European Food Safety Authority

EFSA Journal 2014;12(4):3638

SCIENTIFIC REPORT OF EFSA

Modern methodologies and tools for human hazard assessment of chemicals[1]

European Food Safety Authority[2,3]

European Food Safety Authority (EFSA), Parma, Italy

*This scientific output, published on 11 July 2014, replaces the earlier version published on 24 April 2014**

ABSTRACT

This scientific report provides a review of modern methodologies and tools to depict toxicokinetic and toxicodynamic processes and their application for the human hazard assessment of chemicals. The application of these methods is illustrated with examples drawn from the literature and international efforts in the field. First, the concepts of mode of action/adverse outcome pathway are discussed together with their associated terminology and recent international developments dealing with human hazard assessment of chemicals. Then modern methodologies and tools are presented including *in vitro* systems, physiologically-based models, *in silico* tools and OMICs technologies at the level of DNA/RNA (transcriptomics), proteins (proteomics) and the whole metabolome (metabolomics). Future perspectives for the potential applications of these modern methodologies and tools in the context of prioritisation of chemicals, integrated test strategies and the future of risk assessment are discussed. The report concludes with recommendations for future work and research formulated from consultations of EFSA staff, expert Panels and other international organisations.

KEY WORDS

mode of action, adverse outcome pathway, integrated testing strategy, physiologically-based models, *in silico*, OMICs

[1] On request from EFSA, Question No EFSA-Q-2012-00262, approved on 31 March 2014.

[2] Correspondence: scer@efsa.europa.eu

[3] Acknowledgement: EFSA wishes to thank EFSA staff: Jean Lou Dorne, Andrea Altieri, Camille Bechaux, Charlotte Bergkvist, Arianna Chiusolo, Gina Cioacata, Andrea Germini, Georges Kass, Anna Maria Rossi, Andrea Terron, Matthew Watts and the staff of the Joint Research Centre of the European Commission: Claude Guillou, Alicia Paini and Andrew Worth for the contribution and support provided to this scientific output and Claude Lambré for reviewing the report, and the Scientific Committee, particularly Jan Alexander, Diane Benford, Qasim Chaudhry and Joseph Schlatter, for reviewing the report and providing comments.

*A minor change of an editorial nature was made; it does not materially affect the contents of this Report. To avoid confusion, the original version of the Report has been removed from the website, but is available on request, as is a version showing all the changes made.

Suggested citation: European Food Safety Authority, 2014. Modern methodologies and tools for human hazard assessment of chemicals. EFSA Journal 2014;12(4):3638, 87 pp. doi:10.2903/j.efsa.2014.3638

Available online: www.efsa.europa.eu/efsajournal

SUMMARY

Hazard assessment of chemicals for humans comprises hazard identification and hazard characterisation through understanding of toxicokinetic (TK) and toxicodynamic (TD) processes. Traditionally, a pivotal toxicological study in test species is identified for a specific chemical to determine a reference point which is then used to derive either a health-based guidance value or a margin of exposure. Over the last decade, a number of modern *in vivo, in vitro* methodologies and *in silico* tools have been developed to investigate TK and TD processes of chemicals, i.e. Mode of Action (MoA)/Adverse Outcome Pathway (AOP) at different levels of biological organisation (organism, organ, cellular and molecular level). These methodologies provide the opportunity to move towards a mechanistic understanding of toxicity and give options for Integrated Testing Strategies (ITS) to reduce animal use in toxicological research. These modern methodologies are reviewed in this scientific report to present their potential use in the future of human hazard assessment of chemicals with a view to anticipating their future use within EFSA's work.

Currently, MoA/AOP information is not often available for specific chemicals, and risk assessors rely often on the dose response assessment to translate external dose to a quantitative reference point for hazard characterisation in test species. However, recent international developments have supported the move towards elucidating such MoA/AOPs and these include the new applications of the WHO framework on MoA, the OECD international programme on AOPs. In addition, two research programmes, TOX-21in the US and SEURAT-1 in Europe, both deal with alternatives to animal testing such *as in vitro* methods and other integrated testing strategies (ITS). Strengths of ITS such as high throughput screening (HTS) assays include the possibility to screen and prioritise chemicals while minimising animal testing. Their limitations include their lack of prediction for a) chemically-induced disease-associated pathways, b) metabolism, c) interactions between different cell types, d) tissue-level cellular interactions, and e) chronic exposure.

A number of modern *in vitro* models based on human cells provide very useful tools to investigate TK processes (absorption, distribution, metabolism, excretion of chemicals (ADME)) in humans. The current updated OECD Test Guideline 417, mainly related to absorption and metabolism, has indicated that such *in vitro* models can provide supplemental TK information which may substantially reduce *in vivo* animal testing. Even though these *in vitro* models have still received little attention in hazard assessment of chemicals for the food safety area, they can provide key information on ADME such as bioavailability, protein binding and identify human transporters and metabolic pathways. These parameters can be used to determine the *in vivo* hepatic clearance of a chemical and then can be scaled up to the whole liver and take into account human variability to build physiologically-based (PB) models. Another challenge that remains to be solved, so as to apply these *in vitro* methods routinely, is *in vitro* to *in vivo* extrapolation (IVIVE) in order to reflect human physiology and metabolism (hepatic and extrahepatic such as intestinal metabolism) and incorporate human variability in quantitative IVIVE (QIVIVE).

PB models are presented for the modelling of TK processes (PB-TK) and for TK and TD processes (PB-TK-TD). PB-TK models provide a quantitative means to address TK processes and are therefore very useful tools in hazard assessment. PB-TK-TD link both the TK and the TD dimensions and are therefore more complex compared with PB-TK models. Their use has been recommended by regulatory authorities worldwide and the World Health Organization (WHO) and the United States Environmental Protection Agency (US-EPA) have highlighted the need to develop a guidance to pursue common principles for their application in chemical hazard assessment and risk assessment as a whole. Reservations regarding their routine use include the need for detailed knowledge of TK for a particular chemical, high level expertise and resources, and the need to validate the models. Consequently, PB-TK and PB-TK-TD models are mostly used in high-tier risk assessment (tier 3). PB-TK and PB-TK-TD models can also be developed using ITS, IVIVE and QIVIVE which remain challenges for the determination of both TK parameters and toxicity parameters. In the food safety area, these models have been mostly applied to pesticides, contaminants and food contact materials and are very useful to deal with key issues in hazard assessment such as interspecies differences,

human variability, biomonitoring programmes, combined exposure to multiple chemicals and *in vitro* to *in vivo* extrapolations.

In silico tools include (quantitative) structure activity relationships ((Q)SARs) and read-across methods that have been developed to predict a number of toxicological properties of chemicals using models, databases and tools. Another tool that is increasingly used in hazard assessment and risk assessment as a whole is the threshold of toxicological concern (TTC). QSARs are typically used in combination with other non-testing methods (such as read-across) and testing methods (such as *in vitro* methods) in the context of ITS and Weight-of-Evidence (WoE) assessments. Both (Q)SARs and read-across methods are increasingly predictive for hazard assessment, particularly for acute toxicity, mutagenicity, genotoxicity and bioacummulation. However, their predictability for TK properties (ADME) and sub-chronic and chronic toxicity is still limited and considerable research is undergoing to address these issues. It is foreseen that combining results from different Q(SAR) models, structural alerts, read-across estimates with *in vitro* and *in vivo* toxicological studies using a WoE approach will improve the use and validation of these tools and increase the overall reliability of *in silico* methods.

OMICs technologies are valuable tools to measure biochemical changes associated with a MoA/AOP, at the level of DNA/RNA (transcriptomics), proteins (proteomics) and the whole metabolome (metabolomics). They provide the means to identify biomarkers in humans and animals for dose response modelling, investigate interspecies differences and their human relevance and incorporate human variability (age differences, inter-ethnic differences, polymorphisms). OMICs technologies can also investigate patterns of gene transcripts, proteins, and metabolites within an AOP using *in vitro* models and provide helpful means to validate ITS using mechanistic *in vitro* assays to reduce animal studies and move towards predictive modelling. Weaknesses of OMICs methods include the need for complex molecular and analytical techniques, highly specialised training and sophisticated bioinformatic tools to analyse huge datasets. Another key issue relates to the sensitivity of the methodologies which may lead to the detection of changes that may not be biologically or toxicologically relevant. Finally, OMICs studies have a complex design and have been most often restricted to well known reference substances to allow researchers to correlate OMICs datasets with standardised endpoints (clinical chemistry, histopathological endpoints). It is foreseen that in the future, publicly available databases combining *in vitro* and *in vivo* OMIC datasets for large databases of compounds with MoA/AOP knowledge will help considerably to a) identify biomarkers associated with specific AOPs, and b) to bring new tools for predictive toxicology. Applications to human hazard assessment of chemicals in the food safety area have already been explored and include benchmark dose modelling from transcriptomic profiling, investigation of epigenomic mechanisms, identification of biomarkers of toxicity (proteomics), and investigation of MoA for single and multiple compounds (metabolomics).

A number of approaches have been developed for the prioritisation and ranking of chemicals according to their toxicological properties. At the US-EPA, the toxicological prioritisation index (ToxPi) decision support framework has been developed and enables the ranking of chemicals using multiple sources of evidence on toxicity and exposure surrogates. Future needs of ToxPi development include further studies to understand the relationship between simple exposure surrogates, tiered screening-level exposure assessments, and population-level biomonitoring data. In addition the US-EPA has also developed, during the NextGen project, a recent tiered approach as a prioritisation tool. In practice, Tier 1 aims to prioritise and screen chemicals using ITS (Toxcast HTS assays, *in vitro* genotoxicity tests, IVIVE TK models…) for further testing in Tiers 2 and 3. Tier 2 uses limited *in vivo* toxicity testing (e.g. short-term *in vivo* transcriptomic studies, *in vivo* studies to identify a point of departure for chemicals with a selective MoA, and IVIVE TK studies to link exposure and internal dose). Tier 3 is equivalent to the traditional toxicological *in vivo* testing in experimental animals.

For the identification of emerging chemical risks, EFSA has recently developed a systematic framework which uses a number of data sources as input, relating to the source of the chemical (industrial chemical, contaminant) and software models as tools to predict the environmental behaviour and potential toxicity of chemicals from structural features and physico-chemical properties

(e.g. QSAR models and PB-TK models). The application of the framework consists of a multi-step selection process starting with a list of chemicals to which a sequence of selection criteria is applied to identify the substances of potential concern. The selection criteria take into account a number of parameters including volumes of production or import data related to the chemical, its environmental persistence, bioaccumulation potential, dispersive uses, toxicity, and any available outcomes of previous risk assessments. Further work is ongoing to test the framework using additional data sources, selection criteria and through the development of databases and software, for the systematic identification of emerging chemical risks in the food chain.

Overall, this report has highlighted the shift towards a MoA/AOP approach in chemical risk assessment to depict TK and TD processes using ITS including *in vitro* methods based on human cells (e.g. HTS assays), OMICs, physiologically-based models and *in silico* tools. This paradigm shift will allow to a) move towards a systems toxicology view for human hazard assessment of chemicals, b) reduce animal use in toxicity testing, and c) provide support for the prioritisation of thousands of chemicals. Key issues remain to be solved and include the need to validate these methodologies, the inclusion of more case studies to test the methods and combine new knowledge and historical data for proof of concept, and the need for publicly accessible databases integrating data from such methods. It can be foreseen that, as knowledge of MoA/AOP advances, risk assessors and toxicologists will be able to refine models and tools for human risk assessment of chemicals (e.g. dynamic AOPs, complex cellular network models, integration of the impact of the human microbiome on TK and TD events).

Finally, recommendations are presented as the result of a general consultation of EFSA Panels and staff dealing with chemical risk assessment and other experts from international bodies (ECHA, OECD, WHO…) performed between April and October 2013. In the context of these modern methods and tools the need for harmonisation of the terminology and definitions is highlighted particularly for human health, animal health and environment risk assessment. A review to highlight the use of these modern methodologies and tools in animal health and environment risk assessment is also emphasised as well as the need for a guidance document on the use of the MoA in risk assessment. Finally, exploration of the applications of these tools to risk assessment of combined exposure to multiple chemicals and multiple stressors (e.g. biological hazards, physical agents etc) is also highlighted.

For TK processes, the lack of human TK data for chemicals represents a key data gap and such TK data are needed to allow the full characterisation of interspecies differences and human variability in these processes. This will provide a basis to link exposure, internal dose and toxicity using PB models. Finally, in the context of exposure to multiple chemicals, such human TK data will provide a scientific basis to set assessment groups based on TK criteria.

Other recommendations include:

a) Improvement of *in vitro* methods for generating TK data to measure human absorption, distribution, metabolism (gut and hepatic) and excretion patterns of chemicals.

b) Development of a guidance on the use of PB models in chemical risk assessment together with the development of prototype physiologically-based models using specific case studies to integrate exposure, toxicokinetic information and toxicity data, for hazard assessment purposes.

c) Developing databases providing critical parameters to build the models (physico-chemical, physiological, toxicological), and bioinformatic tools/algorithms in order to analyse and integrate such data.

Further work is needed to explore the application of *in silico* tools in chemical risk assessment, including the systematic and harmonised approach for the use of QSAR. Further development of the read-across methodologies are recommended, particularly using QSAR, physico-chemical properties and toxicological data together with refinements to the TTC approach.

For OMICs technologies, validation and standardisation of OMICs technologies for their use in human hazard assessment are needed, together with detailed guidance on criteria for acceptability. Further exploration of the use of OMICs in human hazard assessment using case studies relevant to the food and feed area is recommended and include benchmark dose modelling, consideration of *in vitro* data, use of biomarkers of exposure and effects. Application in other areas, such as animal health and ecological risk assessment, and nutrition are also foreseen.

Future work on ITS is recommended using specific chemicals as case studies. Testing should concentrate on differentiating chemicals with specific MoA and chemicals with non specific and or multiple MoA, giving opportunities to rank potencies for prioritisation. Alternative test species bridging *in vitro* methods and mammalian tests should be further explored. Further exploration of these new methodologies for hazard assessment are needed for both regulators and industry, to screen large sets of chemicals, prioritise chemicals, and to assess chemicals for a specific purpose. In the case of exposure to multiple chemicals, a better understanding of MoA/interaction of multiple substances using predictive and alternative methodologies will again allow to improve the basis for setting assessment groups.

Regarding the future of chemical risk assessment, exploration of new risk assessment frameworks to bring a systems toxicology perspective to risk assessment using case studies is needed. Weight of evidence and uncertainty analysis methodologies are also essential for the integration of data from new methodologies in the mode of action framework and chemical risk assessment as a whole. Finally, reinforcing collaborations with international institutions is critical for EFSA and highly recommended in order to integrate these new methods and facilitate international harmonisation.

TABLE OF CONTENTS

BACKGROUND AS PROVIDED BY EFSA

Methodologies to perform risk assessment are harmonised within EFSA, whether the hazard is of chemical or biological origin, and follow four steps namely hazard identification, hazard characterisation, exposure assessment, and risk characterisation. Since its foundation, a number of activities have been ongoing at EFSA to keep up to date with key methodologies and tools available to the risk assessor.

Modern methodologies and tools for chemical risk assessment are numerous for both exposure assessment and hazard assessment. For exposure assessment, advances in statistical methods have enabled risk assessors to develop probabilistic methods quantifying variability and uncertainty. Depending on the needs of the risk assessment, such methods can replace deterministic approaches. In relation to hazard assessment, a number of tools and methods are available such as OMICs technologies (e.g. genomics, proteomics, metabolomics, toxicogenomics, etc.) (Herrero et al., 2011; Kean, 2011) and other profiling techniques i.e. systems biology, biomarkers and biological pathway perturbations (NRC, 2007), *in vitro* and *in silico* methods such as quantitative structure activity relationships (QSAR), and biologically-based models such as physiologically-based toxicokinetic and toxicodynamic models. It is underlined that the possible applicability of tools such as *in vitro* and *in silico* methods would also contribute to the 3Rs (reduce, replace, refine) moving towards the reduction of animal use in toxicological research.

Keeping up to date with such modern methodologies and tools for hazard assessment, understanding their strengths and weaknesses and the purpose they may serve in evidence-based approaches for the prioritisation and ranking of chemicals according to their toxicological properties is a priority subject for the science strategy of EFSA. Since hazard identification and characterisation of chemicals specifically investigate toxicokinetic and toxicodynamic processes, the scientific report should focus on new and emerging methods and tools to measure and quantify these two processes. Additionally, the report should also contribute to further discussions on the integration of these methods and tools in human risk assessment as a whole i.e. integrate hazard and exposure assessment using a harmonised and consistent approach.

Needs for EFSA's future activities

In the discussions supporting the development of an EFSA Science Strategy, modern tools and methods in risk assessment have been identified as one of the priority tasks for the coming years. Hence, it is proposed to establish an internal task force composed of staff members from the three science Directorates, with the aim to prepare a scientific report reviewing the state of the science of new and emerging tools available for the hazard identification and hazard characterisation of chemicals. Such a document would also prepare further discussion on the possible integration of such methodologies and tools, and their applicability, to the human risk assessment of chemicals using a harmonised and consistent approach.

When preparing its scientific report, the task force will consider:

- Relevant work previously done and on-going in EFSA, e.g. the applicability of QSAR analysis to the evaluation of the toxicological relevance of metabolites and degrades of pesticide active substances for dietary risk assessment (EFSA, 2011), the use of physiologically-based toxicokinetic models and biomarkers for the hazard identification and characterisation of cadmium (Amzal et al., 2009; EFSA, 2009a, 2009b).
- On-going international activities, such as (i) the possible integration of OMICs in risk assessment frameworks, involving US-EPA, WHO, and OECD (US-EPA, 2004; OECD, 2009a), (ii) the Tox-21 project, involving the National Toxicology Program (NTP), the National Institutes of Health (NIH) and the Chemical Genomics Center (NCGC) in the US, and discussing new strategy for toxicity testing. (Schmidt, 2009), (iii) the investigation of tools for mode of action elucidation of chemicals and risk assessment including the

'effectopedia' by the working group of the International Program on Chemical Safety (IPCS) of the World Health Organisation (WHO) in collaboration with EFSA, the European Chemicals Agency and the European Commission Joint Research Centre.

TERMS OF REFERENCE AS PROVIDED BY EFSA

The EMRISK unit is requested to establish an internal task force with staff members from the science Directorates.

The specific aims of the task force are:

- To review and produce an EFSA scientific report regarding the state of the science of modern methods and tools available and their applicability to the hazard identification and characterisation of chemicals with specific regard to toxicokinetics and toxicodynamic processes, i.e. biologically-based models such as physiologically-based toxicokinetic models and OMIC technologies, respectively.

- Since hazard identification and characterisation of chemicals specifically investigate toxicokinetic and toxicodynamic processes, the scientific report should focus on new and emerging methods and tools to measure and quantify these two processes. The technical report should not review methods for exposure assessment which are beyond the scope of this exercise.

- To consider the applicability they may have in evidence-based approaches for the prioritisation and ranking of chemicals according to their toxicological properties.

- The report should include a discussion on the strengths and weaknesses of such methodologies and tools for hazard assessment and their possible integration in human risk assessment as a whole i.e. in relation to exposure assessment, using a harmonised and consistent approach.

- To discuss the outcomes of this work with the Scientific Committee for further consideration.

- To present the outcome of this work to the Scientific network on harmonisation of risk assessment.

CONTEXT OF THE SCIENTIFIC OUTPUT

This internal mandate is in direct relation with EFSA's Science Strategy (2012-2016) and complements the work of the Scientific Committee and all units dealing with the scientific evaluation of chemical hazards (pesticides, contaminants, food contact materials, food and feed additives) and the work of the SCER unit on the chemical hazards database.

EVALUATION

1. Introduction

Over the last 50 years, human risk assessments of chemicals have been performed for thousands of substances by international and national bodies dealing with food safety and consumer safety such as the Joint Expert Committee on Food Additives (JECFA) of the Food and Agriculture Organization (FAO)/World Health Organization (WHO), the US Environmental Protection Agency (US-EPA), the US Food and Drug Administration (FDA), the European Food Safety Authority (EFSA), the European Chemicals Agency (ECHA) within the REACH directive, The European Commission's Scientific Committee of Consumer Safety for Cosmetic Products (SCCS), the European Medicines Agency for pharmaceuticals - to cite but a few. The founding regulation of EFSA has defined risk assessment as 'a scientifically based process consisting of four steps: hazard identification, hazard characterisation, exposure assessment and risk characterisation' (EC, 2002). The WHO has described; 'principles and methods for the risk Assessment of chemicals in food;' in the Environmental Health Criteria monograph 240 (WHO, 2009).

Hazard assessment involves **hazard identification** which is 'the identification of the type and nature of adverse effects that an agent has an inherent capacity to cause in an organism, system or (sub)-population' and **hazard characterisation** which is 'the qualitative and, wherever possible, quantitative description of the inherent properties of an agent or situation having the potential to cause adverse effects'. The hazard characterization step should, where possible, include an assessment of dose-response and an evaluation of uncertainties (WHO, 2009). In traditional practice, an **apical endpoint** is identified as: **'**an observable outcome in a whole organism, such as a clinical sign or pathologic state, that is indicative of a disease state that can result from exposure to a toxicant' (Krewski et al., 2011). For a given chemical, the apical endpoint is identified on the dose-response relationship from the pivotal toxicity study in test species (rat, mouse, rabbit, dog). The apical point occurs at lower dose than other effects and is the basis to identify a Reference Point (RP) such as No-Observed-Adverse-Effect Level (NOAEL), Lowest-Observed-Adverse-Effect-Level (LOAEL) or the lower confidence limit of the Benchmark Dose (BMDL) (EFSA, 2009a, 2013).

For threshold (non-genotoxic) chemicals, the RP is often divided by a 100-fold default uncertainty factor to allow for inter-species and inter-individual variability in toxicokinetics and toxicodynamics to establish a health-based guidance value (HBGV), also sometimes referred to as a Reference Value (RV) (EFSA SC, 2012). In the food safety area, these include the Acceptable Daily Intake (ADI) for food and feed additives and pesticides and the Tolerable Daily Intake (TDI) for contaminants and chemicals in food contact materials and, for acute effects, the Acute Reference Dose (ARfD). For genotoxic and carcinogenic compounds, the WHO has developed the Margin of Exposure (MOE) approach in which the RP is divided by the human exposure for risk characterisation (WHO, 2005). The MOE has also been applied by EFSA (EFSA, 2005). In this context, the Scientific Committee (SC) of EFSA has considered that MOE values of 10,000 or more, when based on a BMDL for a 10 % extra incidence of tumours in an animal study 'of low concern from a public health point of view'. The SC noted that the magnitude of a MOE only indicates a level of concern and does not quantify risk (EFSA, 2005; 2009a,b; EFSA SC, 2012).

Over the last decade, a number of research programmes such as TOX-21 in the USA and SEURAT in Europe have investigated the use of new methodologies and tools using *in vivo, in vitro* and *in silico* approaches to investigate toxicokinetic and toxicodynamic processes of chemicals at the organism, organ, cellular and molecular level (NRC, 2007). Two key reasons have been driving such efforts, firstly the need to assess thousands of chemicals particularly under the REACH regulation, and secondly the need to find alternatives to animal testing. These methodologies and tools provide the opportunity to move towards a mechanistic understanding of toxicity for hazard assessment (e.g. mode of action/adverse outcome pathways) due to the obligation under the 3Rs principles – reduce, replace, refine the use of animals use in toxicological investigations (3Rs: reduce, replace, refine) (SCHER,

SCENHIR, SCCS, 2012, or due to the ban on testing chemical ingredients in animals-as under the EU Cosmetics Regulation 1223/2009[4]). Recent reviews have discussed such options and include the joint report of the three non-food committees of the European Commission 'New challenges in Risk Assessment', and the report of the US-EPA on 'Next Generation (NexGen) Risk Assessment: incorporation of recent advances in molecular, computational, and systems Biology' (SCHER, SCENHIR, SCCS, 2012; US-EPA, 2013; Goodman et al., 2014).

The purpose of this report is to provide a review of these modern and emerging methodologies and tools to depict and potentially predict toxicokinetic and toxicodynamic processes in the context of human hazard assessment of chemicals. First, the mode of action/adverse outcome pathway concept is discussed in the context of toxicokinetics and toxicodynamics of chemicals and illustrated with new developments and international research efforts (WHO, OECD, Tox-21 programme in the US and the SEURAT programme in Europe). Methodologies and tools to investigate toxicokinetic processes (*in vitro* tools, physiologically-based toxicokinetic, physiologically-based toxicokinetic-toxicodynamic models), key *in silico* tools (QSAR, read-across) and the Threshold for Toxicological Concern (TTC)) for hazard assessment are then presented and illustrated with examples. The report also provides a brief account of the principles behind the key OMICs technologies at the level of DNA/RNA (transcriptomics), proteins (proteomics) and the whole metabolome (metabolomics), with recent examples of applications in human hazard assessment of chemicals. Future perspectives for the potential applications of these methods and tools in the context of prioritisation/ranking of chemicals and the future of chemical risk assessment are also discussed. Finally, recommendations for future work at EFSA are formulated based on consultations of EFSA staff, expert Panels and other international organisations.

2. Mode of action: definitions, recent advances, TOX-21 and SEURAT

2.1. Definitions

Mode of action information for a particular chemical, i.e. the events leading to adverse effects (toxicokinetics and toxicodynamics), is not often available and risk assessors rely often on dose response assessment to translate external dose to a quantitative RP for hazard characterisation in test species. Toxicokinetics (TK) describes the processes leading to the internal concentrations of a chemical or its metabolites(s) through knowledge of absorption (A), distribution (D), metabolism (M) and excretion (E) (ADME). Toxicodynamics (TD) describes the processes that lead to the toxic effects of a chemical or its metabolites(s) once it has reached the organ(s) or tissue(s). Such information has the potential to improve hazard assessment, particularly to assess key uncertainties related to the RP. These uncertainties include the relevance of the test species to the human situation (qualitative and quantitative interspecies differences) and human variability in TK and TD processes. Figure 1 illustrates the level of knowledge for TK and TD processes, for a particular chemical, which can range from very basic (external dose and toxicity) to a full quantitative understanding (external dose to internal dose to target organ dose and metabolism (TK) to specific target organ toxicity (TD).

The definition of **Mode of action (MoA)** has evolved over time and derives from earlier works by the US-EPA (US EPA, 1996, 2005) and the WHO. MoA analyses have been applied to a number of case studies for non-genotoxic and genotoxic chemicals (WHO, 2006a,b). The current WHO definition for MoA is 'a biologically plausible sequence of key events leading to an observed effect supported by robust experimental observations and mechanistic data'. MoA describes key cytological and biochemical events – that is, those that are both measurable and necessary to the observed effect – in a logical framework (Boobis et al., 2006; WHO, 2009; Meek et al., 2014). MoA does not imply full understanding of **mechanism of Action,** which refers to a detailed molecular description of individual biochemical and physiological key events leading to a toxic effect (Boobis et al., 2006; WHO, 2009; EFSA, 2008). In the US, MoA has been used as a term to reference a mechanistic understanding of the impact of a chemical on human health and to reference other terms from epidemiology including

[4] Regulation (EC) No 1223/2009 of the EuropeanParliament and of the Council of 30 November 2009 on cosmetic products. OJ L342, 22.12.2009, p. 59-209.

'disease signature' and 'network perturbations'. In contrast, toxicologists would refer to the same concept using the terms '**toxicity pathway**, MoA, adverse outcome pathway or mechanism of action' as used by the National Research Council (NRC) report, Science and Decisions: Advancing Risk Assessment (2009) (NRC, 2009) and the Nextgen report of the US-EPA (US-EPA, 2013).

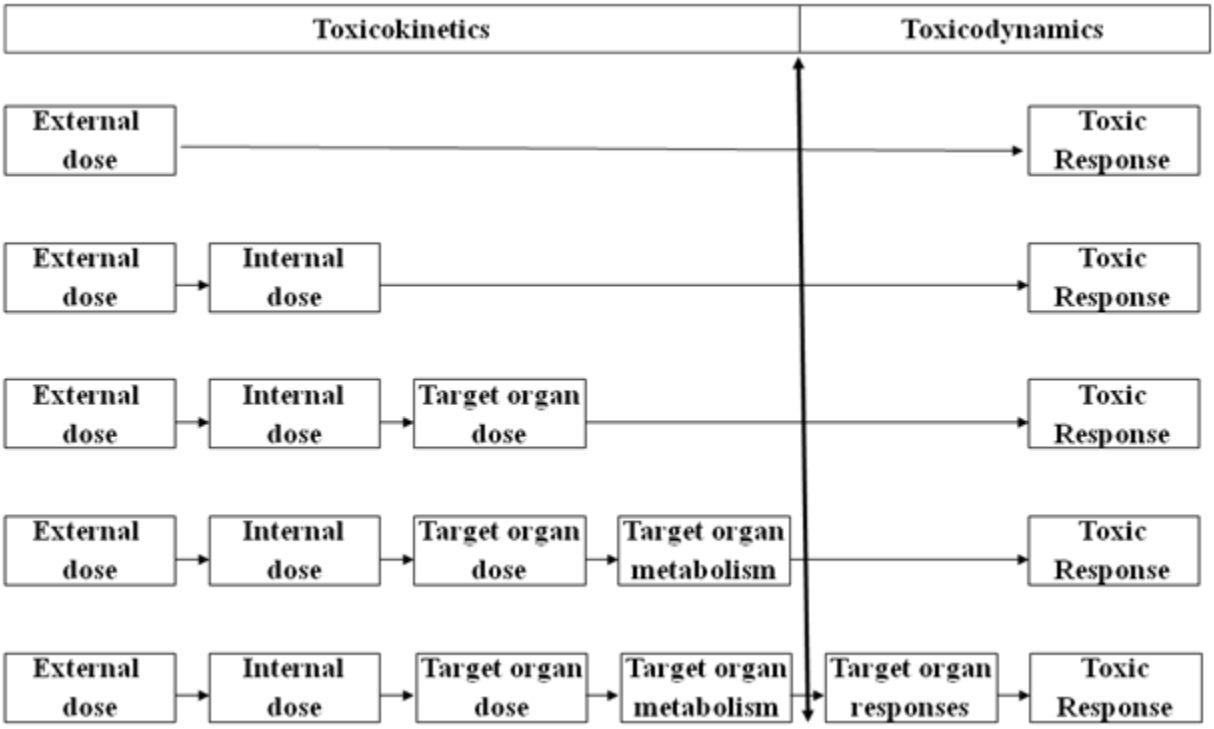

Figure 1: Levels of knowledge of toxicokinetic and toxicodynamic processes

The concept of **Adverse Outcome Pathway (AOP)** emerged from the field of ecotoxicology (Ankley et al., 2010). AOPs are regarded by the WHO as equivalent to MoA in the human health context and an AOP has been defined as 'a sequence of events from the exposure of an individual or population to a chemical substance through a final adverse (toxic) effect at the individual level (from a human health perspective) or population level (from an environmental perspective)' (Ankley et al., 2010; Meek et al., 2014; OECD, 2013). AOPs are characterised by a number of **Intermediate Key Events (IKE)** and key events which individually correspond to 'an empirically observable precursor step that is itself a necessary element of the MoA or is a biologically-based marker for such an element' which are then incorporated into the toxicity pathway and MoA for an adverse effect (Boobis, 2005; US-EPA, 2005; OECD, 2013). Such key events should be definable and make sense from a physiological and biochemical perspective and in a toxicity pathway. Early key events including the **Molecular Initiating Event** (**MIE**) have been defined by the OECD as the 'initial point of chemical-biological interaction within the organism that starts the pathway' (OECD, 2013). The US-EPA has defined AOPs as 'the mechanistic or predictive relationship between initial chemical-biological interactions (i.e. MIE) and subsequent perturbations to cellular functions sufficient to elicit disruptions at higher levels of organisation, culminating in an adverse phenotypic outcome in an individual and population relevant to risk assessment (e.g. disease progression or organ dysfunction in humans)' (Ankley et al., 2010). The authors note that, although commonly used, the AOP term is a misnomer since pathways are not intrinsically adverse or non-adverse but rather pathways which, when perturbed in specific ways, can lead to adverse effects, and the same can be said for the term 'toxicity' pathways (Ankley et al., 2010; US-EPA, 2013).

