge of modelled system (figure 13) should aid in model acceptance and credibility. The graphs (figure 13 A&B) are built on a previous published matrix (Figure 8, Patterson & Whelan, 2017). If a model falls in the bottom left region (testable and with full knowledge), confidence and credibility around the model is likely high. However, if a model falls in the top right region of the matrix (not testable with no knowledge of the system), confidence in the model is likely low due to the uncertainties associated with it. In other words, regulators are unlikely to make decisions with the models in the top right region of the matrix. In some sense, testable models do not really predict, but retrodict. Only when there are no test data, is a model used to make true predictions. In the case of untestable models based on known physics (e.g. Mars Rover), the models need to be simple but with explanatory power in order for them to be consider credible.

The experts identified the need for establishing a framework to examine the credibility of computational models in biology. The framework should lay out the requirements for validating models with different degrees of knowledge and testability (e.g., quantitative validation, rational-empirical validation). In order to pinpoint where we are at the present moment on the matrix, a new matrix was drawn up (figure 13B). This could aid in quantifying the uncertainty we have now with animal models and can help to convince the regulators that models built with in vitro, in silico and NAMs can be just as or maybe even more reliable and trustworthy.

There was a short discussion about the complexity of biological systems. Biological systems are complex and dynamic, resulting in model solutions being patterns in state space rather than single-valued. Thus, similarly a measured value is representative of a particular, and perhaps unknown, starting state of system. Hence, to handle these issues, we need systems thinking and experience-based validation, perhaps involving the merging of all experiences to establish generic digital twins (see Patterson et al., 2016). It was suggested that there is a need to move to 'credible uncertainty' (instead of 'certain credibility') for complex systems.

There is disagreement amongst modellers as to the meaning of the terms, model evaluation, verification, and validation. Regardless of the appropriate term to use, the

purpose is to ensure that the model is predicting what it is designed to predict, and it is a reasonable representation of reality. After confirming that the model is a reasonable representation of reality for our purposes, several analyses may be used to "validate" a model, including sensitivity analysis, robustness analysis, assumption justification, model argumentation, structured calibration, predictive performance, proper scoring rules, and relation to reality. To "verify" a model, one needs to revisit model scope and check model equations and code. Finally, the following key elements were given to achieve model credibility by regulators:

- Understand the model;
- Understand the data underpinning the model;
- State clearly the assumptions and hypothesis encoded;
- Consider the gap between the model and reality, based on available observations.

This gap consideration can be a description of what is lacking in the model. The outcomes of sensitivity analyses can be used to explain some of these deficits. But sensitivity analysis is a characterization of uncertainty in input mostly, so it's dependent on input. In other words, sensitivity analysis cannot be performed on parameters that are missing in the model. One possible approach is to start with a more complex model and then remove parameters which it is not sensitive. The potential problem with this approach is that when there are many parameters with large uncertainties, they may be a flaw the uncertainty analysis. Another possible approach is to build both a simple version and a complex version of the model to examine their differences. A decision tree for model development could help to understand how complex should be the model for the purpose used (e.g. risk assessment). A high tier model would have smaller gap between simulation and reality, and at the same time, the utilization of sensitivity and uncertainty analysis could help to gain trust also with a high tier (highly complex) model.

Additional observations were made by the experts about mixtures and co-exposure. What happens if not only one chemical is the dominant moiety for toxicity? How can interaction between chemicals be considered, such as synergism or inhibition, in the evaluation of the model predictions? The Matrix approach could be extended to address model evaluation for mixtures and co-exposure.

An extensive discussion took place concerning terms validation / verification / evaluation / assessment of fitness-for-purpose / checking validity / evaluating for a given purpose or for a specific application. Is it easier to describe the process by which we "check" models rather than deciding best words to use? EFSA opinion on uncertainty can be a useful resource (EFSA, in preparation) the document is currently in a draft phase, in the process of internal testing at EFSA of its applicability.

In summary for the session on assessing model credibility, the following key points were highlighted:

- Model credibility: biological systems are so complex, we need systems thinking and experience based validation.
- Model Verification vs Validation, proposed a simple mathematical framework.
- Define harmonized terminology. A first step would be to define and agree in terminologies such as robustness analysis, model argumentation, structured calibration and proper scoring rules.

## 5.3 Supporting information: The credibility matrix updated to our needs

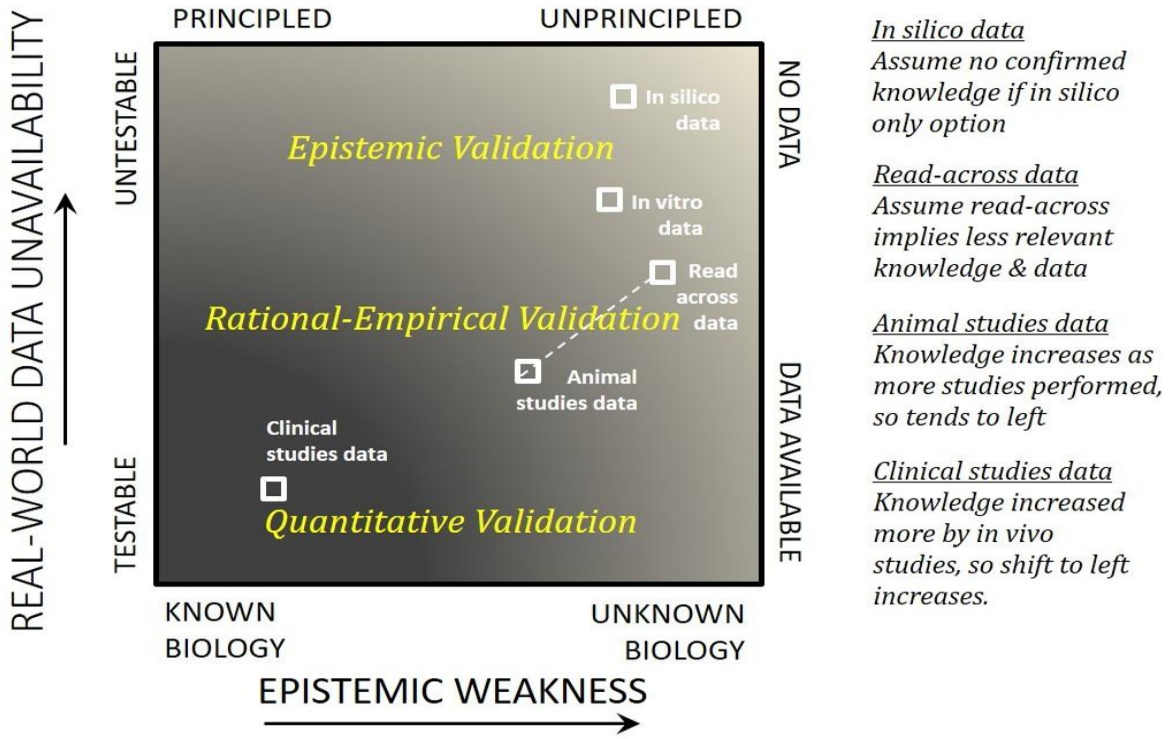



**Figure 13A.** Development of schematic diagram in figure 8 illustrating relationship between availability [or unavailability] of real-world data and epistemic strength [or weakness] of a model and its likely credibility indicated by grey level [darker is more credible] with possible approaches to validation highlight in yellow italics. The tracks show the possible development of a model from purely in silico data, through in vitro data, data from animal studies to clinical studies with increasing knowledge of biology and availability of real-world data leading to better probability of credible predictions.

Recently, Patterson and Whelan (2017) have published a 'credibility' matrix (see figure 8) in which the unavailability of real-world data to support model predictions is plotted as function of the epistemic weakness of the model. The likelihood of establishing credibility in the predictions from a model is represented as a greyscale in the 'credibility' matrix. Two developments of this matrix are shown in figure 13A and 14. Credibility is taken to mean the willingness of others, i.e. not the modeller, to make decisions based on the predictions from the model. It is expected that a model based on known biology whose predictions can be tested by comparison to real-world observations, i.e. in the bottom left corner in figures 13A and 14, will have a high credibility compared to one for which the biology is unknown and no real-world data is available which would be in the top right corner in the matrix. Data might be unavailable because of our inability to control and measure the real-world. Patterson and Whelan proposed strategies for validating or confirming computational models with different level of availability of real-world data; and these are shown superimposed on Figures 13A and 13B, i.e. epistemic, rational-empirical and quantitative validation approaches.

In figure 13A, the path of development of a computational model is shown based on discussions at the workshop. Starting in the top right corner, as an *in silico* model for which the biology is unknown and there is no real-world data. Such a model is largely

heuristic; but, its predictions can be used to design *in vitro* experiments that generate some real-world data, thus allowing the model to be translated downwards in the matrix. Further modelling, utilising the *in vitro* data, should allow the design of useful animal studies that both yield more real-world data and begin to confirm knowledge and understanding; thus, allowing the model to translate further downwards and leftwards. The same position might be achieved by 'read across' data, which by its nature implies initially less relevant data and knowledge. Finally, the predictions from the model, supported by data from animal studies, should enable the design of clinical studies. In turn, the clinical studies yield data that can be used to confirm the predictions, through a process of quantitative validation, which places the model in the bottom left corner with a high probability of stakeholders using its predictions in decision-making. This scenario might be described as the traditional approach to developing *in silico* models of biological processes and is illustrated by the locus in figure 13A. In this figure, the boxes describing likelihood of credibility have been removed to allow a fuzzy classification of the model. The same locus is shown using Greek lettering in figure 13B, i.e. a-α–β-d. In this figure, the boxes describing likelihood of gaining credibility have been retained, which means that the exact position of the model in each box depends on the case being considered, as does its allocation to a particular box when a model is shown on a boundary. Also, shown in figure 13B, is an alternative approach, a-b-c-d, which is not dependent on animal tests. The initial *in silico* model starts from the same position, i.e. the top right, and might consist of a simple model of an observed phenomenon described for instance in a QSAR. This leads to a more sophisticated, but still heuristic, model 'b' based on the understanding gained from model 'a'. The predictions from models 'a' and 'b' are used to design *in vitro* tests that enable the development of model 'c', which can be validated using the rational-empirical approach thus enhancing its credibility. Finally, this leads to the development of clinical studies and model 'd' supported by its predecessors and quantitatively validated or confirmed using clinical data; thus, placing it in the bottom left corner, i.e. a model whose predictions stakeholders, including regulators, practitioners, and patients, will use to make decisions. There are few examples of this alternative approach but it is proposed as the approach to which the modelling community should aspire.

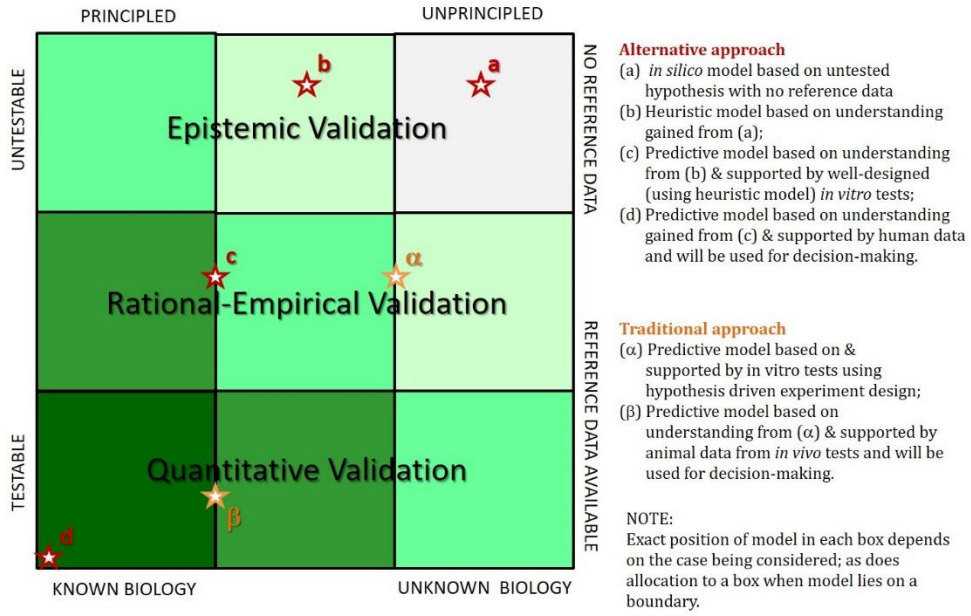



**Figure 13B.** Further development of matrix from figure 8 showing comparative loci for a traditional approach and for an alternative approach. The rationale for the locations of the model types, indicated by stars, are given in the side-bar.

# 6 Conclusions/Recommendations

In order to facilitate acceptance and use of this new generation of PBK models, which do not rely on animal/human *in vivo* data in the regulatory domain, experts were invited by EURL ECVAM to (i) identify current challenges in the application of PBK modelling to support regulatory decision making; (ii) discuss challenges in constructing models with no *in vivo* kinetic and dynamic data and opportunities for estimating parameter values using *in vitro* and *in silico* methods; (iii) present the challenges in assessing model credibility relying on non-animal data and address strengths, uncertainties and limitations in such an approach; (iv) establish a *good kinetic modelling practice* workflow to serve as the foundation for guidance on the generation and use of *in vitro* and *in silico* data to construct PBK models designed to support regulatory decision making. The use of a ***matrix*** to underline and quantify the uncertainty associated with the new generation of PBK models compared to the models developed using animal models would be desirable.

The experts noted that there is a lack of transparent, accessible and easy-to-use software and/or platforms that could easily build and solve PBK models. Such tools would improve the likelihood of adoption of these models within the regulatory community[11]. Development and refinement of existing web applications and PBK model platforms to be able to perform IVIVE and reverse dosimetry in an automated way is needed; with the flexibility to be interoperable with AEPs and AOPs. In addition, there should be an increase in **communication and training** of the regulatory community, such as risk assessors and risk managers. As reported in the EURL ECVAM TK strategy there should be **more data available** (in libraries and in databases)to build both QSAR and PBK models and the in vitro methods to produce these data should be standardised. In vitro methods for which we would need **more standardisation** are (from high to low priority):

1. Liver metabolism (clearance and/or $V_{max}$ and $K_m$)
2. Absorption in lung and intestine (for skin absorption there already available an OECD test guideline)
3. Protein binding
4. Renal excretion
5. Transporters

A recommendation from the experts was to ***develop/refine/adapt good modelling practice,*** as well as ***to generate harmonized terminology/ontologies.*** This working group proposed the drafting of a guidance document for good modelling practice for PBK[12], which could be extended to other in silico biokinetic models. With the increasing demands for alternative methods within the risk assessment framework, the need to develop PBK models has also increased. Existing guidance documents of WHO (2010) and EPA (2006) and the less PBK model-specific documents of EFSA (2014) and CEN (2015) require updating with respect to the current trends, since science and risk assessment are continuously evolving. The challenge is the use of in vitro data or in silico predictions to build these models and integrate and use of them within IATA or AEP/AOP concepts. EFSA and ECHA are currently working on developing guidance on the use of toxicokinetics and metabolism data and bioaccumulation in chemical risk assessment. As an example Schultz et al., (2015) propose a strategy for structuring and reporting read across predictions of toxicity; supplementary information includes a

[11] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3258408/
[12] PBK, PBBP, PBPK, PBTK… are representing the same type of models, Physiologically Based Kinetic models.

document for reporting the information which includes fields for toxicokinetics and metabolism – for assessing similarities between target and source chemicals.