In addition, a number of authors noted that although there is a rather naive view of the MoA/AOP which has been conceptualised as a series of linear key events; it is recognised that an AOP involves a

number of independent interacting **cellular response networks** 'interconnected pathways composed of the complex biochemical interactions of genes, proteins, and small molecules that maintain normal cellular function, control communication between cells, and allow cells to adapt to changes in their environment'. Such independent networks of key events may play a significant role in their homeostatic regulation and will depend on interspecies differences and human variability, which will need to be considered to develop AOPs (Meek et al., 2014; Vinken et al., 2013). Figure 2 summarises the AOP concept in human and ecological risk assessment in relation to different levels of biological organisation and toxicity pathways from the molecular, cellular, organ, organism level (human risk assessment) through population level (ecological risk assessment) (modified from Ankley et al., 2010, Meek et al., 2014; OECD, 2013).

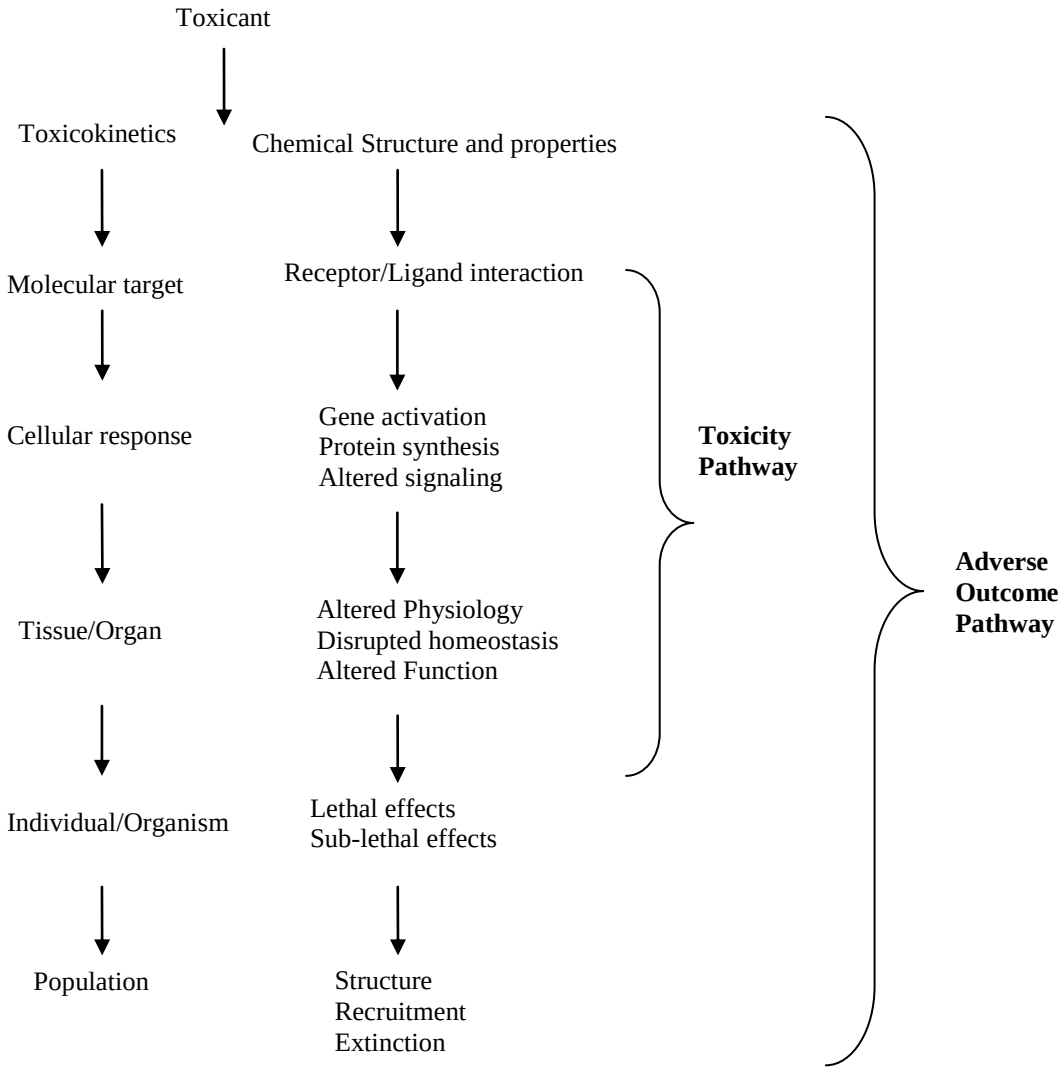

Figure 2: Levels of biological organisation, toxicity pathway and Adverse Outcome pathway

AOPs can have a number of applications including the establishment of (quantitative) structure-activity relationships, the development of novel *in vitro* toxicity screening tests and the elaboration of prioritisation strategies and new testing strategies (Andersen et al., 2012; OECD, 2013; US-EPA, 2013). Such new testing strategies have been designated **Integrated Testing Strategies (ITS)** or equivalently **Integrated Approach on Testing and Assessment (IATA)** and are increasingly used to depict MoA/AOPs as alternatives to animal testing (ECHA, 2013; OECD, 2013). Two recent examples include the use of IATA and ITS strategies at OECD and ECHA, respectively. The OECD

proposed an eight steps IATA strategy for testing skin corrosion and irritation of chemicals which are sequentially addressed: 1) existing human and/or animal data, 2) structure-activity relationships, 3) pH, 4) systemic toxicity via dermal route, 5) use of validated and accepted *in vitro* or *ex vivo* tests for skin corrosion, 6) the use of validated and accepted *in vitro* or *ex vivo* tests for skin irritation, and 7-8) use of a confirmatory *in vivo* rabbit test in a stepwise manner if a negative result is obtained with the *in vitro/ex vivo* skin irritation assays (OECD, 2013). ECHA's guidance on information requirements and Chemical Safety Assessment under the REACH Regulation includes a sequential ITS for skin irritation and/or corrosion. This ITS mostly follows the OECD approach with additional aspects on some elements such as the use of other toxicity data, or weight-of-evidence analysis of existing and relevant data. In addition, validated and accepted *in vitro* tests can be used to identify non-irritants and non-corrosives in order to avoid any *in vivo* tests (ECHA, 2013).

2.2. Recent advances

A number of international efforts have been put together to investigate MoA/AOPs of chemicals for risk assessment purposes. Four key international activities are summarised below: the new developments of the application of the MoA framework by WHO, the OECD guideline on developing and assessing AOPs, the TOX-21 and the SEURAT-1 research initiatives.

2.2.1. New developments in the application of the WHO/IPCs mode of action framework

The WHO/International Programme on Chemical Safety (IPCS), within the context of the working group on MoA, has recently published a thorough account of 'New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis' (Meek et al., 2014). The modified framework has been incorporated within a roadmap which contains a number of feedback loops that consider dose–response relationships, species concordance analysis using a weight of evidence approach and provide options for continuous refinement of fit-for-purpose testing strategies and risk assessment. The authors discussed that the framework can also be used to hypothesising effects resulting from chemical exposure, using information on putative key events in established MoA from appropriate *in vitro* or *in silico* systems and other lines of evidence. Finally, this is also expected to contribute to improving transparency in explicitly addressing weight of evidence considerations in MoA/species concordance analysis based on both conventional data sources and non-standard methods (including *in vitro* and *in silico* methods) (Meek et al., 2014).

A number of cases studies have also illustrated the use of the MoA framework in chemical risk assessment and ITS: 1. Investigation of the relevance of the test species to the human situation, using limonene as an example of qualitative differences in MoA between rat and humans that is not relevant to the human situation; 2. Use of TK and TD data in species concordance analysis; 3. Evaluation of epidemiological data; 4. Guiding more efficient testing strategies (ITS); 5. Prioritising substances for further testing; 6. Categorisation of chemicals; 7. Identifying critical data needs and testing strategies in read-across (Meek et al., 2014).

2.2.2. OECD guideline on developing and assessing adverse outcome pathways

In 2012, the OECD launched a new programme on the development of AOP. In this context, the AOP concept is applied to both human and ecological risk assessment as 'an analytical construct that describes a sequential chain of causally linked events at different levels of biological organisation that lead to an adverse health or ecotoxicological effect'. In a risk assessment context, 'AOPs are the central element of a toxicological knowledge framework being built to support chemical risk assessment based on mechanistic reasoning' (OECD, 2013). This AOP programme of the OECD is coordinated together with the WHO/IPCS work on MoA since both concepts are closely related and addresses three key OECD activities:

- Test guidelines programme to identify new *in vitro* test methods as candidates to become OECD test guidelines, e.g. two recent methods identified in the AOP for protein binding leading to skin sensitisation. These methods investigate gene expression in human

keratinocytes and cell surface markers (CD86) in monocytic cells respectively) (OECD, 2013);

- QSAR Toolbox to identify new methods/profilers for grouping chemicals;
- Hazard Assessment activities to develop ITS for defined hazard endpoints.

The OECD published the 'guidance document on developing and assessing AOP' in 2013, which provided a framework for collecting relevant chemical, biological and toxicological information on the effects of chemicals to develop AOPs. Since this is a new area of activity at OECD, it has been highlighted that the first version of this guidance will be revised in the future as expert groups and member countries get more experience in developing and assessing AOPs. Overall, the guidance provides insights into the type of information required for the identification and the documentation of an AOP, how to present such information and how to undertake the assessment of an AOP regarding its relevance, adequacy and its potential application for regulatory purposes. Further work is ongoing to develop the detailed guidance on the use of AOPs for integrated testing strategies and risk assessment. In addition, an OECD template has been created to allow researchers and stakeholders to develop AOPs and improve the consistency in the AOP development process (OECD, 2013). The AOP development process is described into six phases: 1. Selection of AO and MIE, 2. Study of the relevant physiology underpinning the process, 3. Determination of the IKE, 4. Graphical representation of the AOP/MIE/IKE/AO, 5. Evaluation of evidence supporting the AOP hypothesis, and 6. Reporting of the AOP using the OECD template.

A discussion is provided on the difference between a **qualitative AOP** 'where the key events have been identified but methods for assessing these events have not been identified and/or assessed in sufficient detail to allow for identification of the applicability domains, threshold values and/or the response relationships to other key events' and **quantitative AOP** 'where the methods for assessing the key events have been identified and sufficient data generated to identify the applicability domain, threshold values and/or the response relationships with other key events.' In relation to such AOP knowledge, a number of applications for risk assessment include priority settings for further testing exercise, when not all key events of an AOP are known; hazard identification and classification and labelling, use of partial knowledge of AOP in the OECD QSAR Toolbox. Physiologically-based toxicokinetic models and toxicokinetic information are currently out of the context of AOP and have been recognised as a key gap in the AOP development, therefore they will have to be addressed in the future. As knowledge of AOP increases, the levels of uncertainty and of evidence (e.g. detail, quality, and quantity of information and data) should be reported (OECD, 2013). Since the launching of this programme, a lot of efforts have been made to develop AOPs relevant to human and ecological risk assessment, drafts of which reports are already available (e.g. skin sensitisation) or will be available by the end of 2014. Table 1 illustrates the current AOP development work ongoing within the OECD programme and includes 18 AOP and 3 case studies. In addition, *in vitro* testing strategies are under development as a result of the AOP for skin sensitisation published on the OECD website. A number of AOP tools such as a web-based AOP Knowledge management and an **AOP Wiki/Effectopedia Knowledge Base** (AOP-KB) are also under development by the JRC and the US-EPA under the auspices of the OECD. **AOP-KB** is a knowledge-aggregation and collaboration tool, which facilitates the collection and dissemination of AOP information. Having delivered the 'AOP–KB Wiki' module in early 2014, the project has now entered its next phase: AOP-KB Effectopedia. Whilst the AOP-KB Wiki covers the qualitative aspect of an AOP, the upcoming Effectopedia module will add the quantitative aspect, i.e. the possibility to capture and run (mathematical) models describing the mechanism leading from one Key Event in an AOP to the next. In addition, Effectopedia will add a graphical user interface to the Knowledge Base. More information can be found on the JRC website[5] and on the US-EPA website[6].

[5] See http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing-safety-assessment-chemicals/improved_safety_assessment_chemicals/adverse-outcome-pathways-aop

[6] See: http://www.epa.gov/ord/priorities/docs/aop-wiki.pdf

Table 1: Current Development of Adverse Outcome Pathways at the OECD

Development of Adverse Outcome Pathways
Skin Sensitisation Initiated by Covalent Binding to Proteins
Nonpolar Narcosis
Acetylcholinesterase Inhibition
Five Cell Signalling Pathways Associated with Cell Proliferation and Differentiation Conserved Across Species
Mitochondrial Toxicity
Embryonic Vascular Disruption and Developmental Defects
Sustained Activation of the Avian Aryl Hydrocarbon Receptor
Mutagenic Modes of Action for Cancer
Upregulation of Thyroid Hormone Catabolism via Activation of Hepatic Nuclear Receptors, and Subsequent Adverse Neurodevelopmental Outcomes in Mammals
Xenobiotic Induced Inhibition of Thyroperoxidase and Depressed Thyroid Hormone Synthesis and Subsequent Adverse Neurodevelopmental Outcomes in Mammals
Heritable Germ Cell-Derived Disease (3 AOPs)
1. Alkylation of DNA in male pre-meiotic germ cells causing inherited mutation;
2. Chemical interaction with tubulin in oocytes leading to inherited aneuploidy
3. Bulky DNA adducts in male pre-meiotic germ cells causing point mutation leading to inherited DNA sequence mutation
Linking Aromatase Inhibition, Androgen Receptor Agonism, Oestrogen Receptor Antagonism, or Steroidogenesis Inhibition, to Impaired Reproduction in Small Repeat-Spawning Fish Species
Neurotoxicant-induced Neuroinflammation: a converging key event in an AOP
From protein alkylation to liver fibrosis
Neurotoxicity induced by GABAA receptor inhibition
Haematotoxicity due to Nitroaromatics and N-hydroxyl anilines
CAR and PPARα-mediated pathways to non-genotoxic rodent liver cancer
CAR and PXR-mediated pathways to rodent liver hyperplasia
Development of Case Studies
Case Studies Using Aquatic Organisms
Hepatotoxicity due to 2,4,6-trinitrotoluene
Energy Metabolism by 2,6-Dinitroluene

2.3. Recent research programmes and initiative in the US and Europe:TOX-21 and SEURAT

2.3.1. The National Toxicology Program and TOX-21

In 2004, the National Toxicology Program (NTP) published 'A National Toxicology Program for the 21st Century: A Roadmap to Achieve the NTP Vision'. The NTP intention was to transform toxicity testing from an observational science based on whole animal testing, to a target-specific and mechanistic one, based on identifying mechanisms of cellular toxicity. In this report, the NTP argued that this was necessary given that: (i) traditional *in vivo* toxicity testing methods are resource-intensive and time-consuming; (ii) such methods are unable to assess the vast backlog of untested chemicals already present in the environment to which humans are exposed; (iii) emerging methods and new technologies were helping revolutionise other fields of biology (Attene-Ramos et al., 2013; NTP, 2004). In 2005, the US-EPA asked the National Research Council (NRC) to develop a long-term plan capable of revolutionising toxicity testing (Attene-Ramos et al., 2013). The NRC's answer was to produce the report 2007 'Toxicity Testing in the 21st Century: A Vision and a Strategy'. It provided a long-term strategy on how toxicity testing should be transformed by the consideration of new methods in molecular and systems biology, computational toxicology, and bioinformatics (NRC 2007; Tice 2011; Attene-Ramos et al., 2013; Tice et al., 2013). Key recommendations from the NRC report (2007) included:

- Proposal to use chemical profiling strategies to measure biological changes induced by chemicals using automated **High-throughput screening (HTS) assays** defined as 'efficiently designed experiments that can be automated and rapidly performed to measure the effect of substances on a biologic process of interest'. HTS assays can evaluate thousands of chemicals over wide concentration ranges to identify chemical mechanisms on gene, pathway, and cell function (US-EPA, 2013).
- Assay parameters should be used as toxicological endpoints as the first steps in identifying MoA/AOP.
- Findings from initial *in vitro* experiments can be used to prioritise chemicals for more in-depth evaluation and for the development of toxicological models.
- Where possible, the assays should be based on human derived cells, cell lines or cellular components to avoid species dependent differences in response to the chemicals.

Since the 2007 report, a number of U.S. Federal Agencies have collectively devised an implementation strategy in response to the NRC report. The collaborating agencies include the National Institute of Environmental Health Sciences/National Toxicology Program (NTP), the US-EPA's National Centre for Computational Toxicology, the National Human Genome Research Institute/National Institutes of Health Chemical Genomics Centre and the Food and Drug Administration (FDA).

This collaborative program is informally known as **Tox-21** and is conducted by the NIH Chemical Genomics Center (NCGC) and the US-EPA via the **ToxCast** program (Attene-Ramos et al., 2013; Tice et al., 2013). ToxCast has been divided in two screening phases. Phase I of ToxCast was completed in 2009 and has screened around 320 compounds in 550 biochemical and cell-based assays. Compounds were mainly pesticides which had already been studied extensively *in vivo* so that toxicity endpoint(s) were already characterised (i.e. target organ, reproductive, developmental…) (Tice et al., 2013). The assays evaluated known toxicity pathways e.g. cytotoxicity, apoptosis induction, DNA damage, perturbation of cell signalling pathways, inflammatory response induction, nuclear receptor modulation, oestrogen receptor, enzyme inhibition and membrane transport inhibition. Of particular current interest are HTS assays that have been evaluating biological signal transduction pathways such as Wingless-related integration site (Wnt), Sonic Hedgehog (SHH), Delta-notch, Tumour Growth Factor-beta (TGFβ), receptor tyrosine kinase (RTK), retinoid and endocrine pathways: oestrogen, thyroid, adrenal, and androgen. Such HTS, developed under phase I of Toxcast, have provided the possibility to classify chemicals according to their molecular interactions with cellular targets. Such classification will potentially give the opportunity, in the near future, to separate and prioritise chemicals that affect a specific MoA affect via a particular signal transduction pathway versus chemicals that have non-specific or multiple MoAs (Shukla et al., 2010; Tice, 2011; Tice et al., 2013; Wetmore et al., 2013).

Phase II of ToxCast has expanded on its chemical library including 667 chemicals selected as chemicals for which there is potential human exposure or which represent potential ecological hazards, and include industrial and consumer products, food additives, 'green' products, cosmetic-related chemicals, and failed pharmaceutical drugs (Sipes et al., 2013). These chemicals lack the traditional toxicity data of phase I but human clinical data and other toxicological studies are available to assess and test the performance of predictive models developed in phase I (Truong et al., 2014). Phase II focused primarily on the detection of chemicals that induce one or more stress response pathway with the rationale that such stress responses would be markers of potential *in vivo* toxicity. These stress response pathways included: antioxidant response, cytotoxicity, DNA damage response, heat shock, and mitochondrial damage (Tice et al., 2013).

Strengths and shortcomings of the Tox-21 have been discussed (Tice, 2011; Attene-Ramos et al., 2013; Tice et al., 2013). The **strengths** comprise the coverage of thousands of compounds which can be screened in a single experiment as a cost-effective way to prioritise chemicals based on their

molecular interactions with their target in the cell. In other words, it investigates MoA while minimising animal testing and the assays are based on human cells and cellular components and not test species. The s**hortcomings** of the assays include, amongst others, the fact that they cannot yet: 1. predict disease-associated pathways and diseases. 2. assess metabolism and interactions between different cell types. 3 replicate tissue-level cellular interactions. 4. Replicate chronic exposure as short term in *in vitro* assays. In addition, no proper methodology for the prioritisation of chemicals has been developed; there are limits as to what compounds can be screened (e.g. volatile compounds and gases). Finally, authors also recognised that Tox-21 program is still in its infancy, it is fundamentally a research and development program and may take decades to achieve the goals originally set by the NRC (Tice et al., 2013). Examples of results generated from Tox-21 are illustrated throughout this report.

2.3.2. The European Commission and the SEURAT initiative

The Safety Evaluation Ultimately Replacing Animal Testing (**SEURAT**) initiative was initiated in 2008 by the Health Directorate of the European Commission's Directorate General for Research and Innovation (DG RTD). The overall aim of SEURAT was to devise and implement a comprehensive EU research programme that will drive a major overhaul in the chemical safety assessment paradigm, ensuring the greatest protection of human health without animal testing. The first execution phase, SEURAT-1, was launched in January 2011. The overall emphasis of SEURAT-1 was on the identification and elucidation of MoAs related to repeated dose systemic toxicity in humans, and to develop MoA-based systems of experimental and computational methods to be applied in human safety assessment. SEURAT-1 comprises a cluster of five complementary research projects:

- SCR&Tox: 'Stem Cells for Relevant Efficient Extended and Normalized Toxicology'

- HeMiBio: 'Hepatic Microfluidic Bioreactor'

- DETECTIVE: 'Detection of endpoints and biomarkers of repeated dose toxicity using *in vitro* systems'

- COSMOS: 'Integrated *In Silico* Models for the Prediction of Human Repeated Dose Toxicity of COSMetics to Optimise Safety'

- NOTOX: 'Predicting long-term toxic effects using computer models based onsystems characterization of organotypic cultures'

The research projects are supported by a central data and knowledge management project (ToxBank) and a coordination action (COACH). More than 70 research partners participated in the SEURAT-1 projects.

The SEURAT strategy has adopted a MoA approach to investigate how any substance may adversely affect human health using ITS. Such knowledge provides a basis to develop complimentary theoretical, computational and experimental (*in vitro*) models and assays for the prediction of quantitative points of departure (POD) for risk assessment. The SEURAT-1 framework builds on the idea that key molecular or biological events are common between different MoAs so that it is the particular chain of causally linked events that makes a MoA unique. It is also recognised that a substance may be 'promiscuous' and may have multiple MoAs.

Examples of the models that are being developed under SEURAT-1, include 3D tissue models that are produced either experimentally using bioreactor systems, or virtually, by using computational biology approaches to allow the qualitative association of a chemical with one or more MoAs and dose-response assessment. In addition, differentiation and characterisation of human pluripotent stems cells are made for large scale production of cell models to be used in high throughput *in vitro* testing. Innovative biomarkers and OMICs readouts are further developed. Complementing the cell and tissue models, computational chemistry, quantitative structure-activity relationships (QSARs), and chemoinformatics tools provide the means to understand and predict key biochemical events such as protein binding and metabolic transformation.

In terms of research strategy, SEURAT-1 investigators first decided on which MoA was of relevance to their particular study or test system, and then selected the associated reference chemicals. Thus MoA was brought to the forefront, with the design, optimisation, and evaluation of *in vitro* test systems being driven by the aim to capture one or more specific MoAs with high sensitivity and selectivity. As a consequence, the specifications of the biological model, the exposure protocol, the biomarkers to be measured, and the reference chemicals to be used as positive controls, all depend on the MoA chosen. The selection of MoAs was performed in the context of OECD AOP activities described above. The SEURAT-1 research initiative has resulted in the emergence of a generic AOP development process in six steps in line with the OECD guidance and template for AOP development.

An important factor to consider is the impact of the TK on the TD of the chemical under investigation and it has been recognised both within the SEURAT-1 and the US ToxCast that this is a key factor to understand the potential use of *in vitro* tools in risk assessment. Indeed, TK can be very different in an *in vitro* system when compared to an *in vivo* one because of a number of factors such as the chemical accumulation in a target organ due to its persistence, and the inhibition of a detoxification enzyme or the induction of a bioactivation enzyme. As a consequence, a central issue for the SEURAT-1 research initiative was to further relate treatment concentrations used in the various *in vitro* test systems to *in vivo* serum and target organ concentrations, and vice versa. The SEURAT-1 Research Initiative will deliver many important computational and experimental tools, and related know how that will be critical components in predictive toxicology approaches. To demonstrate the potential of these tools and how they can be assembled in an integrated manner, the cluster will undertake a proof-of-concept exercise to demonstrate how a MoA-based testing strategy can be used to predict aspects of repeated dose target organ toxicity. Two recent examples of AOP development within the context of SEURAT include drug-mediated cholestatic liver injury and skin sensitisation (OECD, 2013; Vinken et al., 2013). Detailed descriptions of the SEURAT-1 progress can be found in the annual reports (www.seurat-1.org).

3. Physiologically-based models and *in silico* tools

The aim of this section is to describe basic tools to investigate TK processes,to then introduce the principles underpinning the building of physiologically-based models to address both TK (physiologically-based TK models (PB-TK)) and TD processes (PB-TK-TD). *In silico* tools such as quantitative structure activity relationships are then discussed together with read-across methods and the threshold of toxicological concern (TTC). Examples are also given to illustrate the applications in relation to human hazard assessment of chemicals.

3.1. Investigating toxicokinetics

In order to build physiologically-based models, chemical specific TK data regarding absorption, distribution in the body, metabolism and excretion (ADME) need to be integrated. Key players in ADME processes include phase I and phase II enzymes as well as efflux transporters (phase 0 and phase III). Phase I enzymes catalyse key reactions such as oxidation, reduction, dealkylation hydrolysis and include the cytochrome P-450 (CYP) superfamily of enzymes and other enzymes (e.g. alcohol dehydrogenase, epoxide hydroxylase, esterases). Phase II enzymes catalyse conjugation reaction and include key enzymes such as UDP-glucuronyltransferases, sulphotransferases glutathione-s-transferases and methyl-transferases. Efflux transporters belong to two main clusters of families: the solute carrier (SLC) families and the adenosine triphosphate (ATP) binding cassette (ABC) carriers. SLC transporters include the human organic anion transporting polypeptides (OATPs) and human organic cation transporters (OCTs) and are often denominated as phase 0 uptake transporters. ABC tranporters are often called efflux pumps or phase III and include examples of importance such as P-glycoproteins and multidrug resistance proteins (MRPs) (Hillgren et al., 2013; Doring and Peztinger, 2014).

It is worth noting that TK information is only part of mandatory *in vivo* animal testing in some but not all legislative frameworks. The current updated OECD Test Guideline 417, mainly related to absorption and biotransformation, indicates that *in vitro* testing using human cells, can be a valid

supplemental TK information and thus may substantially reduce *in vivo* animal testing. A number of *in vitro* models have been developed, mostly in the pharmaceutical industry, to measure ADME processes. Additionally, a number of *in silico* models have been developed to predict ADME parameters with chemical structural features (e.g. Quantitative Structure-Activity Relationships) (see Section 3.3). These *in vitro* models have still received little attention in hazard assessment of chemicals for the food safety area and some are highlighted below for the investigating of TK processes (Blaauboer, 2010; Punt et al., 2011; Coecke et al., 2013).

3.1.1. Absorption

Absorption has been defined by the WHO as: 'the process by which a substance is transferred from the site of administration into the circulation. For chemicals in food, absorption usually refers to passage across the gut wall into the circulation, although for some chemicals, uptake may be only as far as the epithelium of the gastrointestinal tract' (WHO, 2009). In food safety, the oral route is the most relevant route of exposure and after oral ingestion, the chemical passes sequentially from the gastrointestinal lumen, through the gut wall, to the liver and becomes bioavailable after it has enters the systemic circulation. **Oral bioavailability** has been defined as 'the product of three fractions: 1. Fraction of dose absorbed, 2. Fraction of absorbed dose passing through the gut into the hepatic portal blood unmetabolised, and 3. Fraction of dose not metabolised in the liver' (Thelen and Dressman, 2009). In other words, oral bioavailability is a function of absorption and first-pass elimination of the chemical in the gastrointestinal tract and liver. Low oral bioavailability has been attributed to compounds, which may limit solubility, dissolution, permeability, affinity for efflux transporters, metabolism in the gut lumen, in the intestine and/or in the liver (Bueters et al., 2013). Transporters have been shown to be of key importance in the absorption, bioavailability and excretion of chemicals. They are mostly expressed in liver, but are also present in extra-hepatic tissues (e.g. kidney, adrenal gland and lung). A number of *in vitro* tools are available to investigate absorption and bioavailability of chemicals as discussed in Coecke et al. (2013). Caco-2 cells are the most widely *in vitro* cell models to estimate intestinal absorption. This cell model has been developed from a human colon adenocarcinoma in culture and it is grown to form a polarised monolayer, which displays similar morphological and functional characteristics as intestinal enterocytes. The cultured cells form tight junctions and express phase I and phase II enzymes and phase 0 and phase III transporters (Alqahtani et al., 2013). However, the predictability of Caco-2 cells has strong limitations particularly to predict absorption for highly lipophilic compounds, substances with low-to-moderate absorption rates, substances that are substrates for transporters and/or substances which undergo first pass metabolism (Turco et al., 2011).

In vitro and artificial membrane methods have also been developed to measure and predict absorption and bioavailability for different routes of exposure (oral, dermal, inhalation…) and to mimick physiologically-based absorption barrier (Faller et al., 2008; Lafond et al., 2011; Mitra et al., 2011). Affinity of chemicals to transporters in the gut and liver is a growing issue in TK and hazard assessment, it is currently tested routinely in the pharmaceutical area to predict biovailability and potential for drug interactions but it is not a routine assay for food regulated substances and contaminants in food. As a recent example, Meyer et al. (2013) tested the affinity to P-Gp for 47 drugs of abuse using Human P-gp (hP-gp) membranes prepared from baculovirus-infected insect cells and control membranes. The affinity for hP-gp was measured and modelled using classical Michealis Menten constants (Vmax and Km) and provides a tool for measuring intestinal transport of xenobiotics.

A number of software and models are also available for the simulation and modelling of intestinal absorption and metabolism (mostly for drugs) including compartmental models such as CAT (compartmental absorption transit) models, ACAT (Advanced compartmental absorption transit model model) (GastroPlus) and the ADAM (Advanced Dissolution model) (SimCYP) as well as dispersion models (Alqahtani et al., 2013). Recently, mechanistic physiologically based absorption models have been developed such as GastrointestinalSim (GI-Sim). The model is a compartmental gastrointestinal absorption and transit model combined with algorithms that describe permeability, dissolution rate,

effects of salts, partitioning into micelles, particle and micelle drifting in the aqueous boundary layer, particle growth and amorphous or crystalline precipitation (Sjogren et al., 2013).

The integration of *in vitro* techniques and models as input parameters in PB-TK to predict absorption and bioavailability is still in its infancy and, in order to use such models routinely, further work is still needed. For example, over the last ten years, a number of *in vitro* alternative assays have been developed to assess dermal absorption and resulted in the OECD test guidelines 428. However, their uses are still limited because of the lack of the biotransformation capacity of such *in vitro* systems. Again, a critical analysis of these tools in relation to their relative capacity to predict human absorption rates *in vivo* as input for PBTK models, is needed (Van der Merwe et al., 2006; Nossol et al., 2011).

3.1.2. Distribution

Once a chemical enters the systemic circulation, it is distributed into interstitial and intracellular fluids. Distribution of the chemical in each organ will depend on key factors such as vascular permeability, regional blood flow, cardiac output and perfusion rate of the tissues and the chemical‘s ability to bind tissue and plasma proteins, lipophilicity and pH partition. Major determinants include the volume of distribution, protein binding and clearance at steady state. A recent approach to determine the volume of distribution for lipophilic compounds uses *in vitro*, physicochemical data and a simplified tissue-composition-based model to estimate tissue–plasma ratio (Poulin and Haddad, 2013).

Protein binding of chemicals is a critical part of TK information particularly to develop physiologically-based models since affinity for plasma proteins varies tremendously between compounds and affects the free concentration of the compound and the whole TK (Bow et al., 2006). The blood to plasma ratio is a critical variable to convert tissue/plasma partitions to tissue/blood, or fraction unbound in the plasma to the fraction unbound in the blood (the free concentration) (Yoon et al., 2012). Using *in vitro* data to gather information on distribution, *in vitro* to *in vivo* extrapolation (IVIVE) methods have been developed. In this context, the *in vivo* plasma concentration of a chemical needs to be linked to a target-tissue response similar to the cellular response in the *in vitro* system and to do so, determination of the *in vitro* and *in vivo* free fractions is critical (Teeguarden and Barton, 2004; Yoon et al., 2012). Experimental *in vitro* systems to measure free concentrations include equilibrium dialysis, ultrafiltration, and ultracentrifugation Authors have recognised that a number of factors affect the free fraction/apparent partitioning of a compound and can complicate data interpretation. These factors include blood-plasma ratio, partitioning due to lipophilicity, plasma and tissue binding active transport (phase 0 and phase III), metabolism (e.g. clearance processes) (Yoon et al., 2012).