The experts elaborated on the importance of being transparent about the model construct and applications to support assessment including **model credibility**. The experts offered several recommendations to address commonly perceived limitations of parameterizing PBK models using non-animal data, such as the application of the free concentrations. One of the key recommendations identified was the need for a guidance document on building, reporting and documenting PBK models using non-animal data, for interpreting the model analysis for various risk assessment purposes particularly in the regulatory context (e.g., incorporation of PBK models in integrated strategy approaches, integration with in vitro toxicity testing and adverse outcome pathways [AOPs]). The uncertainty and variability in PBK modelling, and fledgling GMP (Loizou et al., 2008) proposed and reported should be further developed and should include guidance for PBK models built using QSARs, in silico data and NAM. The use of a ***matrix*** in a new risk assessment paradigm, to underline and quantify the uncertainty associated with the new generation of PBK models, compared to the models developed using animal models, would be desirable. With the information gain during the workshop we adapted the workflow reported in figure 1, now depict in figure 14.

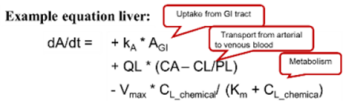
Finally the experts of the workshop recommend the establishment of an international working group for PBK models in addition to an international working group on the selection and standardization of in vitro methods for kinetic parameters necessary for PBK modelling. The first working group should establish criteria for model construction and model evaluation. A group of peer reviewing scientists should be available to put into place the peer reviewing system. US EPA already uses independent scientist to peer review their models. There should be criteria to select the people that will review the models, and provide them with templates and check lists to assist them in this process.

In Table 4 we provide a summary of recommendations, status and solutions/actions taken from the latest workshop held at EURL ECVAM in 2011 and from the 2016 workshop. This was done to make a clear overview of where efforts across the international community are put on and will be with respect to PBK model development and implementation.

## 0. Exposure Scenario

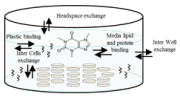
## 1. Definition of conceptual model

## 2. Translation to math equation




## 3. Define parameters

- Physiological and anatomical: tissue volumes, blood flow rates [Literature, databases, in vivo]
- Physicochemical: Partition coefficients [in vitro, databases, in silico prediction QSARs] → in vitro data, to avoid in vitro artefacts, consider the free concentration.
- Biochemical uptake constant, metabolic parameters [in vitro, databases, in silico prediction QSARs] → in vitro data, to avoid in vitro artefacts, consider the free concentration.
  → applicability domain of QSARs



## 4. Solving the equation

R packages (de Solvo, httk); Berkeley Madonna; Matlab, acsIX.

PBK platforms: Megen, Rvis, PopGen, PLATHEM, MERLINEXPO.

Simcyp, PKSim, Gastroplus.

## 5. Evaluation of model performance

- Sensitivity and Uncertainties analysis
  - Matrix to gain model credibility
- Goodness of fit:
  - In vitro data [→ in vitro data, to avoid in vitro artefacts, consider the free concentration].
  - Human in vivo data
- Read across approach

## 6. Model applicability

## 7. Model reporting and dissemination

**Figure 14.** Proposed workflow after discussion and following the recommendation from the workshop.

Table 4 Summary of recommendations, status and solutions/actions from the latest workshop held at EURL ECVAM in 2011 and from the 2016 workshop.

| ***Recommendation (Bessems, 2014) EPAA – EURL ECVAM Workshop 2011.*** | ***Status*** | ***Recommendation EURL ECVAM 2016*** | ***Possible solutions/Actions taken*** |
|---|---|---|---|
| | | **Training & Communication** | CEC at conference (applied for EUROTOX 2018 and SOT 2018); Scitovation PBK model course (November 2017);<br>PBK course for master students, online course through Kansas State University Global Campus (early 2018). Additionally webinars and ad hoc meetings such as the CAAT academy webinar/course (September 2017). |
| ***Set up databases for kinetic data*** | *Ongoing efforts by scientific community (interspecies DB https://www.interspeciesinfo.com/), see table 3.* | **Databases of input parameters from in vitro and in silico data** | Ongoing activities<br>ECVAM KinPar database<br>HESI group on databases (Bier, *contact person Dr. M. Embry*)<br>EFSA TK plate<br>EURL ECVAM databases of in vitro and in vivo biotransformation rates in fish and mammalian species. |
| | | ***Funding** scheme to develop software* | *To be discussed* |
| ***Develop free to use, readily accessible PBK model web applications.*** | *The scientific community is engaged and examples are, among others, Megen/Rvis; COSMOS KNIME biokinetic models; PLETHEM.* | ***Open source of libraries of PBK models (already reviewed)*** | *Make available reviewed code into one place accessible to all.*<br>*Such a repository is an aim in the proposed COST Action Kinetics 2.0.* |
| | | ***Guidance for GMP for new generation of PBK models*** | *Proposal sent at OECD and endorsed by WPHA, EAGMST (interested WNT and QSAR toolbox working group)* |
| | | ***Decision Tree** for new generation model construct* | *To be refined and elaborated and included in OECD PBK guidance* |
| | | ***The matrix approach to gain credibility*** | *Develop case studies – to be done* |

| | | | |
|---|---|---|---|
| ***Permanent international group of PBK model reviewing experts*** | *Was not done until now and was also highlighted by the 2016 PBK model workshop.* | ***TASK FORCE*** *for model peer review* | *HESI PBK working group (March 2017) contact person Dr. M. Embry;*<br>*Results from ECVAM international survey to establish list of experts (March 2017).*<br>*COST Action proposed, which includes work package on PBK models. The group of reviewing experts can be identified in there.* |
| | | ***Scoring*** *system for model peer reviewing and creation of a template.* | *HESI PBK working group (March 2017) contact person Dr. M. Embry.*<br>*Can also be included in the COST Action Kinetics 2.0 (RIVM).* |
| ***Develop in vitro tools for high throughput measurements of portioning and expand the applicability domains of various tools such as absorption methods*** | | ***Standardization of in vitro methods for:***<br>1. Liver metabolism<br>2. Absorption in lung and intestine<br>3. Protein binding<br>*4.* Renal excretion<br>*5.* Transporters | *Workshop in Leiden, NL, on in vitro methods for toxicokinetics in October 2017*<br>*Included in COST Action Kinetics 2.0.* |
| ***Develop high throughput and low cost analytical facilities to measure chemicals in physiological media.*** | *To be done* | | *Not discussed* |

# 7 Next steps/actions

An action of the PBK workshop was a proposal submitted to the EGMAST group within the OECD, in early December, to seek their support to develop further this guidance document (see ANNEX). There was emphasis and need of a guidance to incorporate PK concepts, physiology, in vitro, in silico approaches, NAMs, with programming, mathematics and statistics; to develop guidance on alternatives.

A **decision tree** needs to be refined and elaborated on, for PBPK model development (similar to figure 1 and described in chapter 4.3) without in vivo data, based on physicochemical data, in vitro data, NAMs and in silico methods. For instance, taking into account PBK predicted internal dose metrics vs. in vitro points of departure from toxicity testing (and how in vitro results link to in vivo adverse outcomes) for a tiered assessment.