3.1.3. Metabolism

In vitro characterisation of metabolism aims to identify key metabolic routes for a compound and to estimate clearance as a surrogate for *in vivo* metabolism. It also provides a basis to develop *in vitro* to *in vivo* extrapolations (IVIVE) and to predict potential interactions between compounds, which gives a basis to take into account TK interactions for hazard characterisation of multiple compounds (chemical mixtures) (Coecke et al., 2006; Chen et al., 2011; Jayaraman et al., 2011; Kumar et al., 2011; Abass et al., 2013; EFSA, 2013). These comprise:

Hepatic xenobiotic metabolism

A number of *in vitro* methods to investigate hepatic xenobiotic metabolism have been reviewed recently by Abass et al. (2013) and Yoon et al. (2012):

Hepatic microsomes are the most frequently used subcellular fractions for drug metabolism studies of new drug candidates and consist of vesicles of the hepatocyte endoplasmic reticulum prepared by standard differential ultracentrifugation.

Hepatocytes including cryo-preserved hepatocytes are used to evaluate the metabolic stability of the compounds and to identify metabolising enzymes as well as enzyme inhibition. It is recognised that for interspecies comparison, hepatocytes in suspensions are more reliable than hepatocytes in culture. This is mainly due to the potential damage caused by cytotoxic substances produced by the cells in and the variations in the expression levels of phase I and phase II enzymes in cultures (Chen et al., 2011; Vasdev et al., 2011; Abass et al., 2013).

Cell line expressing CYP cDNA (single-enzyme systems) as recombinant xenobiotic metabolising enzymes, are well established and commercially available. The enzymes can be expressed in bacterial, yeast, and mammalian cell lines and human lymphoblast or baculovirus- infected insect cells (Abass et al., 2013).

Immortalised cell lines have been isolated from primary tumours of the liver parenchyma, developed after chronic hepatitis or cirrhosis such as HepaRG derived from a hepatocellular carcinoma. In contrast to the HepG2 cell line, HepaRG cells express a large panel of liver-specific genes including several human CYP isoforms (e.g. CYP1A2, CYP2B6, CYP2C9, CYP2E1 and CYP3A4), phase II enzymes (e.g. glutahione-s-transferases and UDP-glucuronyltransferases), transporters (e.g. P-gp) and nuclear transcription factors over six weeks in culture. An important limitation of HepaRG cells is that they have been derived from a single donor and do not provide measures of inter-individual variability for each of CYP isoforms (Abass et al., 2013).

Liver slices constitute a powerful tool to study biotransformation *in vitro*, even though they begun to slowly fall out of use in the prediction of TK and drug metabolism for a number of reasons. These include issues associated with drug movement into and out of the slices, lower enzyme activities and the increased use of hepatocytes to study similar reactions.

Using these *in vitro* assays, classical Michaelis Menten kinetic parameters are determined such as the maximum rate of catalysis (Vmax) and the Michaelis Menten constant (Km) as the substrate concentration that gives half maximal velocity of an enzymatic reaction. The ratio between Vmax and Km is then calculated to give the intrinsic clearance (CLi). CLi can then be scaled up to the whole liver to determine the *in vivo* hepatic clearance of a particular compound using the milligrams of microsomal protein per gram of liver, liver blood flow, and the size of the liver (Abass et al., 2013; Yoon et al., 2012). Examples of IVIVE for hepatic clearance are numerous in the literature. Two recent examples recently illustrated the method for diuron (phenylurea herbicide), and carbosulfan (carbamate insecticide) (Abass et al., 2013). For enzyme inhibition, Rostkowski et al. (2013) proposed a tool for the prediction of which human CYP (amongst 1A2, 2C9, 2C19, 2D6 and 3A4), a given molecule is likely to inhibit. Finally, *in vitro* biotransformation systems and modern modelling techniques can also nowadays account for inter-individual variability in metabolism (genetic/epigenetic) using cells from different individuals (Chaudhry et al., 2010). In a recent study, the impact of different CYP2C8 genotypes on rosiglitazone plasma levels and possible drug-drug interactions was investigated using *in vitro* metabolism data obtained from human wild-type and variant enzymes of CYP2C8 which were then converted to whole organ metabolic CLi. IVIVE extrapolation of the *in vitro* data to *in vivo* and incorporation into a PB-PK model resulted in reasonable predicted values which were within a 1.2-1.7-fold range of the observed values (Yeo et al., 2013).

Intestinal xenobiotic metabolism

As highlighted previously (see Section 3.1.1), xenobiotic metabolism in the intestine is also of considerable relevance in metabolism particularly in food safety since the main route of exposure is the oral route. However, intestinal metabolism has long been underestimated as a consequence of the technical difficulty to dissociate the role of the intestine from that of the liver in *in vivo* experiments and of the lack of *in vitro* models sufficiently viable and fully representative of the physiology and anatomy of the intestine. Recently the precision-cut slice model (PCIS), widely used for the liver and kidney, has been adapted for the small and large intestine and its applicability to investigate intestinal metabolism has been reviewed recently (Groothuis et al., 2013). PCIS can be prepared from animal and human tissues from all regions of the intestine and allow investigation of species differences and regional gradients of activities of metabolising enzymes. PCIS are viable for 8-24 h of incubation and show high activity of xenobiotic metabolising enzymes, and are good surrogates of *in vivo* activity. They have been successfully used to study drug-drug interactions such as induction, inhibition and regulation of xenobiotic metabolising enzymes, transporters and nuclear factors. Their uses as models for chemical-intestinal metabolism and chemical-induced intestinal toxicity are still limited but these methods appear to be promising to contribute to the reduction and replacement of animal experiments (Groothuis et al., 2013).

3.1.4. Excretion

Excretion of chemicals is the last step in TK processes and may involve different organs (kidney, liver, lung, skin). The *in vitro* estimation of excretion has not been investigated much and no *in vitro* method is currently available (Coecke et al., 2013), however, some progress has been made with pharmaceuticals (Kusuhara and Sugiyama, 2009). A main difficulty to develop models in organs, such as the kidney, relates to the relative spatial complexity of its tubular transport systems compared to the more homogenous architecture of the liver. In addition, a multitude of transporter proteins such as uptake pumps (phase 0) and efflux pumps (phase III) have been identified and are involved in both absorption and excretion which complicates the picture. However, Yang et al. (2010) argued that as knowledge on such transporters increases, it will be feasible to develop assays to identify whether a compound is a substrate for a particular transporter in the kidney (or intestine or liver) or not. In terms of PB-TK models, Tonnellier et al. (2012) noted that default assumptions are often made regarding excretion of chemicals and research efforts are needed to test the accuracy of such assumptions and how they affect the overall sensitivity of the models. For example, investigating the interplay between transporters and kidney excretion may provide indication of the likelihood that a compound's renal clearance might deviate from expectations based on glomerular filtration alone.

3.2. Physiologically-based toxicokinetic models and application in hazard assessment

3.2.1. Principles

The WHO has defined **TK models** as **'**mathematical descriptions simulating the relationship between external exposure level and chemical concentration in biological matrices over time'. TK models take into account ADME of the administered chemical and its metabolites (WHO, 2010). The integration of physiological parameters into TK models results in a physiologically based -TK models which is defined as 'a model that estimates the dose to target tissue by taking into account the rate of absorption into the body, distribution and storage in tissues, metabolism and excretion on the basis of interplay among critical physiological, physicochemical and biochemical determinants' (WHO, 2010).

In a reverse way (reverse dosimetry), PB-TK models can be used to estimate the external dose or exposure concentration needed to achieve given target organ concentrations, measured for example using biomarkers in humans (Verner et al., 2009).

PB-TK models are based on a compartmental approach that separates the organism's body into a series of biologically relevant anatomical compartments of defined volumes. The number of compartments varies from model to model, i.e. one compartment model to multi-compartment models

depending on data quality, tissue compartments of interest (i.e. site of pharmacological or toxicological activity), purpose of the model, and the physico-chemical properties and behaviour of the chemical in the organism (lipophilic/hydrophilic). All compartments are in general connected in anatomical order based on the blood circulatory system to form an integrated model. The transfer of chemicals between compartments is thus governed by blood flow rates and tissue solubility (partition coefficients). Each compartment can also have several sub-compartments consisting of a vascular section, an interstitial space, and a cellular space (Gerlowski and Jain, 1983; EFSA, 2013). Recently, Rowland et al. (2013) described the PB-TK major components as system-specific properties, chemical/drug properties, and the structural model. System-specific properties include organ mass or volume, blood flow, and tissue composition. Chemical/drug properties include tissue affinity, plasma-protein binding affinity, membrane permeability, enzymatic stability, and transporter activities. The structural model comprises the anatomical arrangement of the tissues and organs of the body, linked by perfusing blood. Unlike empirical models, dictated by the observed chemical/drug data, a PB-TK structural model is independent of the chemical/drug and is the same for all mammalian species, although the degree of complexity often varies with the intended application.

PB-TK models are increasingly developed from IVIVE or quantitative IVIVE (QIVIVE). The main difficulty is the measurement of the free and internal cell concentrations in these *in vitro* systems mainly determined by either abiotic processes (i.e. chemical stability of the compound over time, adsorption to the plastic devices, binding with the medium components) or cellular processes (mechanism of transport across the membranes, biotransformation, bioaccumulation). Consequently knowledge of the expression/activity of phase I and phase II enzymes and transporters in the *in vitro* model in use is critical. However, it has been acknowledged that many cell cultures have often totally or partially lost their metabolic and transport capacities resulting in an unbalanced situation compared to the *in vivo* situation (Yoon et al., 2012). This conclusion has been highlighted by Tice et al. (2013) for the Toxcast assays of the TOX21 program: a HTS assay to measure the free concentration of a compound *in vitro* is not yet available and xenobiotic metabolism is lacking in virtually all HTS assays. Due to the above factors, the uncertainty about the actual level of exposure of cells *in vitro* is even greater after repeated treatments than after single dosing *in vitro*. Parallel to PB-TK models, integrated mass balance/fate, cell population of *in vitro* experiments should be developed, to understand the TK behaviour of a toxicant at the cellular level. Such peculiarities of *in vitro* TK may lead to large errors in the interpretation and use of the data generated, if ignored. Hence, if such IVIVE and QIVIVE are considered to replace animal testing in the future, concentration measurements of the parent compound and/or of the metabolites in *in vitro* systems have to be considered as critical parts of the experimental design (Yoon et al., 2012).

In practice, two approaches for PB-TK modelling are used namely a bottom up and a top down, approach. The ‘bottom-up’ approach includes each organ and tissue of the body as a distinct entity so that the interactions of a chemical/drug with all components of the body are integrated to allow for mechanistic insights into the global behaviour of the system to make valid extrapolations. The ‘top down approach’ uses simplified models for which the tissues are combined together (‘lumping tissues’). This approach is used increasingly to estimate TK parameters for complex models coming from experimental data a. A key issue for such top down approach is the need to validate methods to reduce the complexity of models while preserving global body characteristics (e.g. cardiac output and body weight, criteria on the ‘lumping of tissues’ based on their kinetic features, mass balance of the chemical/drug).

Recently, a number of generic PB-TK models, applicable to a large number of substances, coupled to parameter databases and QSAR modules have been developed to model inter-individual variability in the ADME processes for pharmaceuticals, environmental contaminants and pesticides, in test species, humans as well as farm animals. Different methodologies are available to build these models ranging from one-compartment model to Markov chain Monte Carlo methods and multi-level (hierarchical) population models used for Bayesian calibration of the models (Bois et al., 2010). Compared to previous PB-TK models, current models usually contain more compartments and even more complex exposure equations. For example, recent models have included transporters from the gastrointestinal

tract (e.g. P-Gp, OATP) and intestinal metabolism together with hepatic metabolism, to predict bioavailability, first pass metabolism and to reflect human physiology and variability more accurately (Pang et al., 2010). It has been recognised that building such complex models require intensive resources and sophisticated software tools as well as detailed knowledge about the chemical deposition in the body and physiological input parameters for animal species and in different subgroups of the human population.

Over the last decade, a number of generic software platforms have been designed to support the TK modelling of pharmaceuticals with a focus on the oral and intravenous routes and metabolism by CYP isoforms and phase II conjugation enzymes. These platforms are available to support generic PB-PK modelling mostly for pharmaceuticals using *in vitro* metabolism data and IVIVE (e.g. Simcyp platform (Rostami-Hodjegan and Tucker, 2007; Jamei et al., 2009) and industrial compounds) using software such as IndusChemFate (Jongeneelen and Berge, 2011) and MEGen (Loizou and Hogg, 2011). Software packages to develop these models range from user-friendly excel spread sheet interfaces to software using complex algorithms such as MC-Sim, PK-Sim, Berkeley Madonna, acslX, MATLAB or PK-BUGS using full Bayesian inference (Loizou and Hoggs, 2011).

Generally, PBTK model are very useful tools in hazard assessment but are often used in high-tier assessment (tier 3) since they require detailed knowledge and high level of expertise and they need to be validated (EFSA, 2013). When available and appropriate, the application and use of these models have been recommended by regulatory authorities around the world (ATSDR 2004; US-EPA, 2007; EFSA, 2008a; WHO, 2010; Meek et al., 2011; OECD, 2011). However, because of the specialised expertise required, these models have not yet been implemented routinely in human hazard assessment and WHO (2010) has highlighted that guidance needs to be developed to pursue common principles and harmonised approaches in relation to those models.

Authors have discussed some key criteria that a generic PB-TK modelling would need to have to be a useful practical tool: (i) user friendly, open access; (ii) database for physiological parameters; (iii) inhalation, dermal, and oral exposure routes and (iv) capability to model multiple parallel metabolic pathways (Yoon et al., 2012). The US-EPA has discussed a number of criteria for the acceptance of PB-TK models in risk assessment: (1) the model represents the species and life stage of relevance for the specific risk assessment, (2) the model has been evaluated and peer-reviewed for transparency, adequacy of its structure and parameters, and (3) the model provides adequate simulations of the concentration of the toxic moiety (parent compound or metabolite(s)) in the target organ (or a surrogate compartment of the body), relevant exposure route(s) and relevant time-course for which the chemical would be present in that target organ/surrogate compartment (US-EPA, 2006). In terms of validation, 4 key aspects to be reviewed have been identified 1. model purpose and structure, 2. mathematical representation, 3. calibration of the parameter estimations, and 4. computer implementation of the model. In the future, PB-TK models may be increasingly used, on a case by case basis depending on the purpose of the risk assessment and data availability, as more generic predictive tools become available (Conolly et al., 2005; US-EPA, 2007; EFSA, 2008; Dorne et al., 2012).

3.2.2. Application in hazard assessment

Generic applications of PB-TK in human hazard assessment are numerous and include interspecies differences, route-to-route extrapolation, analysis of human variability in TK to integrate differences between subgroups from varying exposure condition, chemical mixtures, high-to-low dose extrapolation (WHO, 2010). Historically, they have been mostly applied to pharmaceuticals (PB-PK) and for hazard assessment of industrial chemicals (solvents). In food safety, PB-TK are increasingly developed for substances with well known TK. These include regulated substances (mostly pesticides (organophosphates, fungicides…)), food contact materials (bisphenol A) and contaminants (persistent organic pollutants), acrylamide, heavy metals (cadmium, lead, mercury…), metalloids (arsenic). Applications of PB-TK models in key areas of hazard assessment are highlighted including

interspecies differences, human variability, biomonitoring programmes, and combined exposure to multiple chemicals and *in vitro* to *in vivo* extrapolations.

Interspecies differences

PB-TK models have been used to investigate interspecies differences/similarities in TK processes, in order to facilitate extrapolations between test species and humans. In the pesticide area, PB-TK models have been developed for the conazole fungicide triadimefon and its primary metabolite triadimenol from rat experimental data and were then extrapolated to humans using *in vitro* metabolic constants from human hepatic microsomes. Human equivalent doses (HEDs) were then calculated from a rat NOAEL dose using the area under the concentration curve in the brain and blood for triadimefon and triadimenol. Such reverse dosimetry PB-TK models are expected to be applied in the future to better estimate the human exposure profile (external dose) based on internal dose of other conazole fungicides and other compounds (Crowell et al., 2011). Interspecies differences in bisphenol A (BPA)-TK were investigated using a 7-compartment oral PB-PK model (brain, liver, fat, slowly perfused, richly perfused, plasma, and gonads) in the infant and adult monkey and a one compartment sub-model for BPA phase II metabolites (e.g. BPA-glucuronide and sulphate). The model showed metabolism of BPA at all ages (post-natal day PND 5 to adult) by the gut wall and liver (Doerge et al., 2010). From the monkey model, Fisher et al. (2011) extrapolated the model to humans, incorporating knowledge of gut wall and liver metabolism (e.g. glucuronidation). The authors could demonstrate that previous human models that did not take into account gut metabolism over-estimated the concentrations of BPA as parent compound in the serum.

A PB-TK model was developed in monkeys for the contaminants perfluorooctanoic acid (PFOA) and perfluorosulphonate (PFOS) and extrapolated to humans taking into account differences in half life between monkeys (several months) and humans (several years). In addition, the model successfully simulated human plasma concentrations using data collected from residents of two communities exposed to PFOA in drinking water. Sensitivity analysis was performed to test whether the model was able to describe the available PFOA and PFOS plasma concentrations adequately. Overall, even though the data were highly variable due to the long half life of PFOS and PFOA in humans, predictions of the model were in good agreements with the experimental data (Loccisano et al., 2011).

Human variability

PB-TK models are particularly suited to explore, understand and predict the determinants of inter-individual variability in TK when the required information is adequately incorporated. Bois et al. (2010) reviewed the prediction and sources of inter-individual variability in ADME parameters. They distinguished three main contributing sources to the total variability: (1) the variation across a population of 'normal' individuals at the same age, e.g. young adults; (2) the variation across the population resulting from their different ages, e.g. infants or the elderly; and (3) the variation resulting from the existence of subpopulations that differ in some way from the 'normal' population, e.g. due to genetic polymorphisms. The authors pointed out that a fourth source of variability, namely the health status, which is frequently disregarded, should also be considered. Human variability in PB-TK has been applied to methylmercury and combined with a Monte Carlo analysis to provide information on the distribution of acceptable ingestion rates across the population (Clewell et al., 2000). Clewell et al. (2004) developed an age-dependent PB-TK model for isopropanol and its metabolite acetone incorporating time dependent changes in physiological and biochemical parameters based on data from the literature. A PB-TK model was combined with Monte Carlo techniques in order to take into account human variability in paraoxonase polymorphism in relation to parathion TK and the inhibition of acetylcholinesterase (Gentry et al., 2002). Finally, chemical specific adjustment factors (CSAFs) have been derived for a number of chemicals in specific subgroups of the population using PB-TK models taking into account variability in physiological parameters and TK processes (Valcke and Krishnan, 2014).

Finally, it has been argued that Bayesian methods for PB-TK modelling have emerged as the best suited approaches, given the large amount of prior information they require (Bernillon and Bois, 2000). A number of applications of posterior Bayesian PB-TK modelling have been published on chloroform (Lyons et al., 2008), dichloromethane (David et al., 2006), methylmercury (Allen et al., 2007), nanoparticles (Péry et al., 2009), tetrachloroethylene (Chiu and Bois, 2006; Covington et al., 2007) and cadmium (Amzal et al., 2009; EFSA, 2009c).

Human Biomonitoring

A number of biomonitoring programmes are currently ongoing to assess environmental exposure of humans to xenobiotics (e.g. EU ESBIO, COPHES; US CDC NHANES; Canadian Health Measures Survey). The goal of these projects is to determine relative trends in exposure to chemicals, across time and subpopulations. Due to the lack of data, there is often little information correlating biomarker concentrations with exposure levels and durations. As a result, it can be difficult to use biomonitoring data to derive Derived No-Effect Level (DNEL) values under the EU REACH program, or Reference Dose/Concentration (RfD, RfC) values of the US EPA based on internal dose (Bartels et al., 2012). PB-PK models have been shown to be of great help for a quantitative interpretation of human biomonitoring data to relate exposure to blood concentrations as exemplified with the NHANES biomonitoring data for cadmium (Clewell et al., 2008; Ruiz et al., 2010).

Verner et al. (2009) developed a mother–infant PB-PK model for Persistent Organic Pollutants (POPs) and using maternal blood levels at the time of delivery, exposure of mothers to several metabolites of 1,1,1-trichloro-2,2-bis(p-chlorophenyl)ethane (DDT), polychlorinated biphenyls (PCB) and other POPs was estimated to subsequently simulate infant blood, breast milk, and cord blood POP concentrations. Simulations were then compared with corresponding measured levels and predictions were strongly correlated with measured concentrations of residues that were above the limits of detection. This study shows how PB-PK model can be validated with individual data and how they can help reduce sampling efforts. In addition, these models enable the use of individual TK profiles of POPs and incorporate them in epidemiological studies to investigate adverse effects on child development. Another example by Verner et al. (2012) illustrates the relevance of PB-TK model to integrate the characterisation of exposure accounting for uptake through multiple pathways and physiological parameters influencing the TK. Using styrene as an example, the authors determined the best times to sample venous blood and end-exhaled air for biomonitoring purposes, characterised inter-individual variability in biological levels following occupational exposure to styrene, and proposed biological limit values using a population-based PB-TK model.

Combined exposure to multiple chemicals (‘Chemical mixtures’)

In the context of combined exposure to multiple chemicals, a number of PB-TK models have enabled to investigate potential TK interactions to calculate potency factors such as interaction-based hazard index (HI) using information on the chemicals’ tissue concentrations (EFSA, 2013). Haddad et al. (2001) proposed a methodology to model occupational inhalation exposure to airborne mixtures of dichloromethane, benzene, toluene, ethylbenzene, and m-xylene. The basis of the proposed methodology related to the characterisation of the change in tissue dose metrics in humans, during mixed exposures using an interaction-based PB-TK model. More recently, PB-TK models for four solvents (styrene, benzene, ethylbenzene and toluene) were developed taking into account metabolic interactions at the level of their oxidation pathway mediated via CYP2E1. The models were calibrated using three joint models of benzene, toluene, ethylbenzene, m-xylene toxicokinetics using Markov chain Monte Carlo simulations and single-substance exposure data (Cheng and Bois, 2011). Sasso et al. (2010) developed a generalised PB-TK model for mixtures using a chemical independent approach based on modules approach that can be directly ‘mapped’ to individual TK models for specific chemicals, while maintaining physiological consistency across different chemicals, the model was applied to a mixture methylmercury, cadmium, lead, arsenic (and metabolites), toluene, and benzene.

In vitro-In vivo extrapolation

A number of QIVIVE models have been developed to extrapolate from an *in vitro* concentration used in an experimental setting to an *in vivo dose*, which could be significant for human exposure. Two examples of QIVIVE that included biotransformation measured *in vitro* into PB-TK models are provided by studies on deltamethrin (Mirfazaelian et al., 2006; Tornero-Velez et al., 2010) and chlorpyrifos/diazinon (Timchalk and Poet, 2008). The deltamethrin example also showed another value of using *in vitro* derived metabolism data in the PB-TK model. Using the age-specific enzyme kinetic parameters determined *in vitro*, it was possible to extend the adult PB-TK model to different life stages. This capability of PB-TK modelling is advantageous for considering potentially sensitive subpopulationssuch as infants and children in risk assessment (Mirfazaelian et al., 2006, Tornero-Velez et al., 2010).

Another recent QIVIVE approach evaluated the complex metabolism of estragole in rats and humans from *in vitro* data. Punt et al. (2008) characterised the multiple steps of estragole bioactivation and detoxification mediated by a number of enzymes (CYPs, UGTs, dehydrogenases, and sulfotransferases) using *in vitro* systems (microsomes or subcellular fractions (S9). The Michaelis-Menten kinetic parameters (Vmax and Km) parameters were then scaled to *in vivo*, based on the microsomal/S9 protein content, and the interplay of these multiple reactions were integrated into a PB-PK model. Simulated concentrations of a metabolite (1-hydroxyestragole glucuronide) from the PB-TK model for rat and human were consistent with the observed *in vivo* data from rat and human urine.

Recently, a 3D dynamic flow model with primary human hepatocytes was used to predict the metabolic clearance of ethoxycoumarin. The model was optimised for cell seeding density, medium composition and extracellular matrix proteins and the hepatocytes were cultured for up to 7 weeks. In addition, the model provided *in vivo* liver-like structure as well as important liver-specific functions which included albumin and total protein production, glucose utilisation, lactate production, and CYP 3A4 activity across multiple tissue donors. The *in vitro* intrinsic clearance of 7-ethoxycoumarin was determined and compared to that in hepatocyte suspension and gave reproducible and stable estimates of clearance that were similar to previously published values. The authors discussed that such tools could be valuable to make accurate QIVIVE to predict metabolic clearances and provide ways to assess chronic effects of chemicals and their metabolites in a complex 3D-environment under dynamic flow more accurate (Choi et al., 2013). Such a model is not strictly speaking a PB-TK but provides a sound approach for QIVIVE to determine TK parameters.

Other recent *in vitro* tools for QIVIVE include models to investigate phase II metabolism such as glucuronidation, as the major phase II metabolic pathway in humans. Wu et al. (2013) reviewed the use of hepatic microsomes incubated with bovine serum albumin in addition to hepatocytes which provide accurate predictions of *in vivo* glucuronidation including both hepatic clearance and intestinal availability (Wu et al., 2013). Such models have the potential to predict *in vivo* metabolism and dose metrics from *in vitro* data and can improve IVIVE for the design of PB-TK models.

3.3. Physiologically-based toxicokinetic-toxicodynamic models and application in hazard assessment

3.3.1. Principles

Toxicodynamic (TD) models have been defined as ‘mathematical descriptions simulating the relationship between a biologically effective dose and the occurrence of a tissue response over time’ (WHO, 2009). When the TD for a specific compound is known at the target organ or site (e.g. a cell or enzyme) it can be linked to the predicted TK from a PB-PK model. Therefore, a PB-TK model is combined with dose response data to get to a PB-TK-TD model. Historically, biologically-based dose-response (BBDR) models were introduced to bring mechanistic information into dose-response assessment and today PB-TK-TD are considered to be the most comprehensive and phenomenologically-based models and thereby the most comprehensive BBDR (Shuey et al., 1994; Setzer, 2001).

PB-TK-TD models are very useful since they provide a highly refined tool, in which it should be possible to reduce uncertainty for higher tier risk assessments of single and multiple chemicals. Generally speaking, PB-TK-TD models can provide a tool to estimate the internal concentration of a chemical and its metabolites, together with its toxicity, by integrating population variability into TK and TD. They also allow better understanding of the mechanistic basis for extrapolation from experimental data (using either IVIVE/QIVIVE or animal studies) to the *in vivo* human situation such as high doses to low dose extrapolation between animals and humans, dose and interspecies differences in bio-activation and detoxification, non-linearity in dose response, qualitative and quantitative response to the same cumulative dose administered by different routes and exposure scenarios (Krishnan and Andersen, 2001). As described for PB-TK models, PB-TK-TD are built using the body as a set of interconnected compartments of differential mathematical equations describing the ADME of a specific chemical and/or its metabolite, and then they connect the internal dose to the dose response of the adverse dynamic effect (from *in vivo* and more recently *in vitro* studies) for the compound and/or its metabolites (s).

IVIVE and QIVIVE approaches can incorporate *in vitro* data into *in vivo* PB-TK-TD based on cellular toxicity assays. As discussed for PB-TK, the first step of QIVIVE is the identification of the consequence of metabolism (detoxification or bioactivation) to identify the toxic entity (parent or metabolite). Again, the collection of *in vitro* data to support the prediction of *in vivo* clearance is a complex process particularly to relate the concentration that would be equivalent to a toxic effect *in vitro* (such as using HTS assays as in Toxcast). Ideally, key variables need to be predicted for a full QIVIVE model that would reflect metabolism and physiology in a holistic manner. Such key variables include intestinal absorption and pre-hepatic clearance, extrahepatic metabolic clearance, renal clearance, and volume of distribution (particularly for acute *in vivo* exposures). Current examples of QIVIVE often use historical *in vivo* data and it has been anticipated by a numbers of authors that such information will be generated in the near future using *in silico* tools and targeted *in vitro* studies, particularly novel *in vitro* systems that better mimic *in vivo* conditions (Yoon et al., 2012). High-throughput *in vitro* toxicity screening can provide an efficient way to identify potential biological targets for chemicals, but relying on nominal assay concentrations may misrepresent potential *in vivo* effects of these chemicals due to differences in bioavailability, clearance, and exposure (Wetmore et al., 2012). However, it has been estimated that at the moment, QIVIVE calculations can be associated with uncertainties of more than an order of magnitude (Rotroff et al., 2010; Yoon et al., 2014).

The use of the nominal concentration in the *in vitro* toxicity assay to characterise the toxicity of a compound is an easy, fast and cost effective process compared with QIVIVE which involves much more complex measurements and understanding. However, *in vitro* toxicity test results are quantitatively meaningless for risk assessment without QIVIVE. Relative potency estimates from *in vitro* toxicity assays are obtained under conditions that do not reflect differences in the bioavailability and clearance of the chemicals, which are the key determinants of the doses *in vivo* that would be associated with tissue exposures equivalent to the *in vitro* assay, Therefore, except for qualitative hazard identification, *in vitro* toxicity assay results can only be interpreted on the basis of QIVIVE (Yoon et al., 2012). In order to understand the inter-individual differences in hepatic clearance, knowledge is required on population distributions of protein binding, hepatic uptake, biliary transport, blood flow in healthy adults and subgroups of the populations. Key issues include better *in vitro* tools to measure multiple determinants of clearance through the manipulation of proteins levels, flow rates and transport inhibitors. New developments include liver bioreactors that are promising tools to investigate the effect of biological conditions on hepatic clearance, protein binding and transport rate constants which can be used in a high troughput context. It is foreseen that, in the future, databases will be generated and can be combined with physicochemical properties to create QSAR models, which may facilitate IVIVE/QIVIVE using clearance data generated from hepatocytes or microsomes (Yoon et al., 2014).

3.3.2. Application in hazard assessment

Currently, applications of PB-TK-TD in human hazard assessment are possible for compounds with well known TK and TD. It is foreseen that, in the future, as databases describing TK and TD parameters for numerous compounds, these models may be applied in a more predictive manner. Examples of PB-TK-TD applied to interspecies differences, human variability, epidemiological studies, combined exposure to multiple chemicals and *in vitro* to *in vivo* extrapolations are presented below.

Interspecies differences

Young et al. (2007) developed a physiologically based pharmacokinetic model for acrylamide (AA) and three of its metabolites: glycidamide (GA) and the glutathione conjugates of acrylamide (AA-GS) and glycidamide (GA-GS) in mice and rats. Then human urinary excretion data and haemoglobin adducts data were used to extrapolate to a human model. GA-DNA adducts and haemoglobin (Hb) adducts with AA and GA were included as pharmacodynamic components of the model. Doerge et al. (2008), estimated probable AA intake in the U.S. population, and used PB-TK-TD modelling to integrate the findings of rodent neurotoxicity and cancer into estimates of risks for humans. These modelling techniques have reduced the uncertainty inherent in extrapolating toxicological findings across species and dose by comparing common exposure biomarkers.

PB-TK-TD models can simulate bioactivation and detoxification of alkenylbenzenes (estragole, methyleugenol and safrole), which are a class of naturally occurring compounds found in herbs and spices, known to be carcinogenic (Punt et al., 2008; Al-Subeihi et al., 2011; Martati et al., 2011). These models can be used to extrapolate from benchmark dose causing 10 % extra tumour incidence (BMD_{10}) down to the so-called virtual safe dose (VSD) (the dose resulting in one in a million extra tumour incidence upon life time exposure). The PB-TK models were built using available literature information and *in vitro* TK parameters (Vmax and Km) from human and rat microsomes for each metabolite (Punt et al., 2007, 2011). Additionally, a PB-TD model was developed, by measuring formation of estragole DNA adducts in rat primary hepatocytes and was further validated *in vivo* with male SD rats (Paini et al., 2012). Recently, Van der Berg et al. (2013) applied the estragole PB-TK/TD model to the hazard assessment of plant food supplements containing estragole to predict *in vivo* effects in humans.

Human variability

Human variability is a key parameter influencing cadmium levels in urine, which are widely accepted as a measure of the body burden and its cumulative amount in the kidneys. The Scientific Panel on Contaminants in the Food Chain (CONTAM Panel) of EFSA carried out a meta-analysis on a selected set of studies to evaluate the dose-response relationship between urinary cadmium and urinary beta-2 microglobulin (B2M). B2M, a low molecular weight protein, was recognised as the most useful biomarker in relation to tubular effects and .a Hill model was fitted to the dose-response relationship between urinary cadmium and B2M for subjects over 50 years of age and the whole population. From the model, a $BMDL_5$ of 4 µg Cd/g creatinine was derived for a 5 percent increase of the prevalence of elevated B2M. A chemical-specific adjustment factor of 3.9, to account for inter-individual variation of urinary cadmium within the study populations, was applied, leading to a value of 1.0 µg Cd/g creatinine. Such a value was also supported by data from occupationally exposed workers and by the results of several individual studies using a variety of biomarkers. A one-compartment model was fitted to a large data set based on non-smoking Swedish women (age range from 58 to 70 years), to allow an estimation of the relationship between dietary cadmium exposure and urinary cadmium concentration. The dietary cadmium exposure that corresponds to the critical urinary cadmium in food cadmium concentration of 1 µg/g creatinine after 50 years of exposure was then estimated using the model corresponding to. A safe average daily dietary cadmium intake 0.36 µg Cd/kg body weight (b.w.) or weekly dietary intake of 2.52 µg Cd/kg b.w. Based on these figures, EFSA established a tolerable weekly intake (TWI) for cadmium of 2.5 µg/kg b.w. (Amzal et al., 2009; EFSA, 2009a,b,c).

Epidemiological Studies

Cord serum levels of PCB-153, a highly persistent polychlorinated biphenyl (PCB) congener, were recently reported to be associated with lower birth weight in a meta-analysis of data from > 7,000 pregnancies (Govarts et al., 2012). Verner et al. (2013) suggested that gestational weight gain, which is associated negatively with PCB levels in maternal and cord blood and positively with birth weight, could substantially confound this association. They thus estimated the influence of gestational weight gain on the association between PCB-153 exposure and birth weight using a pharmacokinetic model and Monte Carlo simulations accounting for variability in physiologic parameters and their correlations. The PB-PK model was evaluated by comparing simulated plasma PCB-153 levels during pregnancy to serial measurements in 10 pregnant women from another study population. The association between simulated plasma PCB-153 levels and birth weight were estimated using linear regression models. The plasma PCB-153 level profiles generated with the PB-PK model were comparable to measured levels in 10 pregnant women. A 118-g decrease in birth weight (95 % CI: -129, -106 g) was estimated for each 1μg/L increase in simulated cord plasma PCB‑153. This decrease in birth weight, was in the same range as the value estimated (-150 g) from a previous meta-analysis. The estimated decrease in birth weight was reduced to -6 g (95 % CI: -18, -6 g) when adjusted for simulated gestational weight gain. This study, based on a pharmacokinetic approach, suggests that the association between prenatal levels of PCBs and birth weight may be strongly confounded by the effect of gestational weight gain on both blood PCB levels and birth weight. Overall, the PB-TK-TD model illustrates that epidemiological associations between pollutants and health outcomes may be attributable partly to TK and can be applied to other pollutants in the future (Verner et al., 2013).

Combined exposure to multiple chemicals

In a study by Hinderliter et al. (2011) inter-individual variability in physiology, metabolism, and physical activity was estimated. This allowed the evaluation of individuals' susceptibility to the potential effects of chlorpyrifos (CPF) using a sensitivity analysis in a PB-TK-TD model. The results indicated that the metabolic capacities of liver CYP and paraoxonase-1 (PON-1) in liver and blood were sufficient to prevent significant toxic responses due to brain and red blood cell acetylcholinesterase (AchE) inhibition (unsaturated) at low dietary CPF exposure in both children and adults.

Several PB-TK-TD models have been developed to address combined exposure to multiple pesticides such as organophosphates (Timchalk et al., 2002; Knaak et al., 2004; Poet et al., 2004; Lee et al., 2011a), carbamates (Zhang et al., 2007; Knaak et al., 2008; Pelekis and Emond, 2009), and pyrethroids (Mirfazaelian et al., 2006; Tornero-Valez et al., 2010; Aylward et al., 2011). For organophosphorus insecticides, a PB-TK model was developed to model potential interactions between CPF and nicotine at a) the TK level to predict CPF's metabolite concentrations (CPF-oxon) in blood and brain, and at b) the TD level to compare prediction of (AchE) inhibition in the brain with experimental data. Results showed that CPF-oxon levels were lower following the expected Vmax increase in rats co-exposed to nicotine and CPF. Authors concluded that that repeated nicotine exposure can alter CPF metabolism *in vivo*, resulting in altered AchE inhibition (Lee et al., 2011b).

Other PB-TK-TD models have been developed for CPF and diazinon (DZ) and the models included a number of important metabolic steps such as CYP450 mediated activation/detoxification, B-esterases, butyrylcholinesterase (B-E) and AchE or PON-1 oxon detoxification. Since both insecticides were shown to inhibit the CYP-mediated metabolism *in vitro* in a concentration-dependent manner, the PBPK model was modified to reflect the TK of the CYP inhibition (competitive vs. non-competitive). In addition, B-esterase metabolism was described as dose-additive, and no PON-1 interactions were assumed. The PBTK model was then compared with previously published rodent oral TK data and TD data (AchE inhibition) for co-exposure to CPF and DZ. No differences between predicted TK and published TK data were shown for either CPF or DZ or their respective metabolites, while TD AchE inhibition was shown to be described using dose-addition. The authors concluded from the model that,

at low environmentally relevant binary doses of CPF and DZ, the TK and TD of the mixture were expected to be linear and dose-additive (Timchalk et al., 2008).

In vitro to *in vivo* extrapolations

The JRC of the European Commission has recently developed a number of Virtual Cell-based (VCB assays). An example took into account the fate of a compound in the experimental *in vitro* system as 1) the partitioning between the plastic wall, headspace, serum proteins, lipids, and 2) the compound dynamics within the cell. The model was then coupled with a cell growth model and a toxic effect model (Zaldívar et al., 2010; Zaldivar Comenges et al., 2011). In addition, the VCB assay model could been coupled to PB-TK models to give a PB-TK-TD establishing the relationship between TK and TD taking into account the real concentration affecting the cells. The model was applied to acetaminophen to predict the internal concentration using the VCB assay and compare it with cell viability data (Péry et al., 2013).

An example of a simple QIVIVE modelling using HTS from the Toxcast assays programme of the US-EPA has been recently published (Wetmore et al., 2013). Rat hepatic metabolic clearance and plasma protein binding were measured for 59 ToxCast phase I chemicals. IVIVE provided estimates of the daily internal dose in rats as an oral equivalent dose (OED). This OED would then result in steady-state *in vivo* blood concentrations equivalent to an AC 50 (concentration at 50 % of maximum activity) or lowest effective concentration (LEC) across more than 600 ToxCast *in vitro* assays. Statistical classification analysis was performed using either the OED or unadjusted AC_{50}/LEC values for the *in vitro* assays to predict the *in vivo* effects of the 59 chemicals. The authors concluded that adjusting the *in vitro* assays with a PB-TK did not improve the ability to predict *in vivo* toxicity as either a discrete response or as a low effect level on a continuous dose scale. However, a comparison of such *in vitro* assays with the lowest oral equivalent dose with *in vivo* endpoints (LOEL) effect level suggested that: 1) the lowest oral equivalent dose may provide a conservative estimate of the POD for a chemical in a dose-response assessment, and 2) the approach may also be used to identify potential MIE leading to adversity HTS assays *s* can provide an efficient way to identify potential biological activity of chemicals (Wetmore et al., 2013).

3.4. *In silico* tools and threshold of toxicological concern

3.4.1. *In silico* tools

Broadly interpreted, '*in silico*' tools available to toxicologists and risk assessors aim to predict toxicity of chemicals and cover a wide range of methodologies that would also comprise molecular modelling approaches and general computational toxicology tools, including theoretical models based on the intrinsic structural and physicochemical properties of chemicals and rule-based expert systems. These computational tools often require chemical structure and/or a few physico-chemical properties, as input to provide a fast method for screening of untested substances. They are also helpful tools for identifying emerging risks in the food chain from those chemicals that have not yet been tested for safety to human health or the environment (EFSA, 2014). This section highlights briefly key *in silico* tools that widely used in chemical hazard assessment namely (Quantitative) Structure Activity Relationship models and read-across methods as well as decision making tools such as the threshold of toxicological concern.

Structure-Activity Relationships (SARs) and Quantitative Structure Activity Relationships (QSARs) are sometimes collectively referred to as (Q)SARs and are mathematical models that relate the structure of chemicals to their biological activities. A SAR provides a qualitative relationship between a particular substructure and the presence or absence of a biological activity, regarding the capacity to modulate a biological activity imparted by another sub-structure (e.g. suspected carcinogens mutagens, and reprotoxicants). A QSAR provides a mathematical relationship between a biological activity and one or more molecular descriptors that are used to predict the activity. The term 'quantitative' refers to the fact that the molecular descriptors are quantifiable on a continuous scale

and thus provide a quantitative relationship with toxicity (which may itself be expressed in quantitative or categorical terms). The molecular descriptors of the chemical are generally their inherent physicochemical properties such as atomic composition, structure, sub-structures, hydrophobicity, surface area charge, and molecular volume. QSARs may be classified based on their dimensionality with 1D-QSAR referring to a system where the effect can be correlated with a single (e.g. physicochemical) property, 2D-QSAR with atomic connectivity or two-dimensional (e.g. pharmacophoric) patterns, and 3D-QSAR with the three-dimensional structure of a compound. Dimensionalities with $n > 3$ ($n = 4, 5, 6$) are referred to as 'multi-dimensional QSAR' or short 'mQSAR' and typically include a multiple representation of the ligand such as 4D-QSAR (Vedani et al., 2000; Tseng et al., 2012) and the protein 5D/6D (Vedani et al., 2006).

QSARs are typically used in combination with other non-testing (e.g. read-across) and testing (e.g. *in vitro*) methods in the context of ITS and Weight-of-Evidence assessments.

Read-across has been defined by ECHA as 'a technique for predicting endpoint information for one substance (target substance), by using data from the same endpoint from (an)other substance(s), (source substance(s))'. ECHA pointed out that the read-across approach has to be considered on 'an endpoint-by-endpoint basis due to the different complexities (e.g. key parameters, biological targets) of each endpoint'. In addition, ECHA used the term **analogue approach** 'when the read-across approach is employed' within a group of a very limited number of substances for which trends are not apparent: i.e. 'the simplest case is read-across from a single source substance to a target substance' (ECHA, 2008). In the case of a high number of substances in a group the term **category approach** is used (ECHA, 2008). A wide range of *in silico* tools are available for grouping chemicals and applying read-across. Being a non-formalised approach, it requires considerable expert knowledge and judgment. The comprehensive guidance on grouping and read-across has been published by the OECD (OECD, 2007) and ECHA (ECHA, 2008).

The concepts of grouping chemicals and read-across has been reviewed and illustrated elsewhere by Enoch et al. (2010). A comprehensive guidance for applying the grouping approach has been published by OECD (2007) and more recently, other systematic expert-driven processes have been proposed for read-across (Wu et al., 2010; Blackburn et al., 2011). Grouping and read-across have been used within the OECD High Production Volume Chemicals Program as an alternative for experimental testing and are currently being applied under REACH. Examples of applications of grouping read-across have been reported (Wu et al., 2010; Blackburn et al., 2011). The reliability of read-across depends on the selection of suitable analogues associated with reliable experimental data. In some cases, it is only possible to identify one or a limited number of suitable analogs, whereas in other cases it is possible to build up groups of chemicals. Schilter et al. (2014) discussed options to group chemicals according to their similarities:

- Physico-chemical properties (e.g. molecular weight, solubility, vapour pressure lipophilicity), play a key role in the bioavailability of chemicals.

- Functional/mechanistic/structural alert groups (e.g. aldehyde, epoxide, ketone, Michael acceptor, nitrosamines, aromatic amines).

- Chemical similarity, e.g. based on the Tanimoto coefficient. If a new substance is very similar to an existing one, it is assumed that minor modifications to its structure are unlikely to affect its properties and, for hazard assessment purposes, the same hazards and potencies can be used.

- Similarity in breakdown or metabolic products. Physical or biochemical processes may generate compounds of similar structure (e.g. ester hydrolysis; oxidation of primary alcohols and aldehydes to carboxylic acids).

When QSAR and read-across approaches are applied to toxicity prediction, they are typically based on data or knowledge relating to both TK and TD processes. In the case of QSARs, some of the underlying parameters (predictor variables) may be associated with TK (e.g. partitioning coefficients) whereas others may be associated with TD (e.g. electronic properties). Often however, the predictor variables included in QSARs already account for both TK and TD contributions to toxicity. QSARs can also be used to predict physicochemical properties that may serve as input parameters in PB-TK models (e.g. protein binding coefficients and partitioning coefficients across biological barriers). Recently, the use of OMIC technologies such as metabolomics has been suggested to optimise the chemical grouping process by providing biologically-based criteria for toxicological equivalence; the authors have named this approach Quantitative biological activity relationship (QBAR) (Van Ravenzwaay et al., 2012).

EFSA (2014a) has discussed a typical workflow, which would first examine existing data and information for possible read-across and grouping using the OECD QSAR Toolbox and the databases discussed above. A second step would be to predict metabolism in the relevant species (human, rat..) using metabolism prediction tools such as the expert systems METEOR (LHASA), OASIS-TIMES or the US EPA MetaPath pesticide database. Another option is to use molecular modelling tools to conduct 3-D docking studies in potential target receptors and enzymes and these studies can also be used to build QSAR models (EFSA, 2014a).

Key databases for QSAR and read-across include the OECD QSAR Toolbox (http://www.qsartoolbox.org/), a hazard identification tool, which contains QSAR relationship methodologies that can be used to group chemicals into categories sharing the same structural characteristics and/or MoA. The systematic grouping of chemicals according to the presence or modulation of a particular effect for all members of the category is based on the presumption of a common chemical structure or MoA/AOP. The Toolbox can be used to provide:

- estimates for all substances in a category;
- extrapolation of the empirical data from tested chemicals to derive estimates for an untested chemical within a category;
- trend analysis estimates (increasing, decreasing or constant) among relevant regulatory endpoint data;
- estimates using category-based statistical models (QSARs).

The QSAR Toolbox is entering the phase 3 of its development over a six to ten year period, and a common terminology/ontology has been developed for the Toolbox by ECHA. The aim is to improve the user-friendly features of the toolbox in terms of architecture, workflows as well as to address new and less experienced user guidance needs. In addition, the quality assurance and range of the databases will be addressed together with ontology harmonisation, improvements of information technology and additional functions, such as: report documentation and options to save searches and relevant scientific approaches; three dimensional molecular docking function possibilities, up to the eventual population of the toolbox with the AOP data (EFSA, 2014a).

Other databases providing toxicity data for chemicals include the eChemPortal hosted by the OECD, which allows simultaneous searching of reports and datasets by chemical name and number and by chemical property. Direct links to collections of chemical hazard and risk information prepared for government chemical review programmes at national, regional and international levels are available. In addition, the eChemPortal provides exposure and use information on chemicals. Other databases include Chembase (www.chembase.com/), ChemIDplus (http://chem.sis.nlm.nih.gov/chemidplus/), ChemSpider: (www.chemspider.com/), Pubchem (http://pubchem.ncbi.nlm.nih.gov/), Carcinogenic Potency Database (http://toxnet.nlm.nih.gov/cpdb/cpdb.html), DSSTox (www.epa.gov/comptox/dsstox/), European chemical Substances Information System

(esis.jrc.ec.europa.eu/), NTP Database: (http://ntp.niehs.nih.gov/), IPCS (www.inchem.org/), ToxNet (http://toxnet.nlm.nih.gov/) (EFSA, 2014a).

QSAR software and models are being used by international and national organisations such as the Toxicity Estimation Software Tool (TEST), the OECD QSAR toolbox models and High-throughput Virtual Molecular Docking (HTVMD) (OECD, 2004, 2012; Rabinowitz et al., 2008; Benfenati et al., 2009; Zhu et al., 2009), MetaCore (Teschendorff and Widschwendter, 2012), and the TOPKAT model (Rakyan et al., 2011). HTVMD models use a ligand-based chemoinformatics strategy that allows for the prediction of relationships between various attributes of ligands and their binding to a number of known targets as a direct agonist, such as the oestrogen receptor (US-EPA, 2013). HTVMD models are increasingly being used in risk assessment and can screen thousands of chemicals for the potential affinity of their 3D structures to the binding sites of active proteins (US-EPA, 2013). Other public and commercial (Q)SAR models and expert systems are available for assessment of chemical toxicity. These include DEMETRA, CAESAR, VEGA, TEST, DEREK, METEOR, Multicase, PASS, OASIS Times. The latter also allows prediction of metabolites as well as assessment of their toxicity.

Overall, (Q)SARs methods are increasingly predictive for hazard identification in relation to acute toxicity, mutagenicity, genotoxicity and bioacummulation. However, applications of Q(SARs) and read-across to the prediction of TK properties (ADME and sub-chronic and chronic toxicity) for chemicals relevant to the food safety area are still limited and considerable research is undergoing in this area (Roncaglioni et al., 2013; Scholtz et al., 2013; Gissi et al., 2014). In addition, a increasing number of Q(SAR) models, and databases are available and their precision, specificity and sensitivity may vary and would need to be evaluated (Roncaglioni et al., 2013; Scholtz et al., 2013). In terms of hazard assessment, combining Q(SARs) from more than one model with additional information from structural alerts, read-across estimates but also from *in vitro* and *in vivo* toxicological studies using a Weight of Evidence (WoE) approach can improve the utility and the validation of these tools and increase overall reliability of *in silico* methods (Scholtz et al., 2013; Roncaglioni et al., 2013; US-EPA, 2013; EFSA, 2014a).

3.4.2. Threshold of Toxicological Concern

The Threshold of Toxicological Concern approach (TTC) is a well known decision making tool that has been used for a number of years for hazard assessment purposes. The reader is referred to the opinion of the Scientific Committee of EFSA on 'exploring options for providing advice about possible human health risks based on the concept of Threshold of Toxicological Concern' (EFSA SC, 2012). The approach is based on a historical toxicological database built on the empirical evidence that for non-cancer effects there are thresholds below which toxicity does not occur, whereas for cancer effects the likelihood of tumour incidence is zero to very small at very low exposure levels. The TTC values are based on the analysis of the distribution of NOAELs for compounds sorted to different categories of toxicity (Cramer classes). These values are based on the 5th percentile NOAELs along with the application of default uncertainty factors of 100-fold allowing for interspecies differences and human variability (Kroes et al., 2005). The Scientific Committee of EFSA concluded on the following human exposure threshold values to be sufficiently conservative to be used in EFSA's work: 0.15 µg/person per day for substances with a structural alert for genotoxicity, 18 µg/person per day for organophosphate and carbamate substances with anti-cholinesterase activity, 90 µg/person per day for Cramer classes II and III, and 1800 µg/person per day for Cramer class I substances. Thus, for chemicals of unknown toxicity, human exposure thresholds (TTC values) can be established below which there is a low probability of adverse effects on health. The combined and stepwise use of TTC values, i.e. the TTC approach, can be used to assess substances of unknown toxicity present at low levels in the diet. Application of the TTC approach requires only knowledge of the chemical structure of the substance concerned and reliable information on human exposure. The extent to which the TTC is accepted depends on the regulatory application and context. In general, the approach is better accepted for the assessment of non-intentionally added substances, such as contaminants, reaction by-products, and metabolites, for which experimental toxicity data are not available and consumer exposure is low compared to the TTC threshold. The TTC approach is also used in the safety

assessment of flavourings, in which read-across of toxicological properties within structurally defined groups is also permitted (EFSA SC, 2012).

3.4.3. Application of *in silico* tools and TTC to human hazard assessment

A wide range of (Q)SAR models are available to predict a number of toxicological properties *in silico* including mutagenicity, genotoxicity, acute toxicity and bioaccumulation as reviewed elsewhere (JRC, 2011; Worth et al., 2011; EFSA, 2014a).

Several surveys have been conducted to establish the extent to which *in silico* tools are accepted and used by regulatory bodies and industry (Lo Piparo et al., 2011; Mays et al., 2012; IEH, 2013). These surveys provide consistent findings, and reveal that grouping and read-across approaches are the most often used approaches across different regulatory sectors. (Q)SAR tools are used much less commonly in hazard assessments, and rarely as stand-alone methods. As an example, the new Commission Regulation (EU) 283/2013 setting out data requirements for pesticide approval does not discuss the potential use of QSARs for ecological and human health hazard assessment (EU, 2013). However, examples for which (Q)SARs can be/are used to generate toxicity data include: a) compounds for which a risk assessment is not explicitly required by legislation (e.g. contaminants, impurities, co-formulants, pesticide residues in food of animal or plant origin, pesticide groundwater metabolites); and b) compounds for which there is an urgent need to inform risk management decisions, such as in the case of incidents of food contamination.

The possible applications of (Q)SARs in the assessment of pesticide residues for dietary risk assessment have been explored in several projects at EFSA, focussing in particular on the use of (Q)SARs for predicting genotoxicity and carcinogenicity (JRC, 2010), as well as developmental toxicity and neurotoxicity (JRC, 2011). While recognising that further efforts are needed to improve and evaluate (Q)SAR models for these endpoints, the Scientific Panel on Plant Protection Products and their Residues (PPR Panel) of EFSA proposed the application of *in silico* tools (QSAR and read-across) for the prediction of genotoxicity and developmental toxicity, to complement the TTC approach in the assessment scheme for pesticide metabolite exposure (EFSA PPR Panel, 2012). Possible use of QSARs has also been foreseen in the QPS (qualified presumption of safety) of botanical food supplements (EFSA, 2014a).

Recently, a database on ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) structure activity relationship, abbreviated admetSAR, has been published (Cheng et al., 2013). The database is an open source, text and structure searchable, and is continually updated. AdmetSAR manages available ADMET-associated properties data from the published literature with over 210 000 ADMET annotated data points for over 96 000 compounds with 45 kinds of ADMET-associated properties, proteins, species, or organisms. A specific chemical profile can be queried in admetSAR using either the CAS registry number, the common name, or structure similarity. Finally, 22 qualitative classification and 5 quantitative regression models are included allowing to estimate ecological/mammalian ADMET properties for novel chemicals (Cheng et al., 2013).

In a recent a recent statement, the Scientific Panel on Food Additives and Nutrient Sources added to Food (ANS Panel) of EFSA has assessed the new scientific information on the food colouring Allura Red AC (EFSA ANS Panel, 2013) that became available since their previous opinion (EFSA, 2009d) using read-across methods. The assessment dealt particularly with the positive findings from an *in vivo* Comet assay in mice. These findings were interpreted together with all the available relevant data from genotoxicity testing, metabolism and carcinogenicity, and in consideration of possible species differences between mouse and rat. The Panel considered the overall relevant data available in a read-across exercise, not only for Allura Red AC but also for a number of other structurally related sulphonated mono azo dyes authorised as food additives, namely: Amaranth, Ponceau 4R, Sunset Yellow FCF, Tartrazine and Azorubine/Carmoisine (EFSA ANS Panel, 2013). The Panel concluded that the new data were insufficient to change the conclusions of the 2009 opinion and that the read-across exercise highlighted a shared pattern of effects for this class of substances that would need

further investigation. The Panel recommended that *in vivo* Comet assay in mice should be performed for all the sulphonated mono azo dyes, in compliance with the most recent and internationally validated experimental protocol, using whole cells and examining a wide range of tissues (EFSA ANS Panel, 2013).

A generic application of the TTC includes to the categorisation of chemicals based on their level of concern for oral systemic toxicity (Cramer classification scheme), and to the prediction of the potential for genotoxicity using *in silico* tools. These *in silico* tools include the widely used software tool Toxtree (Patlewicz et al., 2008; Lapenna and Worth, 2011), which is freely available (at th://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/qsar_tools/toxtree and Sourcetoxtree.sourceforge.net/). A recent specific example includes the use of the TTC approach for the risk assessment of Alternaria toxins by the CONTAM Panel of EFSA since no toxicological data were available (EFSA CONTAM Panel, 2011).

4. OMICs: principles and application to human hazard assessment, strengths and limitations

The term '**OMICs'** refers to a broad field of studies in biology, ending in the suffix '-omics', such as transcriptomics, proteomics and metabolomics, and associated 'bioinformatics' (US-EPA, 2002). The OMIC technologies are rapidly developing in life sciences and their application to toxicology and ecotoxicology is one of the promising methodologies for evaluation and estimation of chemical risks (OECD, 2009a). Potential applications of OMICs in risk assessment applied to the food and feed area has been reviewed elsewhere (Pielaat et al., 2013).

The US-EPA defines **genomics** as **'**the study of all the genes of a cell or tissue, at the DNA (genotype), mRNA (transcriptome), or protein (proteome) level'. The main difference between genomics and genetics is that genetics scrutinises the functioning and composition of the single gene whereas genomics addresses all genes and their interplays in order to identify their combined influence on the growth and development of the organism. In the 1990s, it has been recognised that molecular biology methods (e.g. Northern blotting, RNAse protection assays, S1 nuclease analysis, plaque hybridisation, slot blots) did not provide sufficient throughput to effectively tackle genomic issues. The development of new methods in the late 1990s such as differential display, high-density filter hybridisation, serial analysis of gene expression, and cDNA- and oligonucleotide-based microarray 'chip' hybridisation (microarrays) has provided new solutions to allow the monitoring of expression levels of thousands of genes simultaneously (Pietu et al., 1999). In toxicology, the term 'toxicogenomics' has been originally coined by Nuwaysir et al. (1999) and refers to the integration of the genomic technologies with bioinformatics as an alternative means to study underlying MoA/AOP of chemicals and a way to potentially address challenges that are difficult to overcome by conventional toxicology methods. This section provides a short overview of the principles of transcriptomics, proteomics and metabolomics with a number of examples of existing potential applications in human hazard assessment of chemicals and their potential for future developments.

4.1. Transcriptomics

4.1.1. Principles of transcriptomics

Transcriptomics deal with the expression level of mRNAs in a given tissue, organ or other cell population, using DNA microarray and other high-throughput technologies that can estimate the quantities of mRNAs (NRC, 2007).

The transcriptome is the set of all RNA molecules, including mRNA, rRNA, tRNA, and other non-coding RNA (e.g. microRNA-transcriptional and post-transcriptional regulation of gene expression), produced in one or a population of cells (Pietu et al., 1999). The term can be applied to the total set of transcripts in a given organism or to the specific subset of transcripts present in a particular cell type. The key aims of transcriptomics are to catalogue all species of transcripts, including mRNAs, non-coding RNAs and small RNAs, to determine the transcriptional structure of genes, splicing

patterns and other post-transcriptional modifications and to quantify the changing expression levels of each transcript during development and under different conditions (Wang et al., 2009).

Two main technologies are used for transcriptomics, namely oligonucleotide microarrays and next-generation sequencing.

Oligonucleotide microarrays (OM) technology is hybridisation-based which is most common approach used for gene expression profiling, it makes use of the information created by genome sequencing (www.genomesonline.org), and from the myriad of expressed Sequences Tags (ESTs) using the first generation Sanger sequencers. Hybridisation-based approaches are high throughput and relatively inexpensive, except for high-resolution arrays that interrogate large genomes. Today, it is possible to design an array of oligomer probes that covers the whole transcriptome of any organism for which the genome sequence is known and the possible open reading frames and gene models have been identified using well-established bioinformatics analysis pipelines. However, these methods have several limitations, including their dependency on prior knowledge of genome sequence, high background levels caused by cross-hybridization and a limited dynamic range of detection. Moreover, inter-experimental expression level comparison is often difficult and requires complicated normalisation methods (Metzker, 2010).

Next Generation Sequencing (NGS) technologies can deliver fast, high-throughput, inexpensive and accurate genome information, including genomic and epigenomic sequencing. NGS include methods for determining the sequence content and abundance of mRNAs, non-coding RNAs and small RNAs (collectively called RNA–seq) and methods for measuring genome-wide profiles of immunoprecipitated DNA–protein complexes (ChIP–seq), methylation sites (methyl–seq) and DNase I hypersensitivity sites (DNase–seq). A key feature is the ability to sequence the whole genome of many organisms and it has allowed large-scale comparative and evolutionary studies to be performed (Metzker, 2010). In addition, the entire transcriptome can be queried, down to an individual base, whether or not a reference genome is available (McGettigan, 2013). This is illustrated with the recent publication of the genome of 1 092 individuals from 14 human populations constructed using a combination of low-coverage whole-genome and exome sequencing as part of the 1000 Genomes Project. In addition, NGS also allow the genome-scale mapping of epigenomic modifications important for transcriptional control, including DNA methylation and covalent modifications of histone proteins. Several large-scale analysis techniques are available that enable the survey of DNA methylation status at nucleotide resolution throughout the genome. NGS platforms for genome and epigenetic techniques are discussed elsewhere (Metzker, 2010). Overall, NGS is likely to replace OM because of their greater accuracy that closely matches quantitative polymerase chain reaction (PCR) and enable gene-expression studies in organisms for which OM are not available. Finally, they are likely to offer a higher throughput compared with microarrays as new developments will likely allow for the analysis of thousands of transcriptome samples in a single sequencing run (Sturla et al., 2014). However, the technology is limited by artefacts and biases that still need to be fully identified and controlled for (McGettigan, 2013).

Analysis of transcriptomic data requires a combination of statistical techniques, bioinformatic tools and databases. The huge amount of data produced by NGS platforms requires powerful information technology tools for data storage, tracking and quality control and data processing. Datasets are transformed using standardisation, normalisation or scaling in order to be able to compare measurements within and between studies. The challenge is to turn the large data sets with relatively high amounts of noise and without obvious biological/toxicological meaning into relevant findings. Advances in bioinformatics and algorithms have recently been reviewed, with focus on state-of-the-art techniques to support experimental scientists in analysing transcriptomic data (Berger et al., 2013). A number of methods for transcriptomics data analysis and interpretation exist and include: mathematical clustering algorithms (e.g. hierarchical clustering), K-means clustering and self-organising maps, and calculation of a measure of similarity between gene profiles. Clustering creates subsets of similar sequences and enables to select, amongst thousands, the sequences with biologically relevant characteristics. Multivariate statistical methods include Principal Component Analysis (PCA)

and Partial Least Squares (PLS). PCA is an unsupervised method which determines intrinsic structure within data sets, without prior knowledge, and that is used to calculate similarity between large data sets, such as microarray measurements. PLS as principal component discriminant analysis are supervised methods that use additional information (biochemical, histopathological or clinical data) to optimise the discrimination between samples (Draghici et al., 2003). In addition, software tools are under development to enable in-depth analysis of any list of inter-related biological data (pathway analysis tools) and many databases are available (Davies et al., 2010). These databases include the early Protein Data Bank, US National Center for Biotechnology Information (NCBI) sequence data sets and the University of California, Santa Cruz Genome Browser164, ENCODE165 and modENCODE166 projects. Data sets are usually generated by different laboratories and can have different dimensionalities and organisation. In order to support formatting, storing and calibrating of datasets, there have been substantial efforts to analyse such databases and online analysis tools have allowed performing a number of integrative data analyses on genomic data (e.g. Galaxy, DAVID119, STRING, Cytoscape, mouseNET).

4.1.2. Application of transcriptomics in human hazard assessment

The uses of transcriptomics for human hazard assessment have been reviewed elsewhere and the general view is that accurate prediction of chemical toxicity with such technologies remains a challenge (OECD, 2013; Thomas et al., 2013a; US-EPA, 2013). Key historical issues include that OM technology and PCR data have generally been limited to single time points, thus providing only snapshot information. However, NGS has allowed broader applications which include information on MoA of chemicals, dose-response assessment, inter- and intra-species differences in TK and TD, *in vitro* to *in vivo* extrapolations, epigenetic mechanisms and toxicity of multiple chemicals. To illustrate this, four examples are presented below in relation to interspecies differences, benchmark dose modelling, epigenetic mechanisms, and combined toxicity of exposure to multiple chemicals (chemical mixtures).