A **public repository** is needed for already developed and peer reviewed PBK models. This will be important, as once there is a repository for PBK models (for example, considering the AOP wiki format), relevant documentation can be introduced and can include a quality certificate following evaluation from an independent peer review. This is line with the work reported in Lu et al., (2016). Such a repository will allow for the curation of more case studies and in the creation of libraries of ad hoc PBK models that could be used for training purposes, for performing risk assessment, for conducting in vitro to in vivo extrapolation, and importantly to inform decision makers efficiently in the current state of science for the use of animal free models in regulatory applications.

***Establish a Task force for model software*** –

1. *A Linkedin group on PBK model developer could be created*
2. *HESI PBK model working groups.*

*These communities should try to establish: How they peer review model codes.* Which universal language to use; will it be possible to write model codes using one common universal language?

*Build a core **expert group for training** for new modellers and risk assessors taking the example of the AOP training provided by JRC to EFSA, a similar training with focus on PBK model development, evaluation and application should be set in to place. This can be done with direct specific courses or as a continuous education course during conferences, like EUROTOX (2018) and SOT (2018).*

***More communication with regulators:** no in vivo data for model evaluation when in vivo data are not allowed to be collected (take responsibility!)* – Use social media (like Linked-In) to establish a direct dialogue between PBK work groups and regulators (both risk assessors and risk managers) to improve communication and understanding of these models.

*There is a need to **create a community** to address ADME/TK and PBK models issues.* With the needs for several international working groups to further work on these tools, this requires cross-talk to ensure compatibility of the in vitro methods with PBK models in addition to cross-talk with regulators to fit the total risk-assessment framework. A COST Action proposal has now been submitted to form the international network to ensure this cross-talk. In addition, such a COST Network would provide the recommended training, repository of PBK models and standardization of in vitro methods and QSARs. However, funding resources for research and development would still be necessary.

# Reference


Armitage, J.M., Wania, F., & Arnot J.A. (2014). Application of mass balance models and the chemical activity concept to facilitate the use of in vitro toxicity data for risk assessment. *Environ. Sci. Technol*., *48* (16), pp. 9770-9779.

ASME V&V 10-2006, Guide for verification & validation in computational solid mechanics, Am. Soc. of Mech.l Engineers, New York, 2006.

Baldrick, P. (2003). Toxicokinetics in preclinical evaluation. *Drug Discovery Today 8*(3), 127-33.

Barton, H.A., Bessems, J., Bouvier d'Yvoire, M., Buist, H., Clewell III, H., Gundert-Remy, U., et al. (2009). Principles of Characterizing and Applying Physiologically-Based Pharmacokinetic and Toxicokinetic Models in Risk Assessment. IPCS project on the Harmonization of Approaches to the Assessment of Risk from Exposure to Chemicals.

Barton, H.A., Chiu, W.A., Setzer, R.W., Andersen, M.E., Bailer, A.J., Bois, F. Y., et al. (2007). Characterizing uncertainty and variability in physiologically-based pharmacokinetic (PBPK) models: state of the science and needs for research and implementation. *Toxicol Sci, 99*(2), 395-402.

Bell S.M., Chang X, Ph.D., Wambaugh JF, Allen D.G., Bartels M., Brouwer, K.L.R.,. Casey W.M., Choksi N, Ferguson S.S., F*raczkiewicz G, Jarabek A.M.*, Ke A, Ph.D., Lumen A., Lynn, SG., Paini A, Price P.S., Ring C, Simon TW, Sipes NS, Sprankle *C,* Strickland J, Troutman J, Wetmore BA, Kleinstreuer NC., (2018) In vitro to in vivo extrapolation for high throughput prioritization and decision making. Toxicology in vitro. https://doi.org/10.1016/j.tiv.2017.11.016

Bessems, J.G., Loizou, G., Krishnan, K., Clewell, H.J., Bernasconi, C., Bois, F., Coecke, S., Collnot, E.M., Diembeck, W., Farcal, L.R., Geraets, L., Gundert-Remy U, Kramer, N., Küsters, G., Leite, S.B., Pelkonen, O.R., Schröder, K., Testai, E., Wilk-Zasadna, I. and Zaldívar-Comenges, J.M. (2014). PBTK modelling platforms and parameter estimation tools to enable animal-free risk assessment: recommendations from a joint EPAA--EURL ECVAM ADME workshop. Regul Toxicol Pharmacol. ,68(1):119-39.

Blaauboer, B., Bayliss, M.K., Castell, J., Evelo, C.T.A., Frazier, J.M., Groen, K., Gulden, M., Guillouzo, A., Hissink, A.M., Houston, B, Johanson, G., de Jongh, J., Kedderis, G.L., Reinhardt, C.A., van de Sandt, J.J.M., Semino, G. (1996). The use of biokinetics and in vitro methods in toxicological risk evaluation. The report and recommendations of ECVAM Workshop 15. ATLA. 1996;24:473-497.

Blaauboer BJ, (2010) Biokinetic modeling and in vitro-in vivo extrapolations. J Toxicol Environ Health B Crit Rev. 2010 Feb;13(2-4):242-52.

Bouvier d'Yvoire M, Prieto P, Blaauboer BJ, Bois FY, Boobis A, Brochot C, Coecke S, Freidig A, Gundert-Remy U, Hartung T, Jacobs MN, Lavé T, Leahy DE, Lennernäs H, Loizou GD, Meek B, Pease C, Rowland M, Spendiff M, Yang J, Zeilmaker M. (2007) Physiologically-based Kinetic Modelling (PBK Modelling): meeting the 3Rs agenda. The report and recommendations of ECVAM Workshop 63. *Alternative to Laboratory Animals: ATLA, 35*(6):661-671.

Bouzom, F., Ball, K., Perdaems, N., & Walther,B. (2012). Physiologically based pharmacokinetic (PBPK) modeling tools: how to fit with our needs?. *Biopharmaceutics & drug disposition*, *33*(2), 55-71.

Bos, P. M., Gottardo, S., Scott-Fordsmand, J. J., van Tongeren, M., Semenzin, E., Fernandes, T. F., ... & Landsiedel, R. (2015). The MARINA risk assessment strategy: a flexible strategy for efficient information collection and risk

assessment of nanomaterials. International journal of environmental research and public health, 12(12), 15007-15021.

CEN, European committee for standardization (2015) CEN Workshop on Standard documentation of large chemical exposure models (WS MERLIN-EXPO); CWA 16938 Brussels https://www.cen.eu/work/areas/chemical/Pages/WS-MerlinExpo.aspx

Chang Xiaoqing, Kleinstreuer Nicole, Ceger Patricia, Hsieh Jui-Hua, Allen Dave, and Casey Warren (2015) Application of Reverse Dosimetry to Compare In Vitro and In Vivo Estrogen Receptor Activity Applied In Vitro Toxicology, 33 -44.

Ciffroy, P., Alfonso, B., Altenpohl, A., Banjac, Z., Bierkens, J., Brochot, C., Critto, A, De Wilde, T., Fait, G., Fierens, T., Garratt, J., Giubilato, E., Grange, E., Johansson, E., Radomyski, A., Reschwann, K., Suciu, N., Tanaka, T., Tediosi, A., Van Holderbeke, M., and Verdonck, F. (2016). Modelling the exposure to chemicals for risk assessment: a comprehensive library of multimedia and PBPK models for integration, prediction, uncertainty and sensitivity analysis - the MERLIN-Expo tool. *Sci Total Environ*. 568:770-84. doi: 10.1016/j.scitotenv.2016.03.191.