Interspecies differences

The contribution of toxicogenomics in defining the MoA of selective acting compounds has been illustrated by studies carried out on the aryl hydrocarbon receptor (AhR) ligand 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). Nault et al. (2013a) have demonstrated, by comparing TCDD with other AhR ligands such as β-naphthoflavone and using DNA microarrays, that divergent gene expression occurred between different AhR ligands and between *in vitro* liver cells (mouse hepatoma Hepa1c1c7 cells) and the liver *in vivo* (C57BL/6 mice). The acute hepatotoxicity of TCDD in mice following a single dose was found to correlate with changes in gene expression, which in turn were correlated with hepatic TCDD levels (Kopec et al., 2013). A comparison between human, mouse, and rat primary hepatocytes showed that with TCDD, 495, 2305, and 711 orthologous genes were differentially expressed in human, mouse, and rat hepatocytes, respectively (Forgacs et al., 2013). However, in that study, only 16 orthologues were differentially expressed across all three species, demonstrating species-specific gene expression profiles of TCDD despite the conservation of the AhR and its signalling mechanism. Similar conclusions were made in a further study (Nault et al., 2013b) when the genome-wide hepatic gene expressions elicited by TCDD were compared *in vivo* between Sprague-Dawley rats and C57BL/6 mice. The functional analysis of the genes that were differentially expressed has identified different orthologues in the rat (nucleotide binding and acetyltransferase activity) in comparison with the mouse (steroid, phospholipid, fatty acid and carbohydrate metabolism). Transcriptomics can provide valuable data regarding not only on the MoA of selective acting compounds but also inform on interspecies differences in MoA.

Benchmark dose modelling

The possibility of deriving BMD and BMDL from transcriptomic data has recently been explored (Thomas et al., 2007, 2011, 2012, 2013a,b). In a first study, dose-response microarray data were analysed using BMD/BMDL calculations and gene ontology (GO) classification in the rat nasal

epithelium following acute formaldehyde exposure (Thomas et al., 2007). The authors matched gene expression patterns to associated GO categories and, from these, computed average BMD and BMDL values for each category. Using these results and comparing them to doses of formaldehyde exposure causing alterations of individual cellular processes, the authors showed that the BMD estimates for the GO categories related to cell proliferation and DNA damage were similar to those measured in previous studies using cell labelling indices and DNA-protein cross-links. Moreover, the BMD estimates were consistent with the BMD estimated for rat nasal tumours (Thomas et al., 2007). This approach was subsequently extended to two case studies, one comprising five chemicals (1,4-dichlorobenzene, 1,2,3-trichloropropane, propylene glycol mono-t-butyl ether, naphthalene and methylene chloride) that were positive in a 2-year cancer bioassay and that were tested in mice (Thomas et al., 2011, 2012), and a second case study in which six chemicals (1,2,4-tribromobenzene, 2,3,4,6-tetrachlorophenol, bromobenzene, 4,4'-methylenebis(N,N-dimethyl) benzenamine, hydrazobenzene and N-nitrosodiphenylamine) were tested in rats (Thomas et al., 2013a). Three of the latter six compounds were found positive in rodent carcinogenicity tests. Multiple dose levels were used for each chemical and four time points (5 days, 2 weeks, 4 weeks and 13 weeks), and the 13-week time point, were used in the rat and mouse studies respectively. The authors analysed target tissues for traditional apical cancer and non-cancer endpoints (e.g. histological and organ weight changes) and transcriptional changes using microarrays. The dose-response changes in gene expression were analysed using a BMD approach and the responses grouped based on either biological processes (Thomas et al., 2011) or signalling pathways (Thomas et al., 2012, 2013a). For chemicals with human exposure data, the transcriptional BMD values were also used to calculate MOEs. The transcriptional BMD values, when compared with those for the traditional non-cancer and cancer apical endpoints, showed a high degree of correlation for specific pathways (> 0.85). Many of the correlated pathways have been implicated in non-cancer and cancer diseases pathogenesis. The results demonstrated that transcriptomic changes in pathways can be used to estimate non-cancer and cancer points of departure for use in quantitative risk assessments and have identified potential toxicity pathways involved in chemically induced responses in rodents (Thomas et al., 2011, 2012, 2013a). Moreover, the authors showed that the correlation between the transcriptional BMD values for the most sensitive pathway and the apical BMD values (> 0.85) was relatively stable over time for both non cancer- and cancer-related endpoints.

Investigation of epigenetic effects of chemicals

Recent work in the area of genomics has highlighted the importance of the epigenetic control of gene expression. A key feature of epigenetics is that they define heritable changes that are superimposed on the genome in the absence of genome sequence variability (Supic et al., 2013). This regulation occurs at the level of DNA methylation or hydroxymethylation, post-translational histone modification and circulating miRNAs that inhibit mRNA translation or accelerate their degradation. Some attempts have been made to evaluate miRNA as a tool in the risk assessment of drug-induced organ injury (Antoine et al., 2013; Shi et al., 2013; Yokoi and Nakajima, 2013). However, at present the lack of a standard quantification method for miRNAs and the small number of confirmatory studies limit the use of miRNA biomarkers in risk assessment. An attempt to incorporate epigenetic transgenerational effects in chemical risk assessment was recently performed by Alyea et al. (2014) who compared several transgenerational studies on the fungicide vinclozolin from a comprehensive suite of dose-response data (NOAEL, reference dose, and human exposure estimates) for both conventional and epigenetic endpoints. Overall, the analysis revealed that vinclozolin transgenerational effects were demonstrated at a 100 mg/kg/day in rats which would be 40-fold and 80-fold higher than the overall LOAEL and NOAEL from rat guideline studies respectively around 80,000-fold higher than the reference dose and 1.2-million fold above human exposure estimates.

The authors concluded that additional research is needed to investigate the interplay between epigenetics and apical endpoints before epigenetics can be considered in human health risk assessment. Finally, the authors recommended focusing research to examine the potential causal relationships between epigenetic alterations and adverse apical endpoints and if such a causal

relationship can be demonstrated the dose response relationship should also be investigated (Alyea et al., 2014).

Recently, the use of zebrafish (*Danio rerio*) to investigate adaptive and adverse responses to chemicals in relation to global transcriptomic responses has been reviewed. The review highlights specific applications in the area of epigenetic effects (e.g. DNA methylation, histone modifications and micro-RNA expression) through the integration of high-throughput screening, OMICs techniques and bioinformatics leading to the discovery of AOPs (Williams et al., 2014).

Combined exposure to multiple chemicals

Transcriptomics may support the hazard characterisation of the combined toxicity of multiple chemicals through the analysis of individual gene expression changes and multivariate statistical analysis of such gene profile changes (Stierum et al., 2005). In a study, Padhi et al. (2008) exposed rats perinatally to the so called 'Northern contaminant mixture' (NCM), (methylmercury (MeHg), polychlorinated biphenyls (PCBs) and organochlorine pesticides (OCs)) separately or together with the goitrogen propylthiouracyl. The study was designed to reflect the blood contaminant profile of human populations in arctic Canada. Post-natal day 14 cerebellum global gene expression resulting from such exposures was investigated using high-density cDNA microarrays validated by quantitative PCR (qPCR) on a subset of 10 genes. A number of differentially expressed genes involved in a number of neural functions were identified between controls and treated animals (e.g. nerve cell differentiation, migration, myelination and synaptic transmission). The comparison of cerebellum gene expression profiles resulting from exposure to the NCM and its individual components in male and female pups revealed inter-gender differences in transcriptomic responses and that co-exposure significantly masked the effects of individual components on cerebellum gene expression.

Toxicity of the phytoestrogen genistein with the anti-androgenic food contaminant vinclozolin on the male reproductive tract and fertility was assessed in rats combining a number of standard reproductive toxicology end points together with testicular mRNA expression profiles using long OM. Overall, the endpoints for reproductive function (decreased sperm counts, reduced sperm motion parameters) were correlated with testicular mRNA expression profiles (Eustache et al., 2009).

4.2. Proteomics

4.2.1. Principles of proteomics

Proteomics 'deal with cell and tissue-wide expression of proteins encoded by a genome. After transcriptomics, proteomics is the next step in OMICs studies. It is more complicated than genomics because, while a particular genome is more or less constant, the proteins that are produced differ from one cell type to another and from time to time in the same cell type (OECD, 2009a)'. Merrick and Bruno have termed a distinct set of expressed proteins that distinguish between health, toxicity or disease as 'toxicity signature' (Merrick and Bruno, 2004).

The identification and quantification of the proteome remains one of the greatest challenges in the post-genomic era. The three key challenges faced in proteomics are caused by (i) the ca. 1,000,000 estimated proteins transcribed from only 20,000 human genes, (ii) the dynamic range from 6 orders of magnitude for the proteins present in a mammalian cell up to 10 orders of magnitude in body fluids and (iii) the constant dynamic flux of the proteome. This section will highlight principles of proteomics including technological tools used to study the proteome and illustrate application of proteomics in human hazard assessment of chemicals.

The main technologies currently applied for separating proteins from complex biological samples (cells, organs…) are gel-based (e.g. two-dimensional gel electrophoresis (2-DE) and gel-free techniques (e.g. liquid chromatography tandem mass spectrometry (LC–MS/MS)) (Cecconi and Zamo, 2011; Yu, 2011; Sabido et al., 2012; Rodriguez-Suarez and Whetton, 2013). These techniques can be used in combination because they focus on different subsets of proteins that are only partially

overlapping and, therefore, are complementary. The subsequent identification of proteins occurs by combining the separation methods with tandem mass spectrometry.

In 2-DE, the proteins in a sample are separated first by iso-electric focusing on an immobilized pH gradient gel under the influence of an electric field, then separated according to their molecular mass by SDS–polyacrylamide gel electrophoresis and visualised after staining. The main drawback of 2-DE is the high inherent variability of the technique, rendering comparison between protein samples very difficult. This problem has been reduced by the introduction of differential gel electrophoresis, which allows two proteins samples labelled with different fluorescent dyes prior to be run together on the same gel along with an internal standard labelled with a third dye. The gel is then scanned and the protein spots analysed by dedicated software. Identification of the individual proteins is done by protein spot picking from the gel, followed by protease digestion and identification by soft ionisation mass spectrometer (using electrospray ionisation or matrix assisted laser desorption ionisation (MALDI) from a target plate. Overall, 2-DE is an accurate and relatively easy technique to perform quantification of individual proteins, and remains currently the most used method of protein separation. However, despite the major improvement obtained with Difference Gel Electrophoresis (DIGE) technology, co-migration of different proteins on the gel and the inability to run hydrophobic (membrane) proteins and proteins with an extreme pI remain major difficulties. Moreover, 2DE and the individual spot analysis by mass spectrometry are time consuming.

As an alternative to gel-based protein separation, the analysis of complex protein samples by liquid chromatography-tandem mass spectrometry (LC-MS/MS) has gained popularity over the past ten years, as the so-called 'shotgun proteomics approach'. In this approach, the protein sample is first digested into a complex mixture of peptides that are separated by reversed-phase high-performance LC and analysed using soft ionisation techniques such as electrospray ionisation mass spectrometry (ESI-MS) or MALDI-MS. In contrast to 2-DE, shotgun proteomics can be used to study hydrophobic proteins as a high throughput platform, since thousands of peptides can be analysed simultaneously. However, the quantification of the individual proteins remains challenging. To this effect, a number of quantitation techniques involving metabolic and chemical labelling of the protein sample with stable isotopes and label-free approaches has been developed for LC–MS/MS-mediated proteome analyses (Rodriguez-Suarez and Whetton, 2013). In contrast to 2-DE and shotgun proteomics that have been used to analyse all the proteins present in a sample, targeted proteomics based on multiple reaction monitoring (MRM) or selective reaction monitoring (SRM) was developed to verify the identity of specific proteins of interest and to follow them with high throughput. MRM requires a triple quadrupole mass spectrometer, with the first quadrupole (Q1) selecting the known precursor ion, Q2 fragmenting the ion and Q3 monitoring the fragment ions (Meng and Veenstra, 2011). The immense data information provided by a single proteomic analysis necessitates specific tools for data processing to enable the extraction of key information on the identity and quantity of the detected proteins.

As discussed previously for transcriptromic data, analysis of proteomics data requires complex statistical methods, databases and bioinformatic tools. Three main families of regression-based methods currently being applied in the analyses of OMICS data: univariate approaches and associated multiple testing correction procedures, dimension reduction techniques, and variable selection approaches. These are reviewed elsewhere (Chadeau-Hyam et al., 2013). Recently, Knudsen et al. (2014) have published online a comprehensive open-source tool for merging multidimensional quantitative proteomics data into a common format ready for subsequent bioinformatic analysis.

Recent technological advances allow now > 2500 proteins to be detected in a single LC-MS/MS run with a dynamic range of 4-5 orders (Rodriguez-Suarez and Whetton, 2013). However, the robustness, reproducibility and mass accuracy of the technology will need to be further increased to face the technological challenge that 75 % of the proteome is present at < 5000 protein molecules per cell (Washburn et al., 2003). As a result, many low-abundance proteins, among which are often found the biomarkers of interest, remain un-analysed. A further problem is that the technologies in proteomics are not yet standardised between laboratories despite the formation of several consortia. This lack of

standardisation will present a major challenge to regulatory agencies whose task will be to help establishing standardised transparent methods to interpret 'omics' data.

Efforts to compare proteomic with transcriptomic data have met mixed success. It has become clear over the past decade that the biological dogma 'DNA → RNA → protein' is not only affected by complex relationship between the 20,000 genes from which an estimated 1,000,000 proteins are produced, but also the differences in dynamic processes between the events of transcription, translation and post-translational modification and, finally, protein turnover. While only a limited correlation between changes in individual mRNAs and the corresponding proteome is generally observed, several studies have found good correlations at the pathway level (Boitier et al., 2011; Collins et al., 2012). However, these correlations tend to be limited to chemical agents where biological or toxicological activity is receptor-mediated.

4.2.2. Application of proteomics to human hazard assessment

The past decade has seen a rapid increase in the application of proteomics in toxicology and hazard assessment, where it is often referred to as toxicoproteomics (Van Summeren et al., 2012). Applications of proteomics to human hazard assessment of chemicals are illustrated for the identification and quantitation of protein targets for *in vivo* (with tissues and biological fluids) and *in vitro* (cell cultures) studies, as well as for the identification of biomarkers of toxicity with the overarching aim to depict MoA/AOP.

Identification of protein targets for toxicants *in vivo*

A promising application of proteomics has been the identification of protein targets for toxicants *in vivo* to investigate MoA. For instance, Fountoulakis et al. (2000) found in a proteomic study in mice, that the levels of approximately 35 proteins out of 256 hepatic proteins (identified by DE) were modified after paracetamol treatment (Fountoulakis et al., 2000). In a related study by Koen et al. (2007), 33 new protein targets forming bromobenzene metabolite adducts in mouse liver, including glutathione S-transferases, protein disulfide isomerases and liver fatty acid-binding protein, were identified by 2-DE combined with MALDI-MS (Koen et al., 2007). In another study, protein samples from livers of rats treated 14 days with troglitazone were separated by DIGE and analysed by MALDI MS (Boitier et al., 2011). This study identified 55 proteins belonging mostly to the pathways of fatty acid metabolism, PPARα/RXR activation, oxidative stress and cholesterol biosynthesis, whose levels were up-regulated, and some to carbohydrate metabolism, whose levels were down-regulated. The same pathways were also identified to be affected at the transcriptomics level (Boitier et al., 2011).

In an integrated study evaluating gentamycin nephrotoxicity in rats (Com et al., 2012), proteomic analysis of the kidney after 14 days of treatment revealed that of the 2000 polypeptide spots detected by 2D-DIGE, 56 different down-regulated proteins and 49 different up-regulated proteins were identified by MALDI-TOF MS. The modulation in protein levels was suggestive of a mitochondrial dysfunction with impairment of cellular energy production, induction of oxidative stress, an effect on protein biosynthesis and on cellular assembly and organisation. In an attempt to access the proteome using non-invasive techniques, proteomics have also been applied to serum and urine. For instance, proteomics were used to identify proteins up- and down-regulated in rat plasma, in response to treatment with doses of cationic nanobubles that caused liver fibrosis and inflammation (Pan et al., 2012) and to identify multi-organ responses induced by paracetamol treatment of mice (Sun et al., 2013). Several studies have assessed the proteome in urine after treatment with toxic chemicals.

Identification of protein targets for toxicants *in vitro*

Attempts have also been made to use proteomics to study complex toxicological responses such as teratogenicity. Meganathan et al. (2012) applied proteomics to the study the *in vitro* effects of the teratogenic agent thalidomide on differentiation of human embryonic stem cells (Meganathan et al., 2012). The authors showed, with the help of genomic and proteomic expression patterns, the differential expression of limb, heart and embryonic development related transcription factors and

biological processes and the effect of thalidomide treatment on the levels of select proteins such as RANBP1 (a RAN GTPase binding protein mediating the translocation of RNA and proteins through the nuclear pore complex).

Discovery and validation of biomarkers of toxicity

The other key application of proteomics in toxicology is in the research and validation of biomarkers. Biomarkers provide important information on exposure, susceptibility and response to a chemical in biofluids, tissues or cell cultures. Therefore, a signature of specific molecular changes at the level of proteins is expected to accompany the development of toxicity. Proteomics present several major advantages over traditional biochemical or immunological approaches when applied to biomarkers (Amacher, 2010; Van Summeren et al., 2012). For instance, biomarker identification is greatly facilitated by MS-based techniques and the use of targeted proteomics based on multiple reaction monitoring (MRM) enables 30-100 candidate protein biomarkers to be simultaneously targeted and measured. A further advantage of MS-based proteomics is that it allows to distinguish multiple posttranslational variants of a protein and to quantitate them. It is therefore not surprising that major efforts have gone into initiatives such as the European InnoMed PredTox project, the Predictive Safety Testing Consortium, the Human Urine and Kidney Proteome Initiative and FDA's Critical Path Initiative for safety science with the aim to identify, develop and validate new biomarkers for preclinical safety evaluation. An example of the successful application of proteomics in biomarker discovery is the identification of glycine amidinotransferase and plasma retinol-binding protein precursor as novel potential biomarkers for nephrotoxicity (Com et al., 2012). Similarly, a recent evaluation of liver protein samples from rats treated with the hepatotoxic agent EMD 335823 separated by label-free LC-MS identified, using SRM, 48 putative liver toxicity biomarkers (Collins et al., 2012).

4.3. Metabolomics

4.3.1. Principles of metabolomics

The US-EPA has defined metabolomics: the evaluation of tissues and biological fluids for changes in metabolite levels that result from toxicant-induced exposure (US-EPA, 2004). The OECD refers to metabolomics as a discipline, which "deals with endogenous metabolite profiles of tissues or organs derived from mass spectrometry or nuclear magnetic resonance spectrometry analyses of plasma or homogenates. Metabolic profiling can give an immediate picture of the physiological state of the tissue (OECD, 2009a). More recently, Sturla et al. (2014) gave a more quantitative definition of metabolomics: 'metabolomics analyses in a comprehensive and quantitative manner all metabolites or low molecular weight organic or inorganic chemicals that are products or substrates of enzyme-mediated processes'.

The metabolome is composed of all the low molecular weight compounds (typically < 1500 Da) of endogenous nature, which are important modulators, substrates, by-products, and building blocks of many different biological processes such as endogenous metabolites (amino acids, carbohydrates, lipids, etc…). It also includes the exogenous metabolites such as drug metabolites, xenobiotics and contaminants living organisms may be exposed to, and in this case has been referred to as the 'exposome' (Wild, 2005, 2012). Metabolomics thus offer a powerful tool for discovering the functional status of an organism and elucidating the consequences of internal (genetic mutations, diseases) and external (environment, food composition, xenobiotics) perturbations. For example, the accumulation of a specific metabolite may either signal an AOP (activation of a toxicity pathway) or the optimisation of a biosynthetic pathway (anabolism).

Currently, two complementary approaches are used in metabolomics: the targeted approach and the non-targeted approach.

The targeted approach, also called metabolic profiling in some instances, enables to perform quantitative analysis (relative abundances and concentrations) of specific sets of metabolites such as

biomarkers of toxicity, or substrates, and products of known enzymatic reactions (Illig et al., 2010; Menni et al., 2012). A key disadvantage is that the method is not applicable to identify new metabolites because the exact structure of the metabolite(s) needs to be known so that they can be used as standards for quantitation (Wishart et al., 2007; Bouhifd et al., 2013). On the other hand, this allows a systematic collection of data with good quality control for toxicological studies and hazard assessment.

The untargeted approach is often referred to as 'global metabolome analysis' or 'metabolic fingerprinting' and provides a comprehensive view of all metabolites in a biological sample. Metabolites different in their relative quantitation and annotation (i.e. MS/chromatographic/ spectroscopic peaks), are generated and the initial raw data may produce thousands of signals. The main applications of this approach are investigation of MoA, hypothesis generation and identification of biomarkers. However, a key weakness is that untargeted metabolomics provides only a relative quantification and some significant metabolites are not always identifiable (Bouhifd et al., 2013).

Analytical techniques to identify endogenous and exogenous metabolites in metabolomics include Nuclear Magnetic Resonance (NMR) and MS, coupled to separation techniques, or using direct flow injection. NMR has a number of advantages: non-discriminating and non-destructive technique, minimal requirements for sample preparation and versatile technique which can be used in high-throughput fingerprinting context and for the analysis of biological fluids (serum, urine etc.) and intact tissues (for example, tumours) The major weakness of NMR spectroscopy is that it is orders of magnitude less sensitive than MS. Modern MS now enables to measure compounds in the femto-molar to atto-molar range. Coupling MS with LC or GC allows the measurement of hundreds of individual species within a single sample. The combination of mass accuracy and real-time tandem MS, along with increasingly comprehensive databases, can automate the identification of metabolites in a routine manner.

The identification of metabolites chemical structure and their quantitative analysis provide information that can be interpreted in the light of biochemical pathways and metabolites causing group segregation in the fingerprinting approach need to be identified using quantitative methods. This will allow both metabolic fingerprinting and profiling. In order to do that, annotation of the metabolome is important and over the past years, the metabolomics community has made efforts and progress towards more robust approaches for such annotation. Thus, the development and enrichment of databases of known metabolites and the accessibility to a few on-line data processing and annotation workflow may now facilitate the tedious work of pre-processing, statistical analysis, and annotation of metabolomics studies (Patti et al., 2012; Wishart et al., 2013). Although these approaches can be very powerful to reveal metabolites of interest, the validation of these biomarkers, particularly for small molecules reflecting metabolic processes, remains essential (Koulman et al., 2009).

Data analysis in metabolomics requires a pre-processing step, a normalisation step and a statistical analysis step. The pre-processing step consists of alignment of spectral data in order to obtain a matrix of the characteristic features of the samples within a batch of analyses. This can be complex and time demanding especially in the case of high-resolution LC-MS (Boccard et al., 2010). Normalisation of the data aims at making the data comparable as much as possible to allow the quantification of signals detected in multiple samples. Recent methods for data normalisation include optimal selection of multiple internal standards (Sysi-Aho et al., 2007). Data analysis is the final step in the schematic work-flow of metabolomics and a number of multivariate statistical tools are available and the selection will depend on the experimental objectives or the type of question investigated including PCA, partial least square discriminant analysis (PLS-DA) and orthogonal PLS-Discriminant Analysis (OPLS-DA). PLS-DA is a method applied to both discriminating or classifying a set of samples in a metabolomic context, and OPLS-DA has been increasingly used to identify metabolites of potential biochemical significance with a graphical output to allow visualisation/discrimination of metabolites (Wiklund et al., 2008). A full review of statistical methods to analyse OMIC data including metbsolomics is included in Chadeau-Hyam et al. (2013).

Guidelines for reporting data, experimental conditions as well as harmonisation regarding ontology and common semantics for metabolomics are needed to facilitate the exchange of information and the re-use and interpretation of results across several scientific disciplines. Through the Metabolomics Standards Initiative (MSI - http://msi-workgroups.sourceforge.net/), the Metabolomics Society has encouraged the scientists to apply standards for reporting on their metabolomics experiments and studies (Fiehn et al., 2007; Sumner et al., 2007). This Core Information for Metabolomics Reporting (CIMR) specifies minimal guidelines and seeks in the long term to cover all application areas, analytical technologies, biological context metadata, chemical analysis and data processing, as well as formats for exchange of data and the ontology.

Recent metabolomic databases include the human metabolome database (HMDB) which contains ca. 2500 metabolites, 1200 drugs and 3500 food components encountered in the human body (Wishart et al., 2007), the MetaMap®-Tox developed at metabonomics and containing rat plasma metabolome for more than 500 references compounds (Kamp et al., 2012; Mattes et al., 2013). Finally, the EU research project COordination of Standards in MetabOlomicS (COSMOS) (http://cosmos-fp7.eu) developed the MetaboLights dataset://www.ebi.ac.uk/metabolights/) as a central repository of metabolomics experiments (Haug et al., 2013). MetaboLights is a cross-species, cross-technique database, which also covers metabolite structures, their reference spectra as well as information on the biological roles of the metabolites, their locations and concentrations, and the experimental data from the metabolic studies uploaded in this repository (Steinbeck et al., 2012). Such efforts provide existing e-infrastructures to report metabolomic studies, make them available to a broader community so that the data can also be used as standards for the interpretation of future metabolomic studies.

4.3.2. Applications of metabolomics in hazard assessment

Overviews and application of metabolomics in a regulatory context have been published (Bouhifd et al., 2013; Ramirez et al., 2013). Examples of the applications of metabolomics are provided below regarding MoA/AOP and toxicokinetic aspects using *in vivo* studies, predictive models using *in vitro* methods and *in vivo* studies of combined exposure to multiple chemicals (chemical mixtures).

In vivo studies and MoA

An increasing number of studies investigating MoA/AOP using metabolomic biomarkers of toxicity *in vivo* are being published. Montoya et al. (2014) have developed a database (MetaMap®Tox) based on a rat plasma metabolome consisting of approximately 300 endogenous metabolites. Male and female Wistar rats were treated with > 500 reference compounds over a period of 28 days and more than 120 specific toxicity patterns of common metabolite changes associated with unique MoAs were established. The authors applied the results to predictive direct/indirect adverse effects on the thyroid. Animals were treated using compounds acting either directly on the thyroid function (e.g. methimazole, ethylenethiourea) or indirectly on the thyroid (e.g. induction of liver enzyme inducers by agents such as aroclor 1254 and boscalid leading to an increased excretion of thyroid hormones). The authors identified metabolites in plasma, which were commonly regulated irrespective of whether the effect on the thyroid was indirect or direct. For example, direct thyroid hormone synthesis inhibitors affected enzymes in the urea cycle, increased the ω-oxidation of fatty acids and decreased glutamate and oxoproline levels whereas indirect thyroid hormone inhibiting compounds interacted with the lipid mediated and liver metabolism (Montoya et al., 2014).

Toxicological and metabolomics studies of 3-chloropropane-1,2-dipalmitate (3-MCPD dipalmitate) were carried out based on an acute oral toxicity test, a 90-day feeding test, and on ultra-performance LC-MS analysis. The results of the 90-day feeding test in male Wistar rats showed that 3-MCPD dipalmitate caused a significant increase in blood urea nitrogen and creatinine differences analysed by PLS-DA of the chromatographic data. Renal tubular epithelium cell degeneration and renal tubular hyaline cast accumulation were the major histopathological changes in rats administered 3-MCPD dipalmitate. The combination of histopathological examination, clinical chemistry and metabolomics analyses in rats resulted in a systematic and comprehensive assessment of the sub-

chronic toxicity of 3-MCPD dipalmitate and provided specific metabolomic markers of toxicity (Li et al., 2013a). Another study investigating the metabolomics of dimethoate toxicity in rats showed alterations in the excretion of a number of endogenous metabolites (e.g. L-tyrosine, dimethylthiophosphate, dimethyldithiophosphate, citric acid, uric acid, suberic acid, glycylproline, allantoin, isovalerylglutamic acid) reflecting perturbation in liver function, antioxidant and nervous systems, as well as the metabolisms of lipids, glucose, fatty acids, amino acids, and collagen in rats (Feng et al., 2012).

In vivo metabolomics and toxicokinetics

Recent applications of metabolomics have included the investigation of toxicokinetic processes. In mice, the use of stable isotope- and mass spectrometry-based metabolomics to underpin the metabolic routes and effects of the drug tempol has been recently published. PCA of the urinary metabolomics data separated tempol metabolites versus endogenous metabolites, which had been altered by the tempol treatment (Li et al., 2013b). In humans, endogenous metabolic markers of hepatic CYP3A, activity, the major human CYP isoform in the liver and the gastrointestinal (GI) tract have been investigated in 24 healthy subjects: CYP3A substrate (midazolam), inhibitor (ketoconazole) and inducer (rifampicin). Metabolomic analyses supported the development of a predictive model for CYP3A activity using midazolam as a probe substrate and a combination of concentrations and ratios of several endogenous metabolites (Shin et al., 2013).

Predictive *in vitro* methods

Many *in vitro* metabolomic studies have also been published with the aim of developing predictive models. For example, the use of human embryonic stem cells combined with the use of LC-MS analysis, as alternative models to identify potential developmental toxicants, has been highlighted recently. In this study the authors demonstrated correlation between teratogenicity and changes in the ratio of arginine to asymmetric dimethylarginine (greater than 10 %) (West et al., 2010; Kleinstreuer et al., 2011). The authors built a predictive model and validated the predictability of the model for eight teratogenic drugs (West et al., 2010).

Combined toxicity of multiple chemicals

Recently, a number of *in-vivo* metabolomics studies have been used to study the toxicological effects of combined exposure to multiple chemicals. For example, the acute renal toxicity of melamine and cyanuric acid in rats has been investigated combining several analytical techniques and endpoints based on metabolomic markers of kidney damage and compared with histopathology results. In the future, these metabolomic markers could be used to model BMDLs to derive a HBGV for the melamine-cyanuric acid mixture and compared to traditional histopathological endpoints(Xie et al., 2010; Kim et al., 2012; Schnackenberg et al., 2012; Sun et al., 2012). In the pesticide area, recent *in vivo* toxicity studies investigating the long-term effects of combined exposure to multiple organophosphate pesticides (dichlorvos, dimethoate, acephate, and phorate) showed global disturbance in lipid metabolism, tricarboxylic acid cycle and oxidative stress (Du et al., 2013). The use of metabolomics coupled with multivariate statistical tools to address polyTK in the context of multicomponent pharmaceuticals or nutraceuticals has been discussed. A research framework has been proposed to integrate the dynamic concentration profile of bioavailable xenobiotics (e.g. *in vivo* absorption, hepatic and gut microbial metabolism.) as well as the human metabolic response profile (Lan et al., 2013).

4.4. Strength and limitations of OMICs technologies

The era of OMICS technologies opened great opportunity to characterise MoA/AOP of chemicals for a number of endpoints addressing different levels of biological organisation (DNA, mRNA, proteome and metabolome level). In the future, it is foreseen that the integration of such OMIC technologies will provide endpoints to quantify key events associated with AOPs for chemically-induced adverse effects

in humans. However, during the development and refinement of the different OMICs technologies, scientists have soon realised that the 'holy grail of prediction' was not yet available.

A self-evident strength of the application of the OMICs technologies to toxicology and hazard assessment is the possibility to measure experimentally comprehensive biochemical profiles of the modifications occurring during an AOP. Pattern recognition in biochemical signals from transcriptomic, proteomic, and metabolomic experiments will lead to the identification and validation of the biomarkers relevant to monitor defined toxicity in humans and animals (Ellinger-Ziegelbauer et al., 2011; Peng et al., 2013). It is foreseeen that these biomarkers can then be used for dose response modelling such as BMD/BMDL modelling to then derive POD for hazard assessment purposes (see transcriptomic examples and Thomas et al., 2007, 2011, 2012). As illustrated above, the combined results from OMIC technologies can potentially produce a comprehensive profile of the molecular events leading to an AOP and allow testing interspecies differences in toxicity and their human relevance (Burgess-Herbert and Euling, 2013). In addition, OMICs allow addressing the relevance of human variability and particularly genetic polymorphism in response to chemical exposures. This is already an important research field in the pharmaceutical area with many clinical applications. The investigation of human variability provides a great opportunity in the food safety area to better integrate inter-individual differences in: a) TK processes for age differences, inter-ethnic differences and polymorphisms in phase I enzymes (CYP, alcohol dehydrogenase, epoxide hydroxylase, esterases), phase II enzymes (UDP-glucuronyltransferases, sulphotransferases glutathione-s-transferases and methyl-transferases) and transporters (OATPs, OCTs, P-glycoproteins, MRPs), b) TD processes for age differences and inter-ethnic differences in toxicity targets (e.g. receptors, ion channels, enzymes…) (Ozdemir et al., 2009; Squassina et al., 2010; Dorne, 2010; US-EPA, 2013). Last but not least, OMICs can provide ways to investigate patterns of gene transcripts, proteins, and metabolites using *in vitro* methods and how these are associated in an AOP. In the long run, this type of experiments may provide helpful means to validate ITS using mechanistic *in vitro* assays to reduce animal studies and move towards predictive modelling (Basketter, 2012).