Coecke S, Pelkonen O, Leite SB, Bernauer U, Bessems JG, Bois FY, Gundert-Remy U, Loizou G, Testai E, Zaldívar JM. (2013) Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches. Toxicol In Vitro. 2013 Aug;27(5):1570-7.

Cramer GM, Ford RA & Hall RL (1978). Estimation of Toxic Hazard - A Decision Tree Approach. Food and Cosmetics Toxicology 16, 255-276.

Edgett, S. J. (2015). Idea-to-Launch (Stage-Gate®) Model: An Overview. Stage-Gate International.

EFSA. (2014). Scientific opinion on good modelling practice in the context of mechanistic effect models for risk assessment of plant protection products. *EFSA journal, 12*(3): 3589

EFSA (2015a). Scientific Opinion on the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs. EFSA Journal 13, 3978. https://doi.org/10.2903/j.efsa.2015.3978

EFSA (2015b). Scientific Opinion on acrylamide in food. EFSA Journal 13, 4104. https://doi.org/10.2903/j.efsa.2015.4104

EFSA (in preparation). Guidance on Uncertainty in EFSA Scientific Assessment - Revised Draft for Internal Testing, https://www.efsa.europa.eu/sites/default/files/160321DraftGDUncertaintyInScientificAssessment.pdf

EMA European Medicine Agency. (2016). Draft "Guideline on the qualification and reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation." http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2016/07/WC500211315.pdf.

Embry, M. R., Bachman, A. N., Bell, D. R., Boobis, A. R., Cohen, S. M., Dellarco, M., & Pastoor, T. P. (2014). Risk assessment in the 21st century: Roadmap and matrix. Critical reviews in toxicology, 44(sup3), 6-16.

Fallahi-Sichani M, Honarnejad S, Heiser LM, Gray JW, Sorger PK. (2013) Metrics other than potency reveal systematic variation in responses to cancer drugs. Nat Chem Biol. 9 (11):708-14.

Federal Register Notice (2016) Request for Nominations for Peer Reviewers for EPA's Draft Biologically Based Dose-Response (BBDR) Model for Perchlorate, Draft Model Support Document and Draft Approach for Deriving a Maximum Contaminant Level Goal (MCLG) for Perchlorate in Drinking Water, In EPA–HQ–OW–2009–0297; FRL–9947–20–OW.

Geraets L, Oomen AG, Krystek P, Jacobsen NR, Wallin H, Laurentie M, Verharen HW, Brandon EF, De Jong WH. (2014). Tissue distribution and elimination after oral and intravenous administration of different titanium dioxide nanoparticles in rats. Part Fibre Toxicol 11: 30.

Ghibellini, G., Leslie, E.M., & Brouwer K.L.R. (2006). Methods to Evaluate Biliary Excretion of Drugs in Humans: an Updated Review. *Mol Pharm. 3*(3): 198–211.

Groothuis, F.A., Heringa, M.B., Nicol, B., Hermens, J.L., Blaauboer, B.J. and Kramer, N.I. (2015). Dose metric considerations in in vitro assays to improve quantitative in vitro-in vivo dose extrapolations. Toxicology, 332, 30-40.

Heringa MB, Geraets L, Van Eijkeren JC, Vandebriel RJ, De Jong WH, Oomen AG. (2016). Risk assessment of titanium dioxide nanoparticles via oral exposure, including toxicokinetic considerations. Nanotoxicology 10: 1515-1525.

ICH (1995). CPMP/ICH/384/95. Note for guidance on toxicokinetics: a guidance for assessing systemic exposure in toxicology studies.

Jamei, M., Turner, D., Yang, J., Neuhoff, S., Polak, S., Rostami-Hodjegan, A., & Tucker, G. (2009). Population-Based Mechanistic Prediction of Oral Drug Absorption. The AAPS Journal, 11(2), 225–237.

Jamei, M. (2016). Recent advances in development and application of physiologically-based pharmacokinetic (PBPK) models: a transition from academic curiosity to regulatory acceptance. Current pharmacology reports, 2(3), 161-169.

Kaizer, J.S., Heller, A.K., & Oberkampf, W.L. (2015). Scientific computer simulation review. Reliab. Eng. Syst. Saf., 138:210218.

Kirchmair, J., Göller, A. H., Lang, D., Kunze, J., Testa, B., Wilson, I. D., ... & Schneider, G. (2015). Predicting drug metabolism: experiment and/or computation?.(2015). Nature Reviews Drug Discovery, 14(6), 387-404

Kramer, N. I., Busser, F. J., Oosterwijk, M. T., Schirmer, K., Escher, B. I., & Hermens, J. L. (2010). Development of a partition-controlled dosing system for cell assays. Chemical research in toxicology, 23(11), 1806-1814.

Kramer, N. I. (2010).Measuring, modeling, and increasing the free concentration of test chemicals in cell assays. Utrecht University.Kroes R, Renwick AG, Feron V, Galli CL, Gibney M, Greim H, Guy RH, Lhuguenot JC, van de Sandt JJM (2007) Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. Food Chem. Toxicol. 45: 2533-2562

Leong, R., Vieira, M. L. T., Zhao, P., Mulugeta, Y., Lee, C. S., Huang, S.-M. and Burckart, G. J. (2012). Regulatory Experience With Physiologically Based Pharmacokinetic Modeling for Pediatric Drug Trials. Clinical Pharmacology & Therapeutics, 91(5): 926–931. doi:10.1038/clpt.2012.19

Loizou, G. D., and Hogg, A. (2011). MEGen: A Physiologically Based Pharmacokinetic Model Generator. *Frontiers in Pharmacology,* 2 ( 56), 1-14, 10.3389/fphar.2011.00056.

Loizou, G. D., Spendiff, M., Barton, H. A., Bessems, J., Bois, F. Y., Bouvier, d. Y., et al. (2008). Development of Good Modelling Practice for Physiologically Based

Pharmacokinetic Models for Use in Risk Assessment: The First Steps. *Reg. Toxicol. Pharmacol., 50*(3), 400-411.

Lombardo, F., Obach, R.S., Varma, M.V., Stringer, R. & Berellini, G. Clearance Mechanism Assignment and Total Clearance Prediction in Human Based upon in Silico Models, *J. Med. Chem., 57*(10), 4397−4405.

Lu,J. Goldsmith,M.-R., Grulke, C.M.,Chang, D.T., Brooks, R.D., Leonard J.A., … & Johnson, J..(2016). Developing a physiologically-based pharmacokinetic model knowledgebase in support of provisional model construction. *PLoS Comput. Biol*., 12 (2), e1004495.

Lumen, A., Mattie, D. R., and Fisher, J. W. (2013). Evaluation of perturbations in serum thyroid hormones during human pregnancy due to dietary iodide and perchlorate exposure using a biologically based dose-response model, Toxicological sciences, 133(2), 320-341.

Marchant, C.A., Rosser, E.M., & Vessey, J.D. (2017). A k- Nearest Neighbours Approach Using Metabolism related Fingerprints to Improve In Silico Metabolite Ranking. Molecular informatics, 36(3).

McNally, K., Cotton, R., and Loizou, G. (2011). A workflow for global sensitivity analysis of PBPK models. Frontiers in Pharmacology 2 (31), 1-21.

McNally, K., Cotton, R., Cocker, J., Jones, K., Bartels, M., Rick, D., 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, 18, 10.1155/2012/760281.

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

Meek, M.E. B., Barton, H,A., Bessems, J.G., Lipscomb, J.C. & Krishnan, K. (2013). Case study illustrating the WHO IPCS guidance on characterization and application of physiologically based pharmacokinetic models in risk assessment. Regulatory Toxicology and Pharmacology, 66(1), 116-129.

Mostrag-Szlichtyng, A, & Worth, A. (2010). In silico modelling of microbial and human metabolism: a case study with the fungicide carbendazim. JRC Technical Report EUR 24377 EN, 2010.

Munro IC, Ford RA, Kennepohl E, Sprenger JG (1996) Correlation of structural class with No-Observed-Effect Levels: A proposal for establishing a Threshold of Concern. Food Chem. Toxicol. 34: 829-867

Paini, A., Leonard, J.A., Kliment, T., Tan, Y.M. & Worth, A. (2017). Investigating the state of physiologically based kinetic modelling practices and challenges associated with gaining regulatory acceptance of model applications. Regulatory Toxicology and Pharmacology 90, 104-115.

Patterson, E.A., Taylor, R.J. & Bankhead, M., (2016). A framework for an integrated nuclear digital environment, *Progress in Nuclear Energy, 87*:97-103.

Patterson, E.A., & Whelan, M.P. (2017). A framework to establish credibility of computational models in biology, *Progress in Biophysics & Molecular Biology, 129*:13-19. doi: 10.1016/j.pbiomolbio.2016.08.007

Pearce, R. G., Setzer, R. W., Strope, C. L., Sipes, N. S., & Wambaugh, J. F. (2017). Httk: R package for high-throughput toxicokinetics. Journal of Statistical Software, 79 (i04).

Pirovano, A., Brandmaier, S., Huijbregts, M.A.J., Ragas, A.M.J., Veltman, K., & Hendriks, A.J. (2015). The utilisation of structural descriptors to predict metabolic constants of xenobiotics in mammals. *Environmental Toxicology and Pharmacology, 39*(1), 247–258.

Poulin P. & Theil F.P. (2000) A priori prediction of tissue:plasma partition coefficients of drugs to facilitate the use of physiologically-based pharmacokinetic models in drug discovery. J. Pharm. Sci. 89, 16–35.

Poulin P., Schoenlein K. & Theil F.P. (2001)Prediction of adipose tissue: plasma partition coefficients for structurally unrelated drugs. J. Pharm. Sci. 90, 436–447.

Price et al., (2003) Modeling interindividual variation in physiological factors used in PBPK models of humans. Crit Rev Toxicol 33(5):469-503.

Przybylak, K.R., Madden, J.C., Covey-Crump, E., Gibson, L., Barber, C., Patel, M., & Cronin, M.T.D. (2017). Characterisation of data resources for in silico modelling: benchmark datasets for ADME properties. Expert Opinion on Drug Metabolism & Toxicology, 1-13.

Punt A., Peijnenburg A.A.C.M., Hoogenboom R.L.A.P., Bouwmeester H. (2017) Non-animal approaches for toxicokinetics in risk evaluations of food chemicals. ALTEX 34(4):501-514.

Rietjens, I.M., Louisse, J., Punt, A. (2011). Tutorial on physiologically based kinetic modeling in molecular nutrition and food research. Molecular nutrition & food research, 55(6), 941-956.

Ring C.L., Pearce R.G., Setzer R.W., Wetmore B.A., Wambaugh J.F.. (2017) Identifying populations sensitive to environmental chemicals by simulating toxicokinetic variability. Environ Int. Sep;106:105-118.

Rotroff D.M., Wetmore B.A., Dix D J., Ferguson S.S., Clewell H.J., Houck K.A., Lecluyse E.L., Andersen M.E., Judson R.S., Smith C.M., et al. (2010). Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. Toxicol. Sci. 117, 348–358

Sala Benito J.V., Paini A., Richarz A.N., Meinl T., Berthold M.R., Cronin M.T., Worth A.P. (2017). Automated workflows for modelling chemical fate, kinetics and toxicity. Toxicol In Vitro. Volume 45, Part 2, Pages 249-257

Schruben, L.W. (1980). Establishing the credibility of simulations. *Simulation, 34*(3):101-105.

Schmitt W. (2008) General approach for the calculation of tissue to plasma partition coefficients. Toxicol. In Vitro 22, 457–467.

Schultz T.W., Amcoff P., Berggren E., Gautier F., Klaric M., Knight D.J., Mahony C., Schwarz M., White A., Cronin M.T. (2015). A strategy for structuring and reporting a read-across prediction of toxicity. *Regul Toxicol Pharmacol., 72*(3):586-601.

Suciu, N., Tediosi, A., Ciffroy, P., Altenpohl, A., Brochot, C., Verdonck, F., et al. (2016). Potential for MERLIN-Expo, an advanced tool for higher tier exposure assessment, within the EU chemical legislative frameworks. *Sci Total Environ, 562*, 474-479., 10.1016/j.scitotenv.2016.04.072.

US EPA (U.S. Environmental Protection Agency). (2006). Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment (Final Report). National Center for Environmental Assessment, Washington, DC. EPA/600/R- 05/043F.

Teeguarden, J.G. and Barton, H.A. (2004). Computational modeling of serum-binding proteins and clearance in extrapolations across life stages and species for endocrine active compounds*. Risk Anal 24*(3), 751-770.

Tonnelier A., Coecke S., Zaldívar J.M. (2012). Screening of chemicals for human bioaccumulative potential with a physiologically based toxicokinetic model. Arch. Toxicol. 86, 393–403.

US FDA (U.S. Food and Drug Administration). (2016). Draft "Physiologically Based Pharmacokinetic Analyses — Format and Content Guidance for Industry". https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM531207.pdf

Van Hoey S., Seuntjens P., van der Kwast J., Nopens, I. (2014) A qualitative model structure sensitivity analysis method to support model selection, In Journal of Hydrology, Volume 519, Part D, 2014, Pages 3426-3435.

Wagner, C., Zhao, P., Pan, Y., Hsu, V., Grillo, J., Huang, S. and Sinha, V. (2015). Application of Physiologically Based Pharmacokinetic (PBPK) Modeling to Support Dose Selection: Report of an FDA Public Workshop on PBPK. CPT: *Pharmacometrics Syst. Pharmacol., 4*(4): 226–230. doi:10.1002/psp4.33

Wambaugh, J.F., Wetmore, B.A., Pearce, R., Strope, C., Goldsmith, R., Sluka, J.P., Sedykh, A., Tropsha, A., Bosgra, S., Shah, I., Judson, R., Thomas, R.S., Setzer R.W. (2015) Toxicokinetic triage for environmental chemicals. Toxicol. Sci., 147 pp., 55-67.

Wang (2010) Confidence Assessment of the Simcyp Time-Based Approach and a Static Mathematical Model in Predicting Clinical Drug-Drug Interactions for Mechanism-Based CYP3A Inhibitors. Drug Metabolism and Disposition, 38 (7) 1094-1104.

Wetmore B.A., Wambaugh J.F., Ferguson S.S., Sochaski M.A., Rotroff D.M., Freeman K., Clewell H.J., III, Dix D.J., Andersen M.E., Houck K.A., et al. (2012). Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. Toxicol. Sci. 125, 157–174.

Wetmore B.A., Wambaugh J.F., Ferguson S.S., Li L., Clewell H.J., Judson R.S., Freeman K., Bao W., Sochaski M.A., Chu T.-M., et al. (2013). Relative impact of incorporating pharmacokinetics on predicting in vivo hazard and mode of action from high-throughput in vitro toxicity assays. Toxicol. Sci. 132,327–346.

Wetmore B.A., Allen B., Clewell H.J., Parker T., Wambaugh J.F., Almond L.M., Sochaski M.A., Thomas R.S. (2014). Incorporating population variability and susceptible subpopulations into dosimetry for high-throughput toxicity testing. Toxicol. Sci. 142, 210–224.