OMICs technologies have also a number of weaknesses. Major weaknesses of OMICs are: a) the need for a complex arsenal of new molecular techniques, analytical tools and highly specialised training, b) the need for sophisticated bioinformatic tools to analyse the innumerable datapoints that are generated, c) the difficulty to interpret and validate the thousands of signals generated during OMICs experiments especially. Another key weakness relates to the sensitivity of the methodologies which may lead to the detection of changes that may not be biologically or toxicologically relevant. Finally, OMICs studies have a complex design and have been most often conducted with well known reference substances. This has allowed researchers to correlate the OMICs datasets with results from standard methodologies such as clinical chemistry or histopathological endpoints. It is foreseen that in the future, publicly available databases combining *in vitro* and *in vivo* OMIC datasets for large amount of compounds with MoA/AOP knowledge will help considerably to move towards identifying biomarkers associated with specific AOPs and to bring new tools for predictive toxicology.

5. Prioritisation of chemicals, systems toxicology and future outlook on chemical risk assessment

5.1. Prioritisation of chemicals

There is wide international recognition that new approaches and frameworks are needed to evaluate the safety of large numbers of chemicals in food, consumer products and the environment. Many international efforts aim to develop methods for the prioritisation of chemicals using ITS so that, when a concern has been identified, the chemicals can be considered either to prioritise chemicals based on their hazard and exposure profile for decision making or further testing. This is a key aspect of EFSA's work in food and feed safety and ECHA's work under the REACH regulation. Recent reviews have proposed a number of methodologies for this purpose including the Nextgen project of the US-EPA, the IPCS/WHO applications of the MoA and the technical report of EFSA on 'the identification

of emerging chemical risks in the food and feed chain' (Thomas et al., 2013b; US-EPA, 2013; EFSA, 2014; Meek et al., 2014).

5.1.1. Prioritisation of chemicals at the US-EPA

The Toxcast research program and the Nextgen report of the US-EPA has proposed a number of approaches for the prioritisation and ranking of chemicals according to their toxicological properties. These approaches include the Toxicological Prioritisation index (TOXPi) decision support framework and a tiered approach to prioritise chemicals for further testing (Reif et al., 2010; Gangwal et al., 2012; Thomas et al., 2013b; US-EPA, 2013).

The ToxPi decision support framework has been originally developed from the results of ToxCast research programme (see Section 2.3.1) under the U.S. EPA's Endocrine Disruptor Screening Program which has screened over 300 pesticides and environmental contaminants for their potential to affect the endocrine systems of humans and wildlife. In order to facilitate the rationale prioritisation of chemicals for further evaluation, Reif et al. (2010) proposed aToxPi score which incorporates data from *in vitro* assays, chemical descriptors, biological pathways. These ToxPis provide a flexible, ranking of each chemical's potential endocrine activity and focusing on estrogen, androgen, and thyroid pathways, putative endocrine profiles were first defined to then derive a relative rank or score for the entire ToxCast library. Recently, a ToxPi visualisation tool (the ToxPi graphical user interface ToxPi GUI) has been developed to integrate the relative contribution of all information sources, including hazard and exposure information for a particular chemical, to an overall priority ranking. A software tool based on this method and the ToxCast Data is available and described in Reif et al. (2013). More recently, Gangwal et al. (2012) proposed the use of ToxPi decision support tool to enable the integration of multiple sources of evidence or toxicity and incorporate exposure surrogates. The results are then transformed into transparent visual rankings to facilitate decision making and prioritise chemicals for further testing. The approach taken by Gangwal et al. (2012) highlights the utility of the ToxPi framework for incorporating exposure information to rank chemicals and improve understanding of key exposure surrogates. However, this analysis has been performed for relatively data rich compounds and demonstrate the need for further studies to understand the relationship between simple exposure surrogates, tiered screening-level exposure assessments, and population-level biomonitoring data (Gangwal et al., 2012).

Thomas et al. (2013b) proposed a tiered approach for the prioritisation of chemicals which allows the integration of results from new methodologies and standard toxicity testing in a flexible manner. This framework has evolved from tiered approaches developed previously to address regulatory mandates, to prioritising and assessing large numbers of chemicals (Meek et al., 2011; Thomas et al., 2013b). The framework is data-driven, can be iteratively refined as knowledge MoA/AOP and models become available, and involves successive tiers of testing using the MOE as the primary metric (Thomas et al., 2013b). Tier 1 aims to use data to prioritise and screen chemicals for immediate regulatory decision, for further testing in Tiers 2 and 3, or in some cases, to add to the weight of evidence in Tier 2 and 3 assessments. This would be particularly relevant with respect to identifying pathways or molecular signatures associated with chemical-induced diseases (US-EPA, 2013). Five components are proposed for tier 1: (1) use of HTS assays to separate chemicals into selective and nonselective modes of action; (2) *in vitro* genotoxicity assays to separate potential genotoxic and non-genotoxic chemicals; (3) IVIVE TK models to convert *in vitro* assay concentrations to applied doses; (4) high-throughput exposure modelling to estimate human exposures to chemicals; and (5) calculation of MOE for the chemical (Thomas et al., 2013b). The second tier of the proposed framework also consists of 5 components: (1) short-term *in vivo* transcriptomic studies to identify a transcriptional point of departure (POD) or RP values for chemicals with a non –selective MoA; (2) *in vivo* studies to identify POD values for chemicals with a selective MoA; (3) IVIVE TK studies to link internal and applied dose; (4) refinement of human exposure estimates; and (5) calculation of a MOE for the chemical. The rationale for identifying chemicals with selective and non-selective MoA has been developed through the analysis of the relationship between transcriptomic profiles and apical responses from histopathological studies. For a number of non-cancer and cancer responses, POD have been identified

and transcriptomic BMDL have been successfully modelled (see Section 4.2, Thomas et al., 2007, 2011, 2012, 2013a). The third tier 3 would be conceptually equivalent to the traditional toxicological *in vivo* testing in experimental animals. In practice, Tier 3 can be applied to chemicals with a high volume of production and for which significant potential for human exposure would occur. Alternatives include specification based on understanding of the toxicological profile acquired in lower tier testing (tier 1 and 2) studies. This could include rodent cancer bioassays, developmental or reproductive toxicity studies. The authors anticipated that, depending on the MOE cut-off values chosen, the majority of chemicals would be screened out in tier 1 or 2 leaving an estimated 3 % to 15 % of chemicals requiring such *in vivo* tier 3 assessment. Still, testing itself for these chemicals can be prioritised by endpoints (MoA/AOP driven) based on the results from both tier 1 and 2 studies (Thomas et al., 2013b).

It is worth noting that the proposed tiered approach described above can be potentially applied to the risk assessment of combined exposure to multiple chemicals using a hazard index approach and dose addition as the default assumption. In the case of interactions of TK and/or TD nature, these hazard indexes can be modified to take into account the magnitude of the interactions, which can be based on either TK data (decrease or increase elimination such as clearance ratios) and/or TD data (e.g. BMDL ratios) (EFSA, 2013).

5.1.2. Application of the WHO/IPCS mode of action framework

The IPCS/WHO has proposed applications of the MoA framework to the prioritisation of chemicals using endocrine disruption potential as an example. This application refers to 'how best to focus on chemicals that are most likely to cause adverse effects without empirically testing all chemicals of regulatory concern'. An expert (QSAR) system, which had already been developed to predict oestrogen receptor binding affinity using MoA/AOP knowledge is described (OECD, 2009b; Schmieder et al., 2003, 2004; US EPA, 2009. In this case, the AOP starts with a MIE, which is direct binding of the chemical to the oestrogen receptor. The authors discussed that such an event can be tested through the development of two *in vitro* assays based on the trout as a model. The first assay can measure the interaction of the chemical with the oestrogen receptor using a competitive binding assay. The second assay can measure the consequences of oestrogen receptor activation or inhibition as a result of tissue uptake and partitioning of the chemical after xenobiotic metabolism using a trout liver slice assay. Another example discussed include the development and use of alternative (*in vitro*) assays to target particular cellular or physiological key events along a specific toxicity pathway. Once the MoA has been established, the key event data can be used for read-across from other chemicals. If a new chemical fits the established MoA, this existing knowledge can be used to justify a more efficient testing strategy, so that not every chemical needs to be evaluated in an *in vivo* test (Meek et al., 2014).

5.1.3. EFSA technical report on the identification of emerging chemical risks in the food and feed chain

EFSA has recently published a technical report presenting a systematic framework for the identification of emerging chemical risks that may occur in the food and feed chain and which may have a direct or indirect impact on human, animal and/or plant health. Such exposure may arise from industrial chemicals that are either intentionally or non-intentionally, produced (contaminants)as well as from certain natural contaminants that may be transferred to the food/feed chain through the environment. The framework uses a number of data sources as input, relating to the source of the chemical (industrial chemical, contaminant) and software models as tools to predict the environmental behaviour and potential toxicity of chemicals from structural features and physico-chemical properties (e.g. QSAR models and PB-TK models). The application of the framework consists of a multi-step selection process initiating with a list of chemicals to which a sequence of selection criteria is applied to identify the substances of potential concern. The selection criteria take into account a number of parameters including volumes of production or import data related to the chemical, its environmental persistence, bioaccumulation potential, dispersive uses, toxicity, and any available outcomes of previous risk assessments. The procedure has two main entry points either for industrial chemicals

registered under REACH Regulation or for substances consistently detected in the environment with a subset of more specific entry points depending on specific objectives and relevant data availability. Further work has been recommended to further test this methodology through: 1) Consideration of additional data sources and selection criteria, 2) Development of databases and software to apply efficiently and systematically the inclusion/exclusion criteria, which characterise the different selection steps of the proposed methodology, 3) Application of this methodology to the systematic identification of emerging chemical risks through the food/feed chain (EFSA, 2014a).

5.2. Systems toxicology and future perspectives for chemical risk assessment

5.2.1. Systems toxicology and integrated testing strategies

There are many definitions of systems biology that differ between different international bodies. **Systems biology** has been defined by the OECD as the 'Study of the mechanisms underlying complex biological processes as integrated systems of many diverse, interacting components' (OECD, 2013). The NIH has defined systems biology as 'an approach in biomedical research to understanding the larger picture - be it at the level of the organism, tissue, or cell - by putting its pieces together'. Systems biology is in stark contrast to decades of reductionist biology, which involves taking the pieces apart. From the definition of systems biology, **systems toxicology** aims to identify toxicity pathways and potentially predict toxicity. In other words, systems toxicology is the means to depict AOPs through the integration of the knowledge at different levels of biological organisation using ITS (see Section 2.1). Recently, Sturla et al. (2014) discussed that systems toxicology aims at 'decoding the toxicological blueprint of active substances that interact with living systems' and 'should allow exploring how biological components function as a network in cells, tissues and organisms'. The authors discussed the development of dynamic AOP models suggesting that such dynamic AOPs will enable the simulation of the population-level effects of an exposure as 'the ultimate goal of Systems Toxicology'. Three steps are proposed to develop the dynamic AOPs following a top down approach: 1. Development of causal computable biological network models, that link the systems interaction of a toxicant with the organ-level responses. Such models can then be used to quantify the biological impact of an exposure in the context of quantifiable endpoints (e.g. histology or physiological measurements). 2. Development of mechanistic knowledge derived from quantitative measurements and dynamic models linking the exposure with the organ-level responses. 3. Representation of the link between exposure and population outcome using mathematical models enabling the simulation of population-level effects of the exposure (Sturla et al., 2014).

Overall, such systems approaches in toxicology take advantage of the historical developments of modern biology such as the human genome project and advances in ITS such as the one developed and explored in the TOX 21 and SEURAT programme and discussed previously, such as HTS assays, OMIC technologies, physiologically-based models and *in silico* tools. In addition to these tools, alternative species (i.e. non-mammalian species) provide *in vivo* models for identifying hazards, integrating dose-response effects, and understanding pathways and apical effects useful for assessing chemical risks to humans and to other species. The shorter life spans of alternative species enable the evaluation of toxicity over the full life span of the intact organism, facilitating the study of the entire aetiology of disease from the MIE to apical outcomes, including more complex phenomena such as birth defects or neurobehavioral impairment. Alternative species studies are progressively playing a more integral role in hazard assessment of chemicals for humans and the environment (ECHA, 2013; Perkins et al., 2013; Villeneuve et al., 2014). Both the European Chemicals Agency (ECHA) and the US-EPA use alternative species tests as part of required tests for endocrine disruptors (US-EPA, 2013 Scholz et al., 2013).

Truong et al. (2014) examined all 1078 ToxCast phase 1 and 2 chemicals (1060 unique chemicals) for developmental and neurotoxicity in the embryonic zebrafish using a rapid *in vivo* approach. Each chemical was tested using broad dose ranges spanning 4 orders of magnitude (6.4 nM to 64 μM) with multiple replicates (n = 32) at each dose. Twenty-two endpoints were simultaneously evaluated and distinct toxicity patterns in response to chemicals were identified. The author then pursued a

concordance analysis of the phase I chemicals tested in Toxcast to test the complementarity of the developmental zebrafish outcomes with the *in vitro* outcomes in xenobiotic metabolism and CYP inhibition assays, and developmental rat or rabbit maternal and pregnancy studies. Concordance of xenobiotic-related *in vitro* assays and morphologically abnormal embryonic zebrafish was concluded in a number of cases (Lieschke and Currie, 2007; Goldstone et al., 2010; Santoriello and Zon, 2012; Scholz, 2013). For the chemicals and endpoints lacking concordance with ToxCast Phase I results, the authors concluded that these results would indicate toxicity pathways or chemical classes that would require more attention in future phases of testing. Hence, the authors propose to integrate the developing zebrafish into the existing *in vitro* HTS assays for the hazard assessment of chemicals. They proposed to use the zebrafish as the 'tier 1' of the hazard identification process where all chemicals are assessed and all those with potential to cause adverse effects will be further screened in the battery of *in vitro* tests and evaluated in the predictive models already developed in the TOX-21 programme. Having a whole-organism system as the first tier provides the ability to detect endpoints that may be missed in a screen using *in vitro* assays, such as metabolism and pathway sensors (Truong et al., 2014).

From a global risk assessment perspective, the systems toxicology approach needs the exposure dimension to complete the picture. Recently, the three non-food committees of the European Commission have emphasised that there is paradigm shift moving from a hazard-driven process to one that is exposure-driven. Exposure assessment is beyond the scope of this document, however, combining toxicity data in a systems toxicology perspective requires the integration of the relevant exposure data (external dose) into an internal dose (TK) to then relate it to the MoA/AOP (TD). For a particular compound, the translation of an external dose to an internal dose will require absorption and biovailability data as well another ADME data such as half-life and clearance. Finally, exposure can also be investigated from a systems perspective as a complete entity: the exposome. The **exposome** refers to the totality of environmental exposures from conception onwards, and has been proposed to be a critical entity for disease aetiology (Wild et al., 2013). The exposome approach is increasingly used by epidemiologists for genome-wide association studies (GWAS), in order to investigate diseases, while relying on questionnaires to characterise 'environmental' exposures. In addition, assessment of the exposome is now facilitated using OMICs technologies and analytical techniques, which are able to measure multiple chemical residues and multiple biomarkers of exposure and effects (e.g. adductomics to measure DNA adducts) (Rappaport and Smith, 2011; Wild et al., 2013).

5.2.2. Future perspectives for the human risk assessment of chemicals

Overall, this report has highlighted a number of new and emerging methods to depict TK and TD processes using a MoA/AOP approach for hazard assessment of chemicals including PB-TK and PB-TK-TD models, *in silico* (QSAR, read-across and TTC) and OMICs (transcriptomics, proteomics and metabolomics). These methodologies can provide a way to bring mechanistic thinking into toxicity testing and give quantitative insights on key issues in TK and TD for hazard assessment, and reduce animal use in toxicological testing. *In vivo* and *in vitro* examples of application of these modern tools to the hazard assessment of chemicals for humans have been highlighted for a number of fundamental issues in hazard assessment: interspecies differences, human variability in TK and TD processes, *in vitro* to *in vivo* extrapolation and combined exposure to multiple chemicals both using traditional *in vivo* standard toxicity combined with *in vitro* techniques, PB-TK, PB-TK-TD models, *in silico* tools. Key recommendations of the three non-food committees of the European Commission and the Nextgen project of the US-EPA are presented and some perspectives on the future chemical risk assessment are given.

5.2.2.1. Recommendations from the three non-food committees of the European Commission

The three non-food committees of the European Commission published a joint opinion on new challenges in risk assessment with specific recommendations (SCCS, SCHER, SCENIHR, 2013). The committees specifically included the new methods described in this report. A key conclusion and recommendation refers to the need to combine these methods into ITS based on Weight-of-Evidence

methods that integrate independent sources of information and information on MoA (Boobis et al., 2008; SCCS, SCHER, SCENIHR, 2013). Key requirements for such a shift include:

- The need for new *in vitro* methods. These should have similar properties to *in vivo* counterparts and allow for testing over longer periods of time (sub-chronic to chronic) to establish clear relationship between *in vitro* endpoints and adverse effects *in vivo*. These *in vitro* systems should also reflect *in vivo* TK.

- New endpoints: Sensitive measurement methods are needed to allow studies to be made at exposure levels that reflect likely human exposures and OMICs are likely to play progressively a key role in such a development.

- MoA should become the central point of a future risk assessment, and this knowledge should be considering the OMICs technologies; particular focus is needed on how they might improve MoA understanding and how they may themselves benefit from such MoA knowledge. The MoA information is also essential for the assessment and the prediction of chemical interactions in mixtures.

- A tiered approach for risk assessment is recommended to use resources in the most efficient way and limit unnecessary animal testing. The tiered approach combines hazard and exposure for individual stressors.

- Comprehensive, validated and up-to-date databases are needed to develop a new paradigm for risk assessment. The most important include databases on effects of various stressors in humans, monitoring data of human exposure to various stressors, extension of the TTC database, and inclusion of MoA/AOP for each type of adverse effect.

- Validation of SARs or QSARs and read-across approaches

5.2.2.2. Challenges ahead and recommendations from the US-EPA NextGen report

The Nextgen report of the US-EPA has reviewed new approaches and frameworks to identify biological patterns and MoA/AOP associated with specific diseases. Such patterns facilitate the grouping and the evaluation of chemicals based on mechanistic understanding of specific diseases (US-EPA, 2013).

Key challenges and data gaps for these new and emerging methods and tools for human risk assessment of chemicals. Such data gaps need to be filled to incorporate new information into risk assessment frameworks:

(1) metabolism of test compounds cannot currently be predicted;

(2) need for an understanding of the biology from a systems perspective;

(3) evaluate methods available to measure key aspects of the biological space across multiple levels of organisation (from cellular level to organ and level of the individual);

(4) availability of the relevant data through the implementation of a knowledge infrastructure.

Recurrent and problematic issues were identified as including problem formation, classifications of adversity and WoE, dose-response modelling (particularly at the low-dose end), human variability including differences in TK, TD, life stages, diseases, nutrition, interspecies differences in TK and TD and consequences for intra-species extrapolation, risk assessment of multiple chemicals (chemical mixtures) and characterisation of uncertainty. These issues have been explored in a number of 'prototype' case studies on specific chemicals for which different level of knowledge were available

including molecular information, *in silico*, toxicological studies in animals and epidemiological evidence in humans (US-EPA, 2013).

Next steps have been identified to facilitate the incorporation of these new methods and the data generated from such methods in risk assessment and the decision-making processes:

- More cases studies on chemicals, are needed to test the incorporation of HTS toxicity data and other novel data types (e.g. data from HTS assays from ToxCast and TOX-21 OMICs…). This will also allow to inform the risk assessment process and demonstrate the added value of such new tools and identify further scientific gaps. Validation of HT toxicity testing schemes and development a framework for such validation is also recommended since the traditional 'validation' schemes do not address this gap. Plans are in place to develop criteria for systematic review of new types of data, disease signatures, adequate weight of evidence for use in risk assessment, and new approaches for risk assessment.

- Using new data types to guide development of NexGen approaches on challenging questions such as population-level risks (using traditional and molecular biology data), with an emphasis on epigenomics and influences of broadly defined environmental factors. Application of these new methods might also better inform our understanding of the combined effects of multiple stressors (multiple chemical exposures, diet, stress, and pre-existing disease).

- Using tier 1 screening and prioritisation approaches for chemicals (Thomas et al., 2013b; see Section 5.1). Results should then be fedback into the testing paradigm for its refinement.

- Develop toxicity values for each tier. Toxicity values informed by new methods will be developed in each tier. This will allow to address needs from screening chemicals for future testing to assessment for potency/category of adverse effects.

- Expand stakeholder discussion and peer review. Levels of confidence in these toxicity values will be characterised depending on the types/quality of the supporting data. New assessments will receive public comments and peer review.

- Collaboration between US-EPA with other national and international agencies involved in risk assessment, testing, and research. This will allow to coordinate and harmonise activities, and to improve data collection, analyses, curation, sharing, and warehousing.

5.2.2.3. Future perspectives

This report has highlighted the shift towards a MoA/AOP approach in chemical risk assessment to depict TK and TD processes using ITS, including *in vitro* methods based on human cells (e.g. HTS assays), OMICs, physiologically-based models and *in silico* tools. This paradigm shift will allow to move towards new approaches for the safety evaluation of chemicals, reduce animal use in toxicity testing and provide support for the prioritisation of thousands of chemicals.

Within the coming year, it is foreseen that the AOP initiative of the OECD will develop 18 AOP and 3 case studies that are applicable to both human health and environmental risk assessment (see Section 2.3). These AOPs will be of great values to screen chemicals according to specific or non-specific MoA. These AOPs may be of qualitative nature to start with but will increasingly move towards a quantitative understanding, which will also provide a basis to move towards a systems toxicology approach and, further down the line, develop dynamic AOPs. However, the OECD recognised that a key gap in the AOP development is the fact that, currently, TK information (ADME) and PB-TK models are out of the scope of the AOP development and will have to be addressed. In this context, the lack of TK information from human cells is a key aspect that needs to be fulfilled in order to identify the ADME of a chemical and of the extent to which metabolism results in bioactivation or detoxification. This will support the integration of TK and TD processes in the MoA/AOP framework

as well as give a starting point for further HTS assays for toxicity testing. Indeed, HTS assays do not have metabolic capacities so that if the metabolite(s) of a compound is/are the toxic form(s), measurements of toxicity may be biased. In addition, such knowledge of human metabolism will provide a basis to investigate interspecies differences in TK and, for known compounds, to test the biological relevance of the test species.

In the future, it is foreseen that HTS assays to investigate TK will need to be designed to allow incorporation of human variability (genetic polymorphisms, subgroups of population). Key issues throughout this report raised the need for validation of *in vitro* methods and OMIC technologies (analytical methods and statistical approaches), physiologically-based models and *in silico* tools. As discussed by the US-EPA, in parallel to such validation, more case studies should be developed to test these methods combining new knowledge and historical data for proof of concept. The need for publicly accessible databases integrating data from these new methods/tools is also a key issue since in the future it can be foreseen that, as knowledge advances, risk assessors and toxicologist will be able to refine models and tools (e.g. dynamic AOPs, complex cellular network models). Finally, methodologies for weight of evidence and uncertainty analysis to integrate data from such new methodologies in the MoA/AOP framework, including testing the biological relevance (species concordance analysis) and report uncertainties in a transparent way, are needed. These future perspectives are further discussed below to provide detailed recommendations.

The human microbiome, which refers to the community of microorganisms that live in or on the human body, is another key aspect that needs to be taken into account to move towards a systems toxicology approach. The Human Microbiome Project (HMP) of the National Institute of Health (NIH) is currently investigating the role of human microbiota and analysing its role in human health and disease. The HMP is currently sequencing the genomes of a number of microorganisms isolated from the human body as well as samples of digestive tract, mouth, skin, nose, and female urogenital tract of human volunteers. These genomes will then be considered for metagenomic analysis. The HMP has opened new horizons for studying how the composition and functional variations of the microbiome affect drug action, fate, and toxicity (pharmacomicrobiomics) particularly in the human gut. The gut microbiome is the most predominant and most diverse microbial community residing in the human body with hundreds of species and its contribution and influence on xenobiotic metabolism is substantial. The integration of OMICs methodologies has provided very useful means to elucidate the microbiome's influence on chemical metabolic profiles through DNA sequence-based phylogeny and metagenomics (Hood, 2012). Recently, the toxicity of a number of chemicals on microbiota has recently been demonstrated from metagenomics and metabolomics analysis for contaminants such as cadmium and arsenic (Liu et al., 2014; Potera, 2014). Ishii et al. (2012) demonstrated that enteric bacteria induce the expression of CYP3A in mouse liver. Recent studies have begun to identify the key events of the regulation between the gut microbiota and its host but the underlying molecular mechanisms of host-microorganism interactions remain largely unknown. In addition, findings obtained from the study of animal models remain to be translated to humans and a potential caveat is that microbiota members differ not only among host species but also between individual host organisms (Sommer and Bäckhed, 2013). However, progresses made in the development of genetic tools, such as whole-genome sequencing, and in the availability of novel genetic models will allow for dissecting the interplay between the microbiome, host genetics and host physiology. Combining these tools for further studies in the upcoming years will greatly deepen our understanding of the molecular targets in the homeostatic interaction between the gut microbiota and the host and thereby their global impact on chemical metabolism (Sommer and Bäckhed, 2013). To conclude, the human microbiome highlights another important source of high inter-individual differences in the metabolism and toxicity of chemicals that will need to be taken into account to move towards a systems toxicology approach.

6. Recommendations for future activities at EFSA on new and emerging tools for human hazard assessment of chemicals in the food and feed safety area

The section below provides recommendations resulting from a consultation of EFSA panels and staff dealing with chemical risk assessment and other experts from international bodies (ECHA, OECD, WHO…).

6.1. Terminology and general considerations

6.1.1. Terminology

Harmonisation of terminology

A major aspect of EFSA's science strategy 2012-2016 is the further development of internationally harmonised risk assessment methodologies including terminology (EFSA SC, 2012). With regards to new and emerging methods and tools for hazard assessment, harmonisation of the terminology and definitions are needed at the concept level (MoA, mechanism of action, AOP, key events, molecular initiating key events, systems toxicology, …) and for each of the tools available (physiologically-based models, *in silico* tools and OMICs).

Harmonisation of terminology also applies beyond the human context, which is the focus of this report, to animal health and environmental risk assessment. As discussed previously in the context of the scientific report dealing with combined exposure to multiple chemicals, harmonisation of the terminology is not straightforward because different communities may use different terminology/interpretation for the same concept, or the same tool or the same terminology for a different concept/tool. This can be exemplified by the concepts of biovailability in toxicokinetics and mode of action, which have different meanings for the human and the ecological risk assessment communities (EFSA, 2013).

6.1.2. General considerations

Animal health and ecological risk assessment

Further work on viewing the potential use of these modern tools and methodologies for chemical risk assessment in the context of animal health and ecological risk assessment is recommended. In the short-term, these reviews (including the current document) could be the starting point of developing (a) guidance(s) to apply these new and emerging methods, in a context-dependent manner and using tiered approaches, and the WoE approach which would be valuable for EFSA. The development of a guidance for the use of the WoE approach in risk assessment including distinct lines of evidence (*in vivo*, *in vitro*, *in silico*, population studies etc.) has been identified by the Scientific Committee as a priority topic for EFSA (EFSA, 2013).

Guidance on the use of mode of action in risk assessment

A specific recommendation relates to the development of a guidance on the use of mode of action in chemical risk assessment, particularly in relation to the recent new developments highlighted by the WHO (criteria for species concordance analysis and human relevance, modified Bradford Hill criteria, use of epidemiological data, data from recent new and emerging methods (*in vitro* and high throughput screening tests, physiologically-based models, AOP developments, OMICs)). Weight of evidence approaches should be explored further to consider the integration of such multi-level and complex information for hazard assessment and risk assessment as a whole (e.g. integration of hazard and exposure data for risk characterisation). In this context, data needs (and the likelihood of getting such data), biological relevance and statistical aspects would need particular attention for WoE approaches.

Risk assessment of chemical mixtures and multiple stressors

Finally, a possible activity in the longer term would be the development of methodologies to apply these new tools to the risk assessment of exposure to multiple chemicals as well as multiple chemicals combined with other stressors (e.g. biological hazards, physical agents…). WoE approaches would have to be considered including biological relevance and statistical aspects. This recommendation has already been formulated in the context of the EFSA scientific report on combined exposure to multiple chemicals, the EFSA colloquium on bee health and the recent scientific report on integrated risk assessment of multiple stressors in bees (EFSA, 2013, 2014b).

6.2. Physiologically-based models and *in silico* models

6.2.1. Investigating toxicokinetics

The first key recommendation is the need for human TK data in hazard assessment to better understand interspecies differences, human variability. Such TK data will ultimately link exposure, internal dose and toxicity using physiologically-based models for risk assessment purposes.

Such TK is needed to:

- understand the relevance of test species to the human situation from a TK point of view (e.g. evolutionary conservation of enzymes and their respective isoforms) parallel to the investigation of species differences in TD (e.g. receptors, signalling pathways…);

- design sound physiologically-based models integrating species differences in TK and TD and/or human variability in TK for the hazard assessment in metabolic, excretion and transport pathways;

- provide a scientific basis to set Assessment Groups based on TK for multiple chemicals particularly when the metabolic route is a key event (bioactivation to a toxic metabolite or TK interactions such as inhibition of cytochrome P-450). Criteria to set these assessment groups using TK data would also need to be considered using a WoE approach including consideration of interspecies differences and human variability (relevance of the metabolic route in test species to the human situation, availability of human data on metabolism (*in vitro*/*in vivo*). This recommendation has already been formulated in the context of the scientific report on combined exposure to multiple chemicals (EFSA, 2013).

Improvement of *in vitro* methods for generating TK data

A key to generating TK data in humans is the improvement of current *in vitro* methods to measure human absorption (bioavailability, …), distribution (volume of distribution, protein binding, hepatic extraction), metabolism (e.g. Vmax, Km, inhibition constants…) for phase I (cytochrome P-450, esterases…) and phase II (UDP-glucuronyl transferases, glutathione-s-transferases), isoforms involved in gut metabolism versus hepatic metabolism, transporters (e.g. transport via P-glycoprotein, Organic Anion Transporter Proteins (OATP),….), and excretion of chemicals.

6.2.2. Physiologically-based models

Further exploration of the use of physiologically-based models in chemical risk assessment is recommended, namely:

- to develop a guidance on the use of physiologically-based models in chemical risk assessment. This includes toxicokinetic and toxicodynamic models incorporating data from standard *in vivo* assays and alternative methodologies (*in vitro* methods and *in silico* data (QSAR, read-across, TTC)). The guidance could explore, through tiered approaches, the relevance and needs for such models in a context-dependent manner (data-poor chemical specific situation, prioritisation, data-rich chemical specific situation, combined exposure)

- to develop prototype physiologically-based models using specific case studies to integrate exposure (external dose), internal dose and TK information and toxicity data for hazard assessment purposes. These models can also be used to refine uncertainty factors used in hazard assessment (categorical or

chemical-specific) as recommended in the scientific report on combined exposure to multiple chemicals.

- It is recommended to develop relatively simple models that may refine the link between exposure (external dose), basic TK data (internal dose), and toxic effects in the short term. For example, case studies could be explored to develop models for single compounds and binary mixtures based on *in vitro* and *in vivo* data. In the mid-term, as knowledge advances, a full exploration of full physiologically-based toxicokinetic models and physiologically-based toxicokinetic-toxicodynamic models that would integrate more complex quantitative knowledge can then be explored and implemented (e.g. inter-species differences, human variability in TK and TD, epidemiological data, *in vitro* models, *in silico* models). Finally, it is worth highlighting that the data used to build the models and their associated uncertainty should be described and analysed in a transparent manner to optimise their use and ensure reproducibility.

- A practical need to further develop such physiologically-based models is the need for databases providing critical parameters such as physicochemical properties, biological and physiological, toxicokinetic and toxicity variables (body weight, age, ventilation rate, Vmax, Km, clearance, bioavailability, half life, AOPs), and bioinformatic tools/algorithms, to analyse and integrate the data.

6.2.3. *In silico* tools

Further work is needed to explore application of *in silico* tools in chemical risk assessment. This will allow to use currently available databases comprising vast amount of physicochemical and toxicological data and validate the available predictive models to reduce animal use. In addition, it will provide the opportunity to explore the applicability domain of the predictive methods and their degree of specificity. It can be foreseen that the domain of applicability of such tools will be encompassing human health, animal health, and ecological risk assessment.

Development of a framework for systematic and harmonised approach for the use of *in silico* tools (SAR, QSAR, read-across) is recommended. It is proposed to further explore their use as potential tools to 1. support the hazard identification of genotoxic compounds by building batteries of models based on structural alerts, toxicity data and existing databases, 2. design physiologically-based models, 3. elucidate the mode of action (including toxicity pathways) for the prioritisation of chemicals. A key aspect of these applications is the need to compare the currently available (Q)SAR tools in a transparent way to allow optimisation and calibration of the models.

Further development is proposed for the read-across methodologies in terms of further investigations into their use in the hazard assessment of chemicals, particularly to integrate (Q)SAR and physicochemical properties with TK and TD data (potency estimates, AOP) using specific chemicals as case studies for 'proof of concept'. This can also be useful to explore category-approaches for prioritisation of chemicals, especially for data-poor substances (e.g. flavourings, emerging contaminants…) using for example the OECD QSAR toolbox or the ADMET-SAR tool.

Potential refinements of the TTC approach for hazard assessment have been previously discussed by the Scientific Committee of EFSA in their recent TTC opinion (EFSA SC, 2012). The key recommendations can be highlighted as: 1. Re-evaluation and update of the Kramer classes and toxicological databases to improve accuracy, applicability, and availability of *in silico* models. 2. Development and refinement of models for the prediction of TK (bioaccumulation in humans, and quantitative simulation of metabolite degradation/formation) and TD (genotoxic potential, carcinogenic potency) with, as far as possible, an understanding of MoA.

6.3. OMICs

Validation of OMIC technologies for their use in human hazard assessment

Further work is needed on the validation and standardisation of OMIC technologies for their use in human hazard assessment. Detailed guidance on the criteria that are needed for their acceptability will support this activity: data needs, statistical and bioinformatic methodologies for data analysis, mechanistic and statistical aspects, biological and toxicological relevance, relevance of *in vitro* OMICs data to the *in vivo* situation.

Case studies relevant to the food safety area

It is recommended to further explore the use of OMICs in human hazard assessment using case studies relevant to the food and feed area to investigate: mode of action, epigenetic mechanisms, *in vitro* to *in vivo* extrapolation, interspecies differences, human variability for both single compounds and exposure to multiple chemicals. Key aspects include: 1) Exploration of the derivation of reference points such as benchmark doses and their limits based on transcriptomic and/or proteomic and/or metabolomic data, with case studies. These should include comparison of the benchmark doses and their limits at different time points/study lengths. 2) Exploration of the use of *in vitro* OMICs data for the ranking and prioritisation of chemicals. 3) Use of biomarkers of effects and exposure generated from OMIC technologies in subgroups of the human population to incorporate human variability in hazard assessment of chemicals. 4) Applications of OMICs in other areas of chemical hazard assessment are recommended including animal health risk assessment, ecological risk assessment.

In the nutrition area, a number of general recommendations can be formulated: 1) Exploration of the use of OMICs in both Hazard and Benefit Assessment. 2) Use of OMICs in Health claims: evaluation of microbiome/metagenomics in gut and systems resilience. 3) A mechanistic OMICs approach to depict mode of action can be applied to novel food safety.

6.4. Prioritisation of chemicals, systems toxicology and the future of chemical risk assessment

Integrated testing strategies and MoA, prioritisation of chemicals

Future work on integrated testing strategies is recommended using case studies of specific chemicals to investigate both MoA from the toxicokinetic and toxicodynamic processes (*in vitro* methods to depict AOPs, QSAR, physiologically-based models..). This will also allow identifying data gaps and research needs. Further investigations are needed on the use of ITS to differentiate, as much as possible, chemicals with receptor or pathway driven specific MoAs, chemicals with multiple or non-specific MoA. This may provide a basis to rank chemical potencies for prioritisation.

The use of alternative test species (including fish species) as a bridge between *in vitro* methods and mammalian tests should be further explored and historical data from the literature could be used in a systematic review/meta-analysis.

In practice, further exploration of new methodologies for hazard assessment are needed for both regulators and industry and these include: 1) Screening of large sets of chemicals to group them based on hazard data (toxicokinetics, toxicity (potency)..). In order to prioritise chemicals, further data for compounds of concern/industry can focus on candidates/use information for risk assessment. 2) Assessment of chemicals for a specific purpose to fill in data gaps for risk assessment or for dossier submission.

For exposure to multiple chemicals, a better understanding of MoA/AOP of multiple substances using predictive and alternative methodologies (including *in vitro* models, QSAR, physiologically-based models and OMICs) is needed for regulated and active substances such as pesticides and contaminants. This will allow improving the basis for setting Assessment Groups.

Future of chemical risk assessment

Three recommendations for the future of chemical risk assessment can be formulated.

1. New flexible risk assessment frameworks for chemicals and applications

Explore the use of new risk assessment frameworks to bring the perspective of systems toxicology in chemical risk assessment using case studies. In this context, applications include the integration of AOP data for the prioritisation of chemicals, the development of dynamic AOP and the integration of the impact of the human microbiome on TK and TD events. A key aspect of AOP development is to explore the use of TK data, since TK information (ADME) and PB-TK models are currently out of its scope. For risk assessment another key aspect is to integrate the exposure dimension with the TK data.

2. Weight of evidence and uncertainty analysis

Weight of evidence and uncertainty analysis methodologies for the integration of data from new methodologies in the MoA/AOP framework and chemical risk assessment as a whole are needed. This includes testing the biological relevance of the MoA/AOP (species concordance analysis, severity of the effect) and reporting uncertainties in a transparent way.

3. International Collaborations

In order to integrate these new methods in new frameworks for risk assessment and explore them through case studies and facilitate international harmonisation, reinforcing collaboration with international institutions such as ECHA, the JRC, the OECD, the WHO, the NTP and the US-EPA is critical for EFSA and highly recommended.

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

- Over the last decade, a number of *in vivo, in vitro* and *in silico* methodologies and tools have been developed to investigate the mode of action (MoA) or Adverse outcome pathway (AOP) leading to human adverse health effects resulting from chemical exposure. These modern methodologies and tools provide two key opportunities to risk assessors dealing with human hazard assessment of chemicals: 1. to move towards a mechanistic understanding of toxicity taking into account both toxicokinetic (TK) and toxicodynamic (TD) processes for hazard assessment, 2. to reduce animal use in toxicological research (3Rs: reduce, replace, refine). These modern methodologies are reviewed in this scientific report to present their potential use in the future of human hazard assessment of chemicals with a view to anticipating their future use within EFSA's work.

- Currently, MoA information are not often available for specific chemicals, and risk assessors rely often on dose response assessment to translate external dose to a quantitative reference point for hazard characterisation in test species. However, recent international developments support the move towards elucidating such MoA/AOPs for human hazard assessment of chemicals. These include the new applications of the WHO framework on MoA, the OECD international programme on AOPs and the TOX-21 and the SEURAT-1 research programmes in the US and Europe, respectively, both dealing with alternatives to animal testing such as *in vitro* methods and other integrated testing strategies (ITS). Strengths of ITS such as high throughput screening (HTS) assays include the possibility to screen and prioritise chemicals in a single experiment and in a cost effective fashion while minimising animal testing. However, they have a number of limitations which include their lack of prediction for a) chemical-

induced disease associated pathways, b) metabolism, c) interactions between different cell types, d) tissue-level cellular interactions, e) chronic exposure.

- Methodologies and tools to investigate TK processes (absorption, distribution, metabolism, excretion of chemicals (ADME)) in humans include a number of *in vitro* systems. The current updated OECD Test Guideline 417, dealing mainly with absorption and metabolism, indicates that *in vitro* testing using human cells can provide supplemental TK information which may substantially reduce *in vivo* animal testing. Even though these *in vitro* models have still received little attention in hazard assessment of chemicals for the food safety area, they can provide key information on absorption, biovailability, protein binding and the identification of human transporters and metabolic pathways such as efflux transporters (phase 0 and phase III) and phase I and phase II enzymes. These parameters can be used to determine the *in vivo* hepatic clearance of a chemical and then be scaled up to the whole liver and take into account human variability to build physiologically-based (PB) models for both TK (PB-TK) and TD (PB-TK-TD) processes. A critical challenge that remains to be solved in order to apply these *in vitro* methods routinely is the *in vitro* to *in vivo* extrapolation (IVIVE) to reflect human physiology and metabolism (hepatic an extrahepatic such as intestinal metabolism). Ideally in the future, IVIVE would also incorporate human variability in quantitative IVIVE (QIVIVE).

- PB-TK models provide a quantitative means to address TK processes and are therefore very useful tools in hazard assessment. PB-TK-TD are more complex than PB-TK since they link both the TK and the TD dimensions. The use of such models has been recommended by regulatory authorities around the world such as the US-EPA and the WHO which have both highlighted the need to develop guidance to pursue common principles for their application in chemical hazard assessment and risk assessment as a whole. However, reservations have been formulated regarding their routine use since they require a) detailed knowledge of TK for a particular chemical which is not often available for both models and detailed knowledge of TD for PB-TK-TD models, b) high levels of expertise and resources, c) the need to validate the models. Consequently, they are mostly used in high-tier assessment. The development of such models using ITS, IVIVE and QIVIVE also remains a big challenge for both TK parameters and toxicity parameters. These models have been applied to the food safety area particularly to pesticides, contaminants and food contact materials and have provided very useful tools to investigate key issues in hazard assessment such as interspecies differences, human variability, biomonitoring programmes, combined exposure to multiple chemicals and *in vitro* to *in vivo* extrapolations.

- I*n silico* tools include (quantitative) structure activity relationships (Q)SARs) and read-across methods for which a large number of models and databases have been developed around the word to predict a number of toxicological properties of chemicals such as the OECD QSAR toolbox. Another tool that is increasingly used in hazard assessment and risk assessment as a whole is the threshold of toxicological concern (TTC). QSARs are typically used in combination with other non-testing methods (e.g. read-across) and testing methods (e.g. *in vitro*) in the context of ITS and Weight-of-Evidence (WoE) assessments. Overall, (Q)SARs and read-across methods are increasingly predictive for hazard assessment particularly for acute toxicity, mutagenicity, genotoxicity and bioacummulation. However, their predictability for TK properties (ADME) and sub-chronic and chronic toxicity is still limited and considerable research is undergoing in this area. In addition, an increasing number of Q(SAR) models and databases are available and their precision, specificity and sensitivity may vary and would need to be evaluated and validated. Finally, it is foreseen that the combination of the results from different Q(SAR) models, structural alerts, read-across estimates as well as *in vitro* and *in vivo* toxicological studies using a WoE approach will improve the utility and the validation of these tools and increase the overall reliability of *in silico* methods.

- Key OMICs technologies to investigate TK and TD processes include transcriptomics, proteomics and metabolomics. OMICs technologies are valuable tools to measure biochemical

changes associated with a MoA/AOP, identify biomarkers in humans and animals for dose response modelling, investigate interspecies differences and their human relevance and incorporate human variability (age differences, inter-ethnic differences, polymorphisms). Finally, OMICs technologies can also investigate patterns of gene transcripts, proteins, and metabolites within an AOP using *in vitro* models and provide helpful means to validate ITS using mechanistic *in vitro* assays to reduce animal studies and move towards predictive modelling. Weaknesses of OMICs methods include the need for complex molecular and analytical techniques, highly specialised training and sophisticated bioinformatic tools to analyse huge datasets. Another key issue relates to the sensitivity of the methodologies which may lead to the detection of changes that may not be biologically or toxicologically relevant. Finally, OMICs studies have a complex design and have been most restricted to well known reference substances to allow researchers to correlate OMICs datasets with classical toxicological endpoints (clinical chemistry, histopathological endpoints). It is foreseen that in the future, publicly available databases combining *in vitro* and *in vivo* OMIC datasets for large amount of compounds with MoA/AOP knowledge will help considerably to identify biomarkers associated with specific AOPs and to bring new tools for predictive toxicology. Applications to human hazard assessment of chemicals in the food safety area have already been explored and include benchmark dose modelling from transcriptomic profiling, investigation of epigenomic mechanisms, identification of biomarkers of toxicity (proteomics), and investigation of MoA for single and multiple compounds (metabolomics).

- A number of approaches have been developed for the prioritisation and ranking of chemicals according to their toxicological properties. At the US-EPA, the toxicological prioritisation index (ToxPi) decision support framework has been developed from the results of the Tox cast research programme and enable to rank chemicals through the integration of multiple sources of evidence on toxicity and exposure surrogates. ToxPi also includes a graphical user interface which allows visualising the relative contribution of each information sources (toxicity, exposure, uses…) in the overall priority ranking. Future needs of ToxPi include further studies to understand the relationship between simple exposure surrogates, tiered screening-level exposure assessments, and population-level biomonitoring data. Another prioritisation tool from the US-EPA is the tiered approach developed during the NextGen project. In practice, tier 1 aims to prioritise and screen chemicals using ITS (Toxcast HTS assays, in vitro genotoxicity tests, IVIVE TK models…) for further testing in Tiers 2 and 3. Tier 2 includes limited *in vivo* toxicity testing (e.g. short-term *in vivo* transcriptomic studies, *in vivo* studies to identify a point of departure for chemicals with a selective MoA; and IVIVE TK studies to link exposure and internal dose). Tier 3 is equivalent to the traditional toxicological *in vivo* testing in experimental animals.

- For the identification of emerging chemical risks, EFSA has recently developed a systematic framework which uses a number of data sources as input, relating to the source of the chemical (industrial chemical, contaminant) and software models as tools to predict the environmental behaviour and potential toxicity of chemicals from structural features and physico-chemical properties (e.g. QSAR models and PB-TK models). The application of the framework consists of a multi-step selection process starting with a list of chemicals to which a sequence of selection criteria is applied to identify the substances of potential concern. The selection criteria take into account a number of parameters including volumes of production or import data related to the chemical, its environmental persistence, bioaccumulation potential, dispersive uses, toxicity, and any available outcomes of previous risk assessments. Further work is ongoing to test the methodology using additional data sources, selection criteria and through the development of databases and software, for the systematic identification of emerging chemical risks in the food chain.

- This report has highlighted the shift towards a MoA/AOP approach in chemical risk assessment to depict TK and TD processes using ITS including *in vitro* methods based on human cells (e.g. HTS assays), OMICs, physiologically-based models and *in silico* tools. This

paradigm shift will allow moving towards a systems toxicology view for human hazard assessment of chemicals and will allow reducing animal use in toxicity testing and providing support for the prioritisation of thousands of chemicals.

- Key issues discussed throughout this report include the need for validation of *in vitro* methods, OMICs technologies (analytical methods and statistical approaches), physiologically-based models and *in silico* tools. In parallel to such validation, more case studies should be developed to test these methods combining new knowledge and historical data for proof of concept. The need for publicly accessible databases integrating data from these new methods/tools is also a key issue since in the future it can be foreseen that, as knowledge of MoA/AOP advances, risk assessors and toxicologist will be able to refine models and tools (e.g. dynamic AOPs, complex cellular network models, integration of knowledge of the human microbiome). Finally, methodologies for WoE and uncertainty analysis are needed to integrate data from such new methodologies in the MoA/AOP framework, including testing the biological relevance (species concordance analysis) and report uncertainties in a transparent way. These future perspectives are further discussed below to provide detailed recommendations.

RECOMMENDATIONS

- The recommendations below result from a general consultation of EFSA expert Panels and staff dealing with chemical risk assessment and other experts from international agencies (ECHA, OECD, WHO…). Harmonisation of risk assessment terminology is a major aspect of EFSA' science strategy. In the context of these methods, harmonisation of the terminology is required This recommendation also applies to animal health and environmental risk assessments.

- General recommendations include the need for 1) reviews highlighting the use of these tools and methodologies for animal health and ecological risk assessment of chemicals. These reviews could be the starting point for developing guidance document (s) and have been identified as a priority topic by the EFSA Scientific Committee. 2) A guidance on the use of Mode of Action in risk Assessment that would include new international developments in the field. 3) Development of methodologies to apply these new tools to the risk assessment of exposure to multiple chemicals as well as to multiple chemicals combined with other stressors (e.g. biological hazards, physical agents…).

- With regards to toxicokinetic processes in hazard assessment, human toxicokinetic data are needed to characterise interspecies differences and human variability. Such data will ultimately a) link exposure, internal dose and toxicity using physiologically-based models for risk assessment purposes, b) provide a scientific ground to set Assessment Groups based on toxicokinetics for multiple chemicals. In addition, *in vitro* methods for generating TK data to measure human absorption distribution, metabolism (gut and hepatic metabolism) and excretion patterns of chemicals, should be improved.

- The development of a guidance on the use of physiologically-based models in chemical risk assessment is recommended together with the development of prototype physiologically-based models using specific case studies to integrate exposure, toxicokinetic information and toxicity data for hazard assessment purposes. It is recommended to first develop relatively simple models for single compounds and binary mixtures based on *in vitro* and *in vivo* data and, as knowledge advances more complex models can be developed. To further develop such models there is a practical need for databases providing critical parameters to build the models (physico chemical, physiological, toxicological) and bioinformatic tools/algorithms to analyse and integrate such data

- Further work is needed to explore application of *in silico* tools in chemical risk assessment making use of the current data and databases available on toxicity, and to develop and validate predictive models and reduce animal use. Further development of *in silico* tools include the systematic and

harmonised approach for the use of QSAR, further development of the read-across methodologies particularly using QSAR, physico-chemical properties and toxicological data and potential refinements of the TTC approach (re-evaluation/update of the Cramer classes and toxicological databases to improve *in silico* models with, as far as possible, an understanding of MoA.

- Regarding OMICs technologies, validation and standardisation for their use in human hazard assessment are needed. This includes detailed guidance on criteria underpinning their acceptability (data needs, statistical and bioinformatic methodologies for data analysis, mechanistic and statistical aspects, biological and toxicological relevance, *in vitro* relevance to the *in vivo* situation). Also, further exploration of the use of OMICs in human hazard assessment using case studies relevant to the food and feed area is recommended and include benchmark dose modelling, *in vitro* data, use of biomarkers of exposure and effects. Application in other areas such as animal health, ecological risk assessment and nutrition are also foreseen.

- Future work on integrated testing strategies is recommended using case studies of specific chemicals to investigate mode of action from a toxicokinetic and toxicodynamic perspective and identify data gaps and research needs. Testing should be focused on differentiating chemicals with or without specific MoA, which will give an opportunity to rank potencies for prioritisation of chemicals.

- The use of alternative test species as a bridge between *in vitro* methods and mammalian tests should be further explored. In practice, further exploration of new methodologies for hazard assessment are needed for both regulators and industry to screen and prioritise large sets of chemicals, and to assess specific chemicals for a specific purpose. For exposure to multiple chemicals, a better understanding of MoA/AOP of multiple substances using predictive and alternative methodologies is needed and will allow improvement of the basis for setting Assessment Groups.

- For the future of chemical risk assessment, exploration of new risk assessment frameworks to bring a systems toxicology perspective to risk assessment using case studies is needed. Weight of evidence and uncertainty analysis methodologies are needed for the integration of data from new methodologies in the MoA/AOP framework and chemical risk assessment as a whole. Finally, reinforcing collaborations with international institutions such as: ECHA, JRC, OECD, WHO, NTP and the US-EPA is critical for EFSA in order to facilitate integration of these new methods and international harmonisation.

REFERENCES

Abass KM, 2013. From *in vitro* hepatic metabolic studies towards human health risk assessment: Two case studies of diuron and carbosulfan. Pesticide Biochemistry and Physiology, 107, 258-265.

Aliferis KA and Jabaji S, 2011. Metabolomics - A robust bioanalytical approach for the discovery of the modes-of-action of pesticides: A review. Pesticide Biochemistry and Physiology, 100, 105-117.

Alqahtani S, Mohamed LA and Kaddoumi A, 2013. Experimental models for predicting drug absorption and metabolism. Expert Opinion on Drug Metabolism and Toxicology, 9, 1241-1254.

Allen BC, Hack CE and Clewell HJ, 2007. Use of Markov chain Monte Carlo analysis with a physiologically-based pharmacokinetic model of methylmercury to estimate exposures in US women of childbearing age. Risk Analysis, 27, 947-959.

Al-Subeihi AA, Spenkelink B, Rachmawati N, Boersma MG, Punt A and Vervoort J, 2011. Physiologically based biokinetic model of bioactivation and detoxification of the alkenylbenzene methyleugenol in rat. Toxicology in Vitro, 25, 267-285.

Alyea RA, Gollapudi BB and Rasoulpour RJ, 2014. Are we ready to consider transgenerational epigenetic effects in human health risk assessment? Environmental and Molecular Mutagenesis, 55, 292-298.

Amacher DE, 2010. The discovery and development of proteomic safety biomarkers for the detection of drug-induced liver toxicity. Toxicology and Applied Pharmacology, 245, 134-142.

Amzal B, Julin B, Vahter M, Wolk A, Johanson G and Akesson A, 2009. Population toxicokinetic modeling of cadmium for health risk assessment. Environmental Health Perspectives, 117, 1293-1301.

Andersen V, Halfvarson J and Vogel U, 2012. Colorectal cancer in patients with inflammatory bowel disease: can we predict risk? World Journal of Gastroenterology, 18, 4091-4094.

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrrano JA, Tietge JE and Villeneuve DL, 2010. Adverse Outcome Pathways: A Conceptual Framework to Support Ecotoxicology Research and Risk Assessment. Environmental Toxicology and Chemistry, 29, 730-741.

Antoine DJ, Dear JW, Lewis PS, Platt V, Coyle J, Masson M, Thanacoody RH, Gray AJ, Webb DJ, Moggs JG, Bateman DN, Goldring CE and Park BK, 2013. Mechanistic biomarkers provide early and sensitive detection of acetaminophen-induced acute liver injury at first presentation to hospital. Hepatology, 58, 777-787.

Anway MD, Cupp AS, Uzumcu M and Skinner MK, 2005. Epigenetic transgenerational actions of endocrine disruptors and male fertility. Science, 308, 1466-1469.

Attene-Ramos MS, Miller N, Huang RL, Michael S, Itkin M, Kavlock RJ, Austin CP, Shinn P, Simeonov A, Tice RR and Xia MH, 2013. The Tox21 robotic platform for the assessment of environmental chemicals - from vision to reality. Drug Discovery Today, 18, 716-723.

Aylward LL, Krishnan K, Kirman CR, Nong A and Hays SM, 2011. Biomonitoring equivalents for deltamethrin. Regulatory Toxicology and Pharmacology, 60, 189-199.

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

Basketter DA, 2012. Non-animal approaches in skin toxicology. Archives in Toxicology, 86, 1159-1160.

Benfenati E, Benigni R, Demarini DM, Helma C, Kirkland D, Martin TM, Mazzatorta G, Ouedraogo-Arras G, Richard AM, Schilter B, Schoonen WGEJ, Snyder RD and Yang C, 2009. Predictive Models for Carcinogenicity and Mutagenicity: Frameworks, State-of-the-Art, and Perspectives. Journal of Environmental Science and Health Part C, 27, 2, 57-90.

Berger B, Peng J and Singh M, 2013. Computational solutions for omics data. Nature Review Genetics, 14, 333-346.

Bernillon Pand Bois FY, 2000. Statistical issues in toxicokinetic modeling: A Bayesian perspective. Environmental Health Perspectives, 108, 883-893.

Blaauboer BJ, 2010. Biokinetic Modeling and *in Vitro-in Vivo* Extrapolations. Journal of Toxicology and Environmental Health-Part B-Critical Reviews, 13, 242-252.

Blackburn K, Bjerke D, Daston G, Felter S, Mahony C, Naciff J, Robison S and Wu SD, 2011. Case studies to test: A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. Regulatory Toxicology and Pharmacology, 60, 120-135.

Boccard J, Veuthey JL and Rudaz S, 2010. Knowledge discovery in metabolomics: An overview of MS data handling. Journal of Separation Science, 33, 290-304.

Bois FY, Jamei M and Clewell HJ, 2010. PBPK modelling of inter-individual variability in the pharmacokinetics of environmental chemicals. Toxicology, 278, 256-267.

Boitier E, Amberg A, Barbie V, Blichenberg A, Brandenburg A, Gmuender H, Gruhler A, McCarthy D, Meyer K, Riefke B, Raschke M, Schoonen W, Sieber M, Suter L, Thomas CE and Sajot N, 2011. A comparative integrated transcript analysis and functional characterization of differential mechanisms for induction of liver hypertrophy in the rat. Toxicology and Applied Pharmacology, 252, 85-96.

Boobis AR, 2005. A tiered approach to assessing the toxicology of pesticides. Toxicology Letters, 158, S30-S31.

Boobis AR, Cohen SM, Dellarco V, McGregor D, Vickers C, Willcocks D and Farland W, 2006. IPCS framework for analysing the relevance of a cancer mode of action for humans. Toxicology Letters, 164, S254-S255.

Boobis AR, Doe JE, Heinrich-Hirsch B, Meek ME, Munn S, Ruchirawat M, Schlatter J, Seed J and Vickers C, 2008. IPCS framework for analyzing the relevance of a noncancer mode of action for humans. Critical Reviews in Toxicology, 38, 87-96.

Bouhifd M, Hartung T, Hogberg HT, Kleensang A and Zhao L, 2013. Review: Toxicometabolomics. Journal of Applied Toxicology, 33, 1365-1383.

Bueters T, Juric S, Sohlenius-Sternbeck AK, Hu Y and Bylund J, 2013. Rat poorly predicts the combined non-absorbed and presystemically metabolized fractions in the human. Xenobiotica, 43, 607-616.

Bow DAJ, Perry JL, Simon JD and Pritchard JB, 2006. The impact of plasma protein binding on the renal transport of organic anions. Journal of Pharmacology and Experimental Therapeutics, 316, 349-355.

Burgess-Herbert SL and Euling SY, 2013.Use of comparative genomics approaches to characterize interspecies differences in response to environmental chemicals: challenges, opportunities, and research needs. Toxicology and Applied Pharmacology, 15;271, 372-385.

Caldwell JC, Evans MV and Krishnan K, 2012. Cutting Edge PBPK Models and Analyses: Providing the Basis for Future Modeling Efforts and Bridges to Emerging Toxicology Paradigms. Journal of Toxicology, 2012, 852384.

Cecconi D and Zamo A, 2011. Proteomics of human cancer tissues and cells. Trac-Trends in Analytical Chemistry, 30, 346-359.

Chadeau-Hyam M, Campanella G, Jombart T, Bottolo L, Portengen L, Vineis P, Liquet B and Vermeulen RC, 2013. Deciphering the complex: methodological overview of statistical models to derive OMICS-based biomarkers. Environmental and Molecular Mutagenesis, 54, 542-557.

Chaudhry AS, Urban TJ, Lamba JK, Birnbaum AK, Remmel RP and Subramanian M, 2010. CYP2C9*1B promoter polymorphisms, in linkage with CYP2C19*2, affect phenytoin autoinduction of clearance and maintenance dose. Journal of Pharmacology and Experimental Therapeutics, 332, 599-611.

Cheng S and Bois FY, 2011. A Mechanistic Modeling Framework for Predicting Metabolic Interactions in Complex Mixtures. Environmental Health Perspectives, 119, 1712-1718.

Cheng F, Li W, Zhou Y, Shen J, Wu Z, Liu G, Lee PW and Tang Y, 2013. admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. Journal of Chemical Information and Modeling, 52, 3099-3105.

Chen Y, Liu L, Monshouwer M and Fretland AJ, 2011. Determination of timedependent inactivation of CYP3A4 in cryopreserved human hepatocytes and assessment of human drug-drug interactions. Drug Metabolism and Disposition, 39, 2085–2092.

Chiu WA and Bois FY, 2006. Revisiting the population toxicokinetics of tetrachloroethylene. Archives of Toxicology, 80, 382-385.

Choi K, Pfund WP, Andersen ME, Thomas RS, Clewell HJ and Lecluyse EL, 2013. Development of 3D Dynamic Flow Model of Human Liver and its Application to Prediction of Metabolic Clearance of 7-Ethoxycoumarin. Tissue Engineering. Part C Methods, doi: 10.1089/ten.TEC.2013.0562

Clewell HJ, Crump KS, Gentry PR and Shipp AM, 2000. Site-specific reference dose for methylmercury for fish-eating populations. Fuel Processing Technology, 65, 43-54.

Clewell HJ, Gentry PR, Covington TR, Sarangapani R and Teeguarden JG, 2004. Evaluation of the potential impact of age- and gender-specific pharmacokinetic differences on tissue dosimetry. Toxicological Sciences, 79, 381-393.

Clewell RA and Clewell HJ, 2008. Development and specification of physiologically based pharmacokinetic models for use in risk assessment. Regulatory Toxicology and Pharmacology, 50, 129-143.

Coecke S, Ahr H, Blaauboer BJ, Bremer S, Casati S, Castell J, Combes R, Corvi R, Crespi CL, Cunningham ML, Elaut G, Eletti B, Freidig A, Gennari A, Ghersi-Egea JF, Guillouzo A, Hartung T, Hoet P, Ingelman-Sundberg M, Munn S, Janssens W, Ladstetter B, Leahy D, Long A, Meneguz A, Monshouwer M, Morath S, Nagelkerke F, Pelkonen O, Ponti J, Prieto P, Richert L, Sabbioni E, Schaack B, Steiling W, Testai E, Vericat JA and Worth A, 2006. Metabolism: a bottleneck in *in vitro* toxicological test development. The report and recommendations of ECVAM workshop 54. Alternatives to Laboratory Animals, 34, 49-84.

Coecke S, Pelkonen O, Leite SB, Bernauer U, Bessems JGM, Bois FY, Gundert-Remy U, Loizou G, Testai E and Zaldivar JM, 2013. Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches. Toxicology in Vitro, 27, 1570-1577.

Collins BC, Miller CA, Sposny A, Hewitt P, Wells M, Gallagher WM and Pennington SR, 2012. Development of a Pharmaceutical Hepatotoxicity Biomarker Panel Using a Discovery to Targeted Proteomics Approach. Molecular & Cellular Proteomics, 11, 394-410.

Com E, Boitier E, Marchandeau JP, Brandenburg A, Schroeder S, Hoffmann D, Mally A and Gautier JC, 2012. Integrated transcriptomic and proteomic evaluation of gentamicin nephrotoxicity in rats. Toxicology and Applied Pharmacology, 258, 124-133.

Conolly RB, Gaylor DW and Lutz WK, 2005. Population variability in biological adaptive responses to DNA damage and the shapes of carcinogen dose-response curves. Toxicology and Applied Pharmacology, 207, 570-575.

Covington TR, Gentry PR, Van Landingham CB, Andersen ME, Kester JE and Clewell HJ, 2007. The use of Markov chain Monte Carlo uncertainty analysis to support a Public Health Goal for perchloroethylene. Regulatory Toxicology and Pharmacology, 47, 1-18.

Crowell SR, Henderson WM, Kenneke JF and Fisher JW, 2011. Development and application of a physiologically based pharmacokinetic model for triadimefon and its metabolite triadimenol in rats and humans. Toxicology Letters, 205, 154-162.

Crump KS, Chen C, Chiu WA, Louis TA, Portier CJ, Subramaniam RP and White PD, 2010. What Role for Biologically Based Dose-Response Models in Estimating Low-Dose Risk? Environmental Health Perspectives, 118, 585-588.