Wetmore B.A. (2015) Quantitative in vitro-to-in vivo extrapolation in a high-throughput environment Toxicology 332, 5 June 2015, Pages 94-101

WHO/IPCS (World Health Organization. International Programme on Chemical Safety). (2010). Characterization and Application of Physiologically Based Pharmacokinetic Models in Risk Assessment. Harmonization Project Document No. 9. Geneva, Switzerland.

Wilk-Zasadna I, Bernasconi C, Pelkonen O, Coecke S (2015) Biotransformation in vitro: An essential consideration in the quantitative in vitro-to-in vivo extrapolation (QIVIVE) of toxicity data. Toxicology. 2015 Jun 5;332:8-19.

Yoon, M., Campbell, J.L., Andersen, M.E. and Clewell, H.J. (2012). Quantitative in vitro to in vivo extrapolation of cell-based toxicity assay results. Critical reviews in toxicology, 42(8), 633-652.

Yoon M., Kedderis G.L., Yan G.Z., Clewell H.J. 3rd. (2015) Use of in vitro data in developing a physiologically based pharmacokinetic model: Carbaryl as a case study. Toxicology. 5;332:52-66.

Yoon M., Efremenko A., Blaauboer B.J., Clewell H.J. (2014). Evaluation of simple in vitro to in vivo extrapolation approaches for environmental compounds. Toxicol. in vitro 28, 164–170.

Yoshida, K., Budha, N., and Jin, J.Y. (2017). Impact of Physiologically Based Pharmacokinetic Models on Regulatory Reviews and Product Labels: Frequent Utilization in the Field of Oncology. Clinical Pharmacology & Therapeutics 101(5) 597-602.

Zaldivar Comenges, J.M., Joossens, E., Sala Benito, J.V., Worth, A., Paini A., (2017). Theoretical and mathematical foundation of the virtual cell based assay – a review. Toxicol. In Vitro Volume 45, Part 2, Pages 209-221

Zhuang, X., and Lu, C. (2016). PBPK modelling and simulation in drug research and development. *Acta Pharmaceutica Sinica B 6*(5):430–440.

## List of abbreviations and definitions

| | |
|---|---|
| ADME | Absorption, distribution, metabolism, excretion |
| AEP | Aggregate Exposure Pathways |
| AOP | Adverse Outcome Pathways |
| BBDR | Biologically Based Dose-Response |
| Caco 2 | Colorectal adenocarcinoma-2 cell line |
| CEN | European Committee for Standardization |
| COSMOS | Computational Tools for Safety Assessment Focusing on Cosmetics Ingredients |
| CSAF | Chemical-Specific Adjustment Factors |
| CWA | CEN workshop agreement |
| EAGMAST | Extended Advisory Group on Molecular Screening and Toxicogenomics |
| EFSA | European Food Safety Authority |
| EMA | European Medicines Agency |
| EPA | Environmental Protection Agency |
| EPAA | European Partnership for Alternative Approaches to Animal Testing |
| ESTAF | EURL ECVAM Scientific Advisory Committee |
| EURL ECVAM | European Union Reference Laboratory for alternatives to animal testing |
| FDA | Food and Drug Administration |
| GI | Gastrointestinal |
| GIVIMP | Good In Vitro Method Practices |
| GMP | Good Modelling Practice |
| GSA | Global Sensitivity Analysis |
| HSE | Health and Safety Executive |
| HSL | HSE's Health and Safety Laboratory |
| httt | High Throughput Toxicokinetics |
| IATA | Integrated Approaches to Testing and Assessment |
| INERIS | Institut national de l'environnement industriel et des risques |
| IPCS | International Programme of Chemical Safety |
| IVIVE | in vitro to in vivo Extrapolation |
| $K_m$ | Michaelis Constant (enzyme-specific constant that describes the substrate concentration in which the velocity of the reaction is half of Vmax) |
| Log D | Logarith of base 10 of the partition octanol-water considering the ionization forms. |
| Log P | Logarithm of base of the partition octanol-water |
| KNIME | Konstanz Information Miner |
| MDCK | Madin–Darby canine kidney cell line |

| | |
|---|---|
| MEGEN | Model Equation Generator |
| MERLINExpo | Modelling Exposure to chemicals for risk assessment: a comprehensive Library of multimedia and PBK models for integration, prediction, uncertainty and sensitivity analysis. |
| MoA | Modes of Action |
| NAM | New approach methodologies |
| NHANES | National Health and Nutrition Examination Survey |
| NIEHS | National Institute of Environmental Health Sciences |
| OchemOnline | Chemical modelling environment |
| OECD | Organization for Economic Cooperation and Development |
| PARERE | (EURL-ECVAM's network for)Preliminary Assessment of Regulatory Relevance |
| PBD | Physiologically Based Dynamic models |
| PBK | Physiologically Based Kinetic models |
| PBPK | Physiologically Based pharmacokinetic models |
| PK | Pharmacokinetics |
| PLETHEM | Population Lifecourse Exposure-To-Health-Effects Model Suite (computational platform) |
| PoD | Point of Departure |
| QSAR | Quantitative Structure-Activity Relationship |
| QIVIVE | Quantitative IVIVE |
| REACH | Registration, Evaluation, Authorisation and Restriction of Chemicals |
| RIVM | Netherlands National Institute for Public Health and the Environment |
| SCCS | Scientific Committee on Consumer Safety |
| TD | Toxicodynamics |
| TK | Toxicokinetics |
| UF | Uncertainties Factors |
| $V_{max}$ | Maximum velocity of an enzymatic reaction |
| WHO | World Health Organization |

## Terminology definition

Validation: In the CEN WS document "Standard documentation of large chemical exposure models" the term "Validation" is used in the context of Model Evaluation. The document attests: Model evaluation is seen here as the assessment of how accurately mathematical models represent the real world, e.g. the complexity of environmental and/or human systems. Typically the evaluation of complex exposure models is difficult to conduct because empirical data are seldom consistent regarding space and time and because key input data are often lacking. Model evaluation is frequently based on comparisons between the output from deterministic simulations and that from single experiments. However, it is now widely recognized that the impact of uncertainty and variability should be integrated in model validation". However, there is not a specific definition for the term "Validation ".

## List of figures

## List of tables

# Annexes

## Annex 1. Pre conference Material

### Proposed questions/statements for discussion

What do we need to provide in order to have PBK models accepted by regulators, in order to reach a whole new level in regulatory decision making?

**Regulatory Acceptance: Challenges in using PBK model to support regulatory decision making**

- Do we need guidance? if not what do we need
- Integration in/with IATA/AEP/AOP: Where is applied and how can PBK model play a role in IATA and AEP/AOP framework.