Darwich AS, Neuhoff S, Jamei M and Rostami-Hodjegan A, 2010. Interplay of metabolism and transport in determining oral drug absorption and“gut wall metabolism: a simulation assessment using ”he "Advanced Dissolution, Absorption, Metabolism (ADAM)" model. Current Drug Metabolism, 11, 716-729.

David RM, Clewell HJ, Gentry PR, Covington TR, Morgott DA and Marino DJ, 2006. Revised assessment of cancer risk to dichloromethane - II. Application of probabilistic methods to cancer risk determinations. Regulatory Toxicology and Pharmacology, 45, 55-65.

Davies MN, Meaburn EL and Schalkwyk LC, 2010. Gene set enrichment; a problem of pathways. Brief Functional Genomics, 9, 385-390.

Dennison JE, Andersen ME and Yang RSH, 2003. Characterization of the pharmacokinetics of gasoline using PBPK modeling with a complex mixtures chemical lumping approach. Inhalation Toxicology, 15, 961-986.

DeWoskin RS, Barone S, Clewell HJ and Setzer RW, 2001. Improving the development and use of biologically based dose response models (BBDR) in risk assessment. Human and Ecological Risk Assessment, 7, 1091-1120.

Doerge DR, Twaddle NC, Woodling KA and Fisher JW, 2010. Pharmacokinetics of bisphenol A in neonatal and adult rhesus monkeys. Toxicology and Applied Pharmacology, 248, 1-11.

Doerge DR, Young JF, Chen JJ, DiNovi MJ and Henry SH, 2008. Using dietary exposure and physiologically based pharmacokinetic/pharmacodynamic modeling in human risk extrapolations for acrylamide toxicity. Journal of Agricultural and Food Chemistry, 56, 6031-6038.

Döring B and Petzinger E, 2014. Phase 0 and phase III transport in various organs: Combined concept of phases in xenobiotic transport and metabolism. Drug Metabolism Reviews, PMID:24483608.

Dorne JL, 2010. Metabolism, variability and risk assessment. Toxicology, 268, 156-164.

Dorne JL, Amzal B, Bois F, Crepet A, Tressou J and Verger P, 2012. Population effects and variability. Methods in the Molecular Biology, 929, 521-581.

Draghici S, Kulaeva O, Hoff B, Petrov A, Shams S and Tainsky MA, 2013. Noise sampling method: an ANOVA approach allowing robust selection of differentially regulated genes measured by DNA microarrays. Bioinformatics, 19, 1348-1359.

Du LF, Wang H, Xu W, Zeng Y, Hou YR, Zhang YQ, Zhao XJ and Sun CH, 2013. Application of Ultraperformance Liquid Chromatography/Mass Spectrometry-Based Metabonomic Techniques to Analyze the Joint Toxic Action of Long-term Low-Level Exposure to a Mixture of Organophosphate Pesticides on Rat Urine Profile. Toxicological Sciences, 134, 195-206.

ECHA (European Chemicals Agency), 2008. REACH Guidance on information requirements and chemical safety assessment Chapter R.6: QSARs and grouping of chemicals. Available online: http://echa.europa.eu/documents/10162/13632/information_requirements_r6_en.pdf

EFSA (European Food Safety Authority), 2005. Opinion of the Scientific Committee on a request from EFSA related to A Harmonised Approach for Risk Assessment of Substances which are both Genotoxic and Carcinogenic. The EFSA Journal 2005, 282, 1-31.

EFSA (European Food Safety Authority), 2008. Scientific Opinion of the Panel on Plant Protection Products and their Residues (PPR Panel) on a request from the EFSA evaluate the suitability of existing methodologies and, if appropriate, the identification of new approaches to assess cumulative and synergistic risks from pesticides to human health with a view to set MRLs for those pesticides in the frame of Regulation (EC) 396/2005. The EFSA Journal 2008, 705, 1-84. doi:10.2903/j.efsa.2008.705

EFSA (European Food Safety Authority), 2009a. Use of the benchmark dose approach in risk assessment. Guidance of the Scientific Committee. The EFSA Journal 2009, 1150, 1-72.

EFSA (European Food Safety Authority), 2009b. Scientific Opinion of the Panel on Contaminants in the Food Chain on a request from the European Commission on cadmium in food. The EFSA Journal 2009, 980, 1-139.

EFSA (European Food Safety Authority), 2009c. Technical report of EFSA prepared by Assessment Methodology Unit on Meta-analysis of Dose-Effect Relationship of Cadmium for Benchmark Dose Evaluation. The EFSA Journal 2009, 254r, 1-62.

EFSA (European Food Safety Authority), 2009d. Scientific opinion on the re-evaluation of Allura Red AC (E 129) as a food additive on request from the European Commission. The EFSA Journal 2009, 7(11):1327, 39 pp. doi:10.2903/j.efsa.2009.1327

EFSA (European Food Safety Authority), 2013.International Framework Dealing with Human Risk Assessment of Combined Exposure to Multiple Chemicals. EFSA Journal 2013;11(7):3313, 69 pp. doi:10.2903/j.efsa.2013.3313

EFSA (European Food Safety Authority), 2014a. A systematic procedure for the identification of emerging chemical risks in the food and feed chain. EFSA supporting publication 2014:EN-547. 40 pp.

EFSA (European Food Safety Authority), 2014b. Towards an integrated environmental risk assessment of multiple stressors on bees: review of research projects in Europe, knowledge gaps and recommendations. EFSA Journal 2014;12(3):3594, 102 pp. doi:10.2903/j.efsa.2014.3594

EFSA ANS Panel (EFSA Panel on Food Additives and Nutrient Sources Added to Food), 2013.Statement on Allura Red AC and other sulphonated mono azodyesauthorised as food and feed additives. EFSA Journal 2013;11(6):3234, 25 pp. doi:10.2903/j.efsa.2013.3234

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2012. Scientific Opinion on evaluation of the toxicological relevance of pesticide metabolites for dietary risk assessment. EFSA Journal 2012;10(07):2799, 187 pp. doi:10.2903/j.efsa.2012.2799

EFSA SC (EFSA Scientific Committee), 2012. Scientific Opinion on exploring options for providing advice about possible human health risks based on the concept of Threshold of Toxicological Concern (TTC). EFSA Journal 2012;10(7):2750, 103 pp. doi:10.2903/j.efsa.2012.2750

Ellinger-Ziegelbauer H, Adler M, Amberg A, Brandenburg A, Callanan JJ, Connor S, Fountoulakis M, Gmuender H, Gruhler A, Hewitt P, Hodson M, Matheis KA, McCarthy D, Raschke M, Riefke B, Schmitt CS, Sieber M, Sposny A, Suter L, Sweatman B and Mally A, 2011. The enhanced value of combining conventional and "omics" analyses in early assessment of drug-induced hepatobiliary injury. Toxicology and Applied Pharmacology, 252, 97-111.

Enoch SJ, Roberts DW and Cronin MT, 2010. Mechanistic category formation for the prediction of respiratory sensitization. Chemical Research in Toxicology, 23, 1547-1555.

Eustache F, Mondon F, Canivenc-Lavier MC, Lesaffre C, Fulla Y, Berges R, Cravedi JP, Vaiman D and Auger J, 2009. Chronic dietary exposure to a low-dose mixture of genistein and vinclozolin modifies the reproductive axis, testis transcriptome, and fertility. Environmental Health Perspectives, 117, 1272-1279.

Faller LD, 2008. Mechanistic studies of sodium pump. Archives of Biochemistry and Biophysics, 476, 12-21.

Fan JH, Chen S, Chow ECY and Pang KS, 2010. PBPK Modeling of Intestinal and Liver Enzymes and Transporters in Drug Absorption and Sequential Metabolism. Current Drug Metabolism, 11, 743-761.

Feng Z, Sun X, Yang J, Hao D, Du L, Wang H, Xu W, Zhao X and Sun C, 2012. Metabonomics analysis of urine and plasma from rats given long-term and low-dose dimethoate by ultra-performance liquid chromatography-mass spectrometry. Chemico-Biological Interactions, 199, 143-153.

Fiehn O, Robertson D, Griffin J, van der Werf M, Nikolau B, Morrison N, Sumner LW, Goodacre R, Hardy NW, Taylor C, Fostel J, Kristal B, Kaddurah-Daouk R, Mendes P, van Ommen B, Lindon JC and Sansone SA, 2007. The metabolomics standards initiative (MSI). Metabolomics, 3, 175-178.

Fisher JW, Twaddle NC, Vanlandingham M and Doerge DR, 2011. Pharmacokinetic modeling: Prediction and evaluation of route dependent dosimetry of bisphenol A in monkeys with extrapolation to humans. Toxicology and Applied Pharmacology, 257, 122-136.

Forgacs AL, Dere E, Angrish MM and Zacharewski TR, 2013. Comparative analysis of temporal and dose-dependent TCDD-elicited gene expression in human, mouse, and rat primary hepatocytes. Toxicological Sciences, 133, 54-66.

Fountoulakis M, Berndt P, Boelsterli UA, Crameri F, Winter M, Albertini S and Suter L, 2000. Two-dimensional database of mouse liver proteins: Changes in hepatic protein levels following treatment with acetaminophen or its nontoxic regioisomer 3-acetamidophenol. Electrophoresis, 21, 2148-2161.

Gangwal S, Reif DM, Mosher S, Egeghy PP, Wambaugh JF, Judson RS and Hubal EAC, 2012. Incorporating exposure information into the toxicological prioritization index decision support framework. Science of the Total Environment, 435–436, 316-325.

Gentry PR, Hack CE, Haber L, Maier A and Clewell HJ, 2002. An approach for the quantitative consideration of genetic polymorphism data in chemical risk assessment: Examples with warfarin and parathion. Toxicological Sciences, 70, 120-139.

Gerlowski LE and Jain RK, 1983. Physiologically Based Pharmacokinetic Modeling - Principles and Applications. Journal of Pharmacological Sciences, 72, 1103-1127.

Gissi A, Gadaleta D, Floris M, Olla S, Carotti A, Novellino E, Benfenati E and Nicolotti O, 2014. An alternative QSAR-based approach for predicting the bioconcentration factor for regulatory purposes. ALTEX, 31, 23-36.

Goldstone JV, McArthur AG, Kubota A, Zanette J, Parente T, Jonsson ME, Nelson DR and Stegeman JJ, 2010. Identification and developmental expression of the full complement of Cytochrome P450 genes in Zebrafish. BMC Genomics, 11, 643-663.

Goodman JE, Boyce CP, Pizzurro DM and Rhomberg LR, 2014. Strengthening the foundation of next generation risk assessment. Regulatory Toxicology and Pharmacology, 68, 160-170.

Govarts E, Nieuwenhuijsen M, Schoeters G, Ballester F, Bloemen K, de Boer M, Chevrier C, Eggesbo M, Guxens M, Kramer U, Legler J, Martinez D, Palkovicova L, Patelarou E, Ranft U, Rautio A, Petersen MS, Slama R, Stigum H, Toft G, Trnovec T, Vandentorren S, Weihe P, Kuperus NW, Wilhelm M, Wittsiepe J, Bonde JP and Eneieco O, 2012. Birth Weight and Prenatal Exposure to Polychlorinated Biphenyls (PCBs) and Dichlorodiphenyldichloroethylene (DDE): A Meta-analysis within 12 European Birth Cohorts. Environmental Health Perspectives, 120, 162-170.

Groothuis GMM and de Graaf IAM, 2013. Precision-cut Intestinal Slices as *In Vitro* Tool for Studies on Drug Metabolism. Current Drug Metabolism, 14, 112-119.

Haddad S, Beliveau M, Tardif R and Krishnan K, 2001. A PBPK modeling-based approach to account for interactions in the health risk assessment of chemical mixtures. Toxicological Sciences, 63, 125-131.

Haddad S, -Tardif G and Krishnan K, 1999. Physiological modeling of the toxicokinetic interactions in a quaternary mixture of aromatic hydrocarbons. Toxicology and Applied Pharmacology, 161, 249-257.

Haug K, Salek RM, Conesa P, Hastings J, de Matos P, Rijnbeek M, Mahendraker T, Williams M, Neumann S, Rocca-Serra P, Maguire E, Gonzalez-Beltran A, Sansone SA, Griffin JL and Steinbeck C, 2013. MetaboLights--an open-access general-purpose repository for metabolomics studies and associated meta-data. Nucleic Acids Research, 41, D781-786.

Herrero M, Simo C, Garcia-Canas V, Ibanez E and Cifuentes A, 2012. Foodomics: MS-based strategies in modern food science and nutrition. Mass Spectrometry Reviews, 31, 49-69.

Hillgren KM, Keppler D, Zur AA, Giacomini KM, Stieger B, Cass CE, Zhang L and Consortium IT, 2013. Emerging Transporters of Clinical Importance: An Update from the International Transporter Consortium. Clinical Pharmacology & Therapeutics, 94, 52-63.

Hinderliter PM, Price PS, Bartels MJ, Timchalk C and Poet TS, 2011. Development of a source-to-outcome model for dietary exposures to insecticide residues: An example using chlorpyrifos. Regulatory Toxicology and Pharmacology, 61, 82-92.

Hood L, 2012. Tackling the Microbiome. Science, 336(6086), 1209.

IEH (Institute of Environment and Health), 2013. Predictive approaches to chemical hazard identification and characterisation: current use by UK government departments and agencies. Institute of Environment and Health, Cranfield University, UK. Available online: http://ieh.cranfield.ac.uk/ighrc/pdf/cr%20reports/cr16[1].pdf

Illig T, Gieger C, Zhai G, Romisch-Margl W, Wang-Sattler R, Prehn C, Altmaier E, Kastenmuller G, Kato BS, Mewes HW, Meitinger T, de Angelis MH, Kronenberg F, Soranzo N, Wichmann HE, Spector TD, Adamski J and Suhre K, 2010. A genome-wide perspective of genetic variation in human metabolism. Nature Genetics, 42, 137-141.

Ishii M, Toda T, Ikarashi N, Ochiai W and Sugiyama K, 2012. Effects of Intestinal Flora on the Expression of Cytochrome P450 3A in the Liver. Yakugaku Zasshi, 132, 301-310.

Jamei M, Marciniak S, Feng K, Barnett A, Tucker G and Rostami-Hodjegan A, 2009. The Simcyp population-based ADME simulator. Expert Opinion on Drug Metabolism & Toxicology, 5, 211-223.

Jayaraman R, Pilla Reddy V, Pasha MK, Wang H, Sangthongpitag K, Yeo P, Hu CY, Wu X, Xin L, Goh E, New LS and Ethirajulu K, 2011. Preclinical metabolism and disposition of SB939 (Pracinostat), an orally active histone deacetylase inhibitor, and prediction of human pharmacokinetics. Drug Metabolism and Disposition, 39, 2219–2232.

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

Jourdan F, Cottret L, Huc L, Wildridge D, Scheltema R, Hillenweck A, Barrett MP, Zalko D, Watson DG and Debrauwer L, 2010. Use of reconstituted metabolic networks to assist in metabolomic data visualization and mining. Metabolomics, 6, 312-321.

JRC (Joint Research Centre), 2010. Scientific Report submitted to EFSA. Applicability of QSAR analysis to the evaluation of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment. Prepared by Computational Toxicology Group, Institute for Health & Consumer Protection, European Commission- Joint Research Centre, Ispra, Italy (Question No EFSA-Q-2009-01076. Accepted for Publication on 5 May 2010).

JRC (Joint Research Centre), 2011. Applicability of QSAR analysis in the evaluation of developmental and neurotoxicity effects for the assessment of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment. Report produced for EFSA. Available online: http://www.efsa.europa.eu/en/supporting/pub/169e.htm

Kamp H, Fabian E, Groeters S, Herold M, Krennrich G, Looser R, Mattes W, Mellert W, Prokoudine A, Ruiz-Noppinger P, Strauss V, Walk T, Wiemer J and van Ravenzwaay B, 2012. Application of in vivo metabolomics to preclinical/toxicological studies: case study on phenytoin-induced systemic toxicity. Bioanalysis, 4, 2291-2301.

Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, Heger A, Agam A, Slater G, Goodson M, Furlotte NA, Eskin E, Nellaker C, Whitley H, Cleak J, Janowitz D, Hernandez-Pliego P, Edwards A, Belgard TG, Oliver PL, McIntyre RE, Bhomra A, Nicod J, Gan X, Yuan W, van der Weyden L, Steward CA, Bala S, Stalker J, Mott R, Durbin R, Jackson IJ, Czechanski A, Guerra-Assuncao JA, Donahue LR, Reinholdt LG, Payseur BA, Ponting CP, Birney E, Flint J and Adams DJ, 2011. Mouse genomic variation and its effect on phenotypes and gene regulation. Nature, 477, 289-294.

Kim TH, Ahn MY, Lim HJ, Lee YJ, Shin YJ, De U, Lee J, Lee BM, Kim S and Kim HS, 2012. Evaluation of metabolomic profiling against renal toxicity in Sprague-Dawley rats treated with melamine and cyanuric acid. Archives of Toxicology, 86, 1885-1897.

Kleinstreuer NC, Smith AM, West PR, Conard KR, Fontaine BR, Weir-Hauptman AM, Palmer JA, Knudsen TB, Dix DJ, Donley ELR and Cezar GG, 2011. Identifying developmental toxicity pathways for a subset of ToxCast chemicals using human embryonic stem cells and metabolomics. Toxicology and Applied Pharmacology, 257, 111-121.

Knaak JB, Dary CC, Power F, Thompson CB and Blancato JN, 2004. Physicochemical and biological data for the development of predictive organophosphorus pesticide QSARs and PBPK/PD models for human risk assessment. Critical Reviews in Toxicology, 34, 143-207.

Knaak JB, Dary CC, Okino MS, Power FW, Zhang X, Thompson CB, Tornero-Velez R and Blancato JN, 2008. Parameters for Carbamate Pesticide QSAR and PBPK/PD Models for Human Risk Assessment. Review of Environmental Contaminants in Toxicology, 193, 53-212.

Knudsen AD, Bennike T, Kjeldal H, Birkelund S, Otzen DE and Stensballe A, 2014. Condenser: A statistical aggregation tool for multi-sample quantitative proteomic data from Matrix Science Mascot Distiller. Journal of Proteomics. doi: 10.1016/j.jprot.2014.02.001

Kopec AK, Boverhof DR, Nault R, Harkema JR, Tashiro C, Potter D, Sharratt B, Chittim B and Zacharewski TR, 2013. Toxicogenomic evaluation of long-term hepatic effects of TCDD in immature, ovariectomized C57BL/6 mice.Toxicological Sciences, 135, 465-475.

Kobayashi CA, Leite Ade L, da Silva TL, dos Santos LD, Nogueira FC, Santos KS, de Oliveira RC, Palma MS, Domont GB and Buzalaf MA, 2011. Proteomic analysis of urine in rats chronically exposed to fluoride. Journal of Biochemical and Molecular Toxicology, 25, 8-14.

Koen YM, Gogichaeva NV, Alterman MA and Hanzlik RP, 2007. A proteomic analysis of bromobenzene reactive metabolite targets in rat liver cytosol *in vivo*. Chemical Research in Toxicology, 20, 511-519.

Koulman A, Lane GA, Harrison SJ and Volmer DA, 2009. From differentiating metabolites to biomarkers. Analytical and Bioanalytical Chemistry, 394, 663-670.

Krewski D, Westphal M, Al-Zoughool M, Croteau MC and Andersen ME, 2011. New Directions in Toxicity Testing. Annual Review of Public Health, 32, 161-178.

Kroes R, Renwick AG, Cheeseman M, Kleiner J, Mangelsdorf I, Piersma A, Schilter B, Schlatter J, van Schothorst F, Vos JG and Wurtzen G, 2004. Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. Food and Chemical Toxicology, 42, 65-83.

Kroes R, Kleiner J and Renwick A, 2005. The threshold of toxicological concern concept in risk assessment. Toxicological Sciences, 86, 226-230.

Kumar V, Schuck EL, Pelletier RD, Farah N, Condon KB, Ye M, Rowbottom C, King BM, Zhang ZY, Saxton PL, Wong YN, 2011. Pharmacokinetic characterization of a natural product-inspired novel MEK1 inhibitor E6201 in preclinical species. Cancer Chemotherapy and Pharmacology, 69, 229–237.

Kusuhara H and Sugiyama Y, 2009. In vitro-in vivo extrapolation of transporter-mediated clearance in the liver and kidney. Drug Metabolism and Pharmacokinetics, 24, 37-52.

Lafond M, Bouza B, Eyrichine S, Bonnin E, Crost EH, Geraert PA, Giardina T and Ajandouz el H, 2011. An integrative in vitro approach to analyse digestion of wheat polysaccharides and the effect of enzyme supplementation. British Journal of Nutrition, 106, 264-273.

Lan K, Xie GX and Jia W, 2013. Towards Polypharmacokinetics: Pharmacokinetics of Multicomponent Drugs and Herbal Medicines Using a Metabolomics Approach. Evidence-Based Complementary and Alternative Medicine. doi: 10.1155/2013/819147

Lapenna S and Worth A, 2011. Analysis of the Cramer classification scheme for oral systemic toxicity - implications for its implementation in Toxtree. JRC Scientific and Technical Report EUR 24898 EN. Publications Office of the European Union, Luxembourg. Available from: http://publications.jrc.ec.europa.eu/repository/

Law FC and Wang Y, 1997. [A review of physiologically based pharmacokinetic models: development and applications]. Yao Xue Xue Bao, 32, 151-160.

Lee MY, Dordick JS and Clark DS, 2010a. Metabolic enzyme microarray coupled with miniaturized cell-culture array technology for high-throughput toxicity screening. Methods in Molecular Biology, 632, 221-237.

Lee S, Poet TS, Smith JN, Busby-Hjerpe AL and Timchalk C, 2010b. Effect of *in vivo* nicotine exposure on chlorpyrifos pharmacokinetics and pharmacodynamics in rats. Chemico-Biological Interactions, 184, 449-457.

Lee S, Poet TS, Smith JN, Hjerpe AL, Gunawan R and Timchalk C, 2011. Impact of repeated nicotine and alcohol coexposure on *in vitro* and *in vivo* chlorpyrifos dosimetry and cholinesterase inhibition. Journal of Toxicology and Environmental Health A, 74, 1334-1350.

Leung HW, 1991. Development and Utilization of Physiologically Based Pharmacokinetic Models for Toxicological Applications. Journal of Toxicology and Environmental Health, 32, 247-267.

Li J, Wang S, Wang M, Shi W, Du X and Sun C, 2013a. The toxicity of 3-chloropropane-1,2-dipalmitate in Wistar rats and a metabonomics analysis of rat urine by ultra-performance liquid chromatography-mass spectrometry. Chemico-Biological Interactions, 206, 337-345.

Li F, Pang XY, Krausz KW, Jiang CT, Chen C, Cook JA, Krishna MC, Mitchell JB, Gonzalez FJ and Patterson AD, 2013b. Stable Isotope- and Mass Spectrometry-based Metabolomics as Tools in Drug Metabolism: A Study Expanding Tempol Pharmacology. Journal of Proteome Research, 12, 1369-1376.

Lieschke GJ and Currie PD, 2007. Animal models of human disease: zebrafish swim into view. Nature Reviews. Genetics, 8, 353-367.

Liu Y, Li Y, Liu K and Shen J, 2014. Exposing to cadmium stress cause profound toxic effect on microbiota of the mice intestinal tract. PloS one, 9(2),e85323.

Lo Piparo E, Worth A, Manibusan M, Yang C, Schilter B, Mazzatorta P, Jacobs MN, Steinkellner H and Mohimont L, 2011. Use of computational tools in the field of food safety. Regulatory Toxicology and Pharmacology, 60, 354-362.

Loccisano AE, Campbell JL, Andersen ME and Clewell HJ, 2011. Evaluation and prediction of pharmacokinetics of PFOA and PFOS in the monkey and human using a PBPK model. Regulatory Toxicology and Pharmacology, 59, 157-175.

Loizou G and Hogg A, 2011. MEGen: A Physiologically Based Pharmacokinetic Model Generator. Frontiers in Pharmacology, 2, 56.

Lyons MA, Yang RSH, Mayeno AN and Reisfeld B, 2008. Computational toxicology of chloroform: Reverse dosimetry using Bayesian inference, Markov chain Monte Carlo simulation, and human biomonitoring data. Environmental Health Perspectives, 116, 1040-1046.

McGettigan PA, 2013.Transcriptomics in the RNA-seq era. Current Opinion in Chemical Biology, 17, 4-11.

Martati E, Boersma MG, Spenkelink A, Khadka DB, Punt A and Vervoort J, 2011. Physiologically based biokinetic (PBBK) model for safrole bioactivation and detoxification in rats. Chemical Research in Toxicology, 24, 818-834.

Mattes WB, Kamp HG, Fabian E, Herold M, Krennrich G, Looser R, Mellert W, Prokoudine A, Strauss V, van Ravenzwaay B, Walk T, Naraoka H, Omura K, Schuppe-Koistinen I, Nadanaciva S, Bush ED, Moeller N, Ruiz-Noppinger P and Piccoli SP, 2013. Prediction of Clinically Relevant

Safety Signals of Nephrotoxicity through Plasma Metabolite Profiling. Biomed Research International. doi: 10.1155/2013/202497

Mays C, Benfenati E and Pardoe S, 2012. Use and perceived benefits and barriers of QSAR models for REACH: findings from a questionnaire to stakeholders. Chemistry Central Journal, 6. doi: 10.1186/1752-153x-6-159

Meek ME, Boobis AR, Crofton KM, Heinemeyer G, Van Raaij M and Vickers C, 2011. Risk assessment of combined exposure to multiple chemicals: A WHO/IPCS framework. Regulatory Toxicology and Pharmacology, 60, S1-S14.

Meek ME, Boobis A, Cote I, Dellarco V, Fotakis G, Munn S, Seed J and Vickers C, 2013. New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis. Journal of Applied Toxicology, 34, 1-18.

Meganathan K, Jagtap S, Wagh V, Winkler J, Gaspar JA, Hildebrand D, Trusch M, Lehmann K, Hescheler J, Schluter H and Sachinidis A, 2012. Identification of thalidomide-specific transcriptomics and proteomics signatures during differentiation of human embryonic stem cells. PLoS One, 7, e44228.

Meng Z and Veenstra TD, 2011. Targeted mass spectrometry approaches for protein biomarker verification. Journal of Proteomics, 74, 2650-2659.

Menni C, Zhai GJ, MacGregor A, Prehn C, Romisch-Margl W, Suhre K, Adamski J, Cassidy A, Illig T, Spector TD and Valdes AM, 2013. Targeted metabolomics profiles are strongly correlated with nutritional patterns in women. Metabolomics, 9, 506-514.

Merrick BA and Bruno ME, 2004. Genomic and proteomic profiling for biomarkers and signature profiles of toxicity. Current Opinion in Molecular Therapeutics, 6, 600-607.

Metzker ML, 2010. Sequencing technologies - the next generation. Nature Reviews. Genetics, 11, 31-46.

Meyer MR, Orschiedt T and Maurer HH, 2013. Michaelis-Menten kinetic analysis of drugs of abuse to estimate their affinity to human P-glycoprotein. Toxicology Letters, 217, 137-142.

Milreu PV, Klein CC, Cottret L, Acuna V, Birmele E, Borassi M, Junot C, Marchetti-Spaccamela A, Marino A, Stougie L, Jourdan F, Crescenzi P, Lacroix V and Sagot MF, 2014. Telling metabolic stories to explore metabolomics data: a case study on the yeast response to cadmium exposure. Bioinformatics, 30, 61-70.

Mirfazaelian A, Kim KB, Anand SS, Kim HJ, Tornero-Velez R, Bruckner JV and Fisher JW, 2006. Development of a physiologically based pharmacokinetic model for deltamethrin in the adult male Sprague-Dawley rat. Toxicological Sciences, 93, 432-442.

Mitra P, Audus K, Williams G, Yazdanian M and Galinis D, 2011. A comprehensive study demonstrating that p-glycoprotein function is directly affected by changes in pH: Implications for intestinal pH and effects on drug absorption. Journal of Pharmaceutical Sciences. doi: 10.1002/jps.22596

Montoya GA, Strauss V, Fabian E, Kamp H, Mellert W, Walk T, Looser R, Herold M, Krennrich G, Peter E and van Ravenzwaay B, 2014. Mechanistic analysis of metabolomics patterns in rat plasma during administration of direct thyroid hormone synthesis inhibitors or compounds increasing thyroid hormone clearance. Toxicology Letters, 225, 240-251.

Nault R, Forgacs AL, Dere E and Zacharewski TR, 2013a. Comparisons of differential gene expression elicited by TCDD, PCB126, βNF, or ICZ in mouse hepatoma Hepa1c1c7 cells and C57BL/6 mouse liver. Toxicology Letters, 223, 52-59.

Nault R, Kim S and Zacharewski TR, 2013b. Comparison of TCDD-elicited genome-wide hepatic gene expression in Sprague-Dawley rats and C57BL/6 mice. Toxicology and Applied Pharmacology, 267, 184-191.

Nicholson JK, Lindon JC and Holmes E, 1999. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica, 29, 1181-1189.

Nicholson JK and Wilson ID, 2003. Understanding 'global' systems biology: Metabonomics and the continuum of metabolism. Nature Reviews Drug Discovery, 2, 668-676.

Nossol C, Diesing AK, Walk N, Faber-Zuschratter H, Hartig R, Post A, Kluess J, Rothkotter HJ and Kahlert S, 2011. Air-liquid interface cultures enhance the oxygen supply and trigger the structural and functional differentiation of intestinal porcine epithelial cells (IPEC). Histochemistry and Cell Biology, 136, 103-115.

NRC (National Research Council), 2007. Toxicity Testing in the 21st Century: a vision and a strategy.Washington, The National Academies Press.

NRC (National Research Council),2009. Science and Decisions: Advancing Risk Assessment. Washington, The National Academies Press.

NTP (National Toxicology Program), 2004. A National Toxicology Program for the 21st Century: Roadmap to Achieve the NTP Vision Durham: National Institute of Environmental Health Sciences. http://ntp.niehs.nih.gov/NTP/About_NTP/NTPVision/NTPRoadmap_508.pdf

Nuwaysir EF, Bittner M, Trent J, Barrett JC and Afshari CA, 1999. Microarrays and toxicology: The advent of toxicogenomics. Molecular Carcinogenesis, 24, 153-159.

OECD (Organisation for Economic Co-operation and Development), 2004. Report from the Expert Group on (Q)SARs on Principles for the Validation of (Q)SARs Series on Testing and Assessment, No. 49. Organisation for Economic Co-Operation and Development, Paris, France. Available online: http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282004%2924&doclanguage=en

OECD (Organisation for Economic Co-Operation and Development), 2007. Guidance on Grouping of Chemicals. ENV/JM/MONO(2007)28, Series on Testing and Assessment Number 80, Organisation for Economic Co-Operation and Development, Paris, France. Available online: http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono(2007)28&doclanguage=en.

OECD, 2009a. Report of the second survey on available omics tools. ENV/JM/MONO(2008)35. Series on testing and assessment Number 100. Available online: http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282008%2935&doclanguage=en

OECD (Organisation for Economic Co-Operation and Development), 2009b. Guidance document for using the OECD (Q)SAR Application Toolbox to develop chemical categories according to the OECD Guidance on grouping of chemicals. ENV/JM/MONO(2009)5, Series on Testing and Assessment No. 102.

OECD (Organisation for Economic Co-Operation and Development), 2011. WHO OECD ILSI/HESI International Workshop on Risk Assessment of Combined Exposures to Multiple Chemicals. Paris, France, OECD Environment Directorate. OECD Environment, Health and Safety Publications. Series on Testing and Assessment. No 108. Available online: http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono%282011%2910&doclanguage=en

OECD (Organisation for Economic Co-Operation and Development), 2012. QSAR toolbox. User Manual. Getting started. Available online: http://www.oecd.org/chemicalsafety/risk-assessment/TB3%200_GettingStarted_rev2.pdf

OECD (Organisation for Economic Co-Operation and Development), 2013. Guidance document on developing and assessing adverse ouctome pathways. Paris, France, OECD Environment

Directorate. OECD Environment, Health and Safety