**Challenges in assessing model credibility**

- Model complexity versus simplification (how confident can you be?)
- Validate/Verify the model code (mass balance) and the model predictions; *Identify challenges in evaluating model performance when no in vivo kinetic data are available for comparison. When model simulations cannot be evaluated/validated using in vivo data, how do we increase confidence in the in vitro and in silico tools used to parameterize a PBK model?*
- Sensitivity analysis Global versus Local
- Uncertainties

**Challenges in constructing models with no in vivo kinetic & dynamic data**

- Try to relay only on the use of in vitro data and in silico prediction when building your PBK model Quality and quantity of input parameters (which are mandatory and which are optional?)
- Develop the conceptual model, identify key model components required to properly describe a chemical's ADME behaviours to achieve the intended purpose; *Identify challenges in developing a conceptual model when in vivo kinetic data (e.g., time course of blood/tissue concentrations) are not available to inform a chemical's ADME behaviours.*
- Translate the mathematical equations into computational codes; *Identify challenges in peer review process when reviewers may not have the programming software used to code the PBK model, or may not have the programming skills to review a PBK model coded in unfamiliar programming language.*
- Define and obtain the model parameters using high-quality *in vitro* measurements or *in silico* predictions; *Identify challenges in applying in vitro and in silico methods to generate values of ADME parameters, or locating existing resources (e.g., databases, publications) that contain required data.*
- Use of in vitro effect (TD) data to link effect to dose: QIVIVE
- Translate the in vitro concentration to actual in vivo human relevant dose. i) Physiology IVIVE, scale up; Pharmacology/pharmacokinetic IVIVE, simulation of effect in vitro concentration to external; Kinetic & dynamic IVIVE, linking in vitro data of effect to external dose [QIVIVE, comparison, forcing etc].

**Regulatory Acceptance: Solutions**

- What to do next? How to apply/implement good modelling practice? Approach OECD/CEN? To develop a guidance document?
- Propose Training to Regulators and Scientists of these tools.
- Propose to develop a Tool box (similar to the QSAR toolbox) to integrate in a common place a library of PBK models (Lu et a., 2016) for application in Risk

assessment; provide links to software's that can be used to build PBK models, and to integrate in the software generic and specific uncertainty and SA.
- Define case studies?

Additional discussion bullets provided by workshop experts:

**Dr. Woody Setzer (EPA, USA)**

- Any model used in risk assessment (or for any other purpose) needs to be accompanied with an honest quantitative assessment of its likely precision (e.g., measured concentrations are likely to be within an order of magnitude of predicted concentrations half the time). Note that this includes not just uncertainty in parameter estimates, but uncertainty about model structure and issues involved in extrapolation (e.g., how likely is it that there are unmodeled processes that result in a two-fold change in predicted concentrations?)
- What tools do we have to give credibility to such statements, and how acceptable would they be to risk assessors?
- One way to think about tiering or risk assessments is that as we go from, say screening and prioritization, to individual risk assessments for critical chemicals with potentially high exposures, we need increasingly precise model predictions. How do we calibrate this (e.g., we need to be within 100-fold for screening, but within 2-fold for a high-visibility assessment).
- All this applies to assessment of variability, too.

**Dr. Cecilia Tan (EPA, USA)**

- Do the regulatory agencies, including internal experts, decision makers, and external reviewers, have the expertise necessary to evaluate a PBK model (including structure, parameters, predictive capability), as well as the appropriate use in risk assessment? If not, how do we help build that expertise?
- How many tiers do we really have for PBK applications? In reality, is it possible that regulatory agencies will accept either (1) screening/prioritization for models that have no in vivo data to calibrate/evaluate; and (2) replacing default uncertainty factors with model predictions for models that have in vivo data to calibrate/evaluate?
- Acknowledging the importance of identify metabolites (and its hazard), what is the level of uncertainty allowed for PBK predictions when the level of uncertainty in exposure predictions may outweigh any uncertainty in PBK predictions?
- While a generic PBK model that includes intrinsic clearance rates and protein binding may be sufficient for many chemicals, how do we identify those chemicals that have some unique PK properties (e.g., perfluroonated chemicals)?

**Dr. George Loizou (HSL, UK)**

- Can an open access free-to-use modelling platform increase uptake and acceptance of PBK modelling in general and for regulatory applications in particular?
- Can such a platform help improve the model peer-review process?
- How can such a resource be maintained?

**Dr. John Paul Gosling (UNI Leeds, UK)**

- What matters more: having a model that effectively captures reality or having a model that is adequate for our purposes?
- How much do we trust existing in vivo data that is used to parameterise PBK models?

**Dr. Minne Heringa (RIVM, NL)**

- What questions are there, for which we need PBK models?
- What features should a PBK model have to answer a certain one of these questions? What criteria should it meet then?
- What parameters do we then need to determine with e.g. in vitro assays and this possible yet?
- How do we ensure the PBK models are user-friendly?

## Annex 2. OECD proposal for a guidance on PBK modelling

### Sent to EGMAST / WPHA – OECD - in December 2016

### A proposal for the development of an OECD Guidance Document for characterising, validating and reporting Physiologically Based Kinetic (PBK) models intended for regulatory application that are based on data derived from non-animal methods

Physiologically Based Kinetic (PBK)[13] modelling is a scientifically-sound approach to predict internal dose metrics for chemical risk assessment applications. Traditionally, the calibration of internal model parameters and the validation of the performance of the PBK model rely heavily on relevant *in vivo* data. However, due to the advancement of scientific knowledge concerning Absorption, Distribution, Metabolism and Excretion (ADME) processes within target species, the improvement in ADME/TK specific *in vitro* and *in silico* tools and the widespread availability of sophisticated modelling platforms, there is a strong shift towards the development of PBK models which rely primarily on non-animal data for their construction. Another important aspect to consider in relation to this new generation of PBK models is that there are typically very little *in vivo* reference data for the chemicals and species of interest to provide a basis for traditional quantitative model validation (i.e. statistical comparison of model predictions with equivalent *in vivo* measurements for a set of reference/target chemicals).

The ultimate aim of the Guidance will be to establish the credibility of this class of PBK models in order to promote their acceptance and use in a regulatory context. The intention is to provide practical guidance to model developers and end-users on i) how to properly characterise their PBK model (e.g. model elements and construction, underlying principles and assumptions, internal parameters and their estimation, model implementation/coding etc.), ii) how to validate the model to assess its performance (e.g. model verification, sensitivity analysis, uncertainty assessment, applicability domain, limitations, etc.) and iii) how to comprehensively report and describe the model in terms and a format that an end-user (e.g. risk assessor) of the data could readily understand and trust.

There are several existing guidance documents on developing and reporting traditional PBK models that are calibrated and evaluated using *in vivo* data e.g. WHO (2010)[14], EPA (2007)[15], CEN CWA 16938 (2015)[16]. The European Medicines Agency (EMA) is currently (2016-2017) working on a guidance document for qualification and reporting of PBK models which are focused primarily on the assessment of drug-drug interactions and understanding TK in children, as well as qualification of PBK modelling platforms commonly used to support TK aspects of the regulatory submission of new drugs[17]. Although aspects of these guidance documents are relatively universal and should be taken into account, they are all based on the premise that model performance must be ultimately assessed by direct comparison on predicted TK profiles with equivalent *in vivo* data.

In a recent international PBK model workshop held at the JRC in November 2016, experts recommended that such a Guidance Document be developed at international level.

---

[13] PBK is synonymous with PBPK, PBBK, PBTK.

[14] http://www.who.int/ipcs/methods/harmonization/areas/pbpk/en/

[15] https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=157668&CFID=72277452&CFTOKEN=72162106

[16] https://www.cen.eu/work/areas/chemical/Pages/WS-MerlinExpo.aspx

[17] http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/clinical_pharmacology_pharmacokinetics/general_content_001729.jsp&mid=WC0b01ac0580032ec5

**GETTING IN TOUCH WITH THE EU**

**In person**

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: http://europea.eu/contact

**On the phone or by email**

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: http://europa.eu/contact

**FINDING INFORMATION ABOUT THE EU**

**Online**

Information about the European Union in all the official languages of the EU is available on the Europa website at: http://europa.eu

**EU publications**
You can download or order free and priced EU publications from EU Bookshop at: http://bookshop.europa.eu. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact).

KJ-NA-28794-EN-N

## JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

**EU Science Hub**
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/619902

ISBN 978-92-79-73848-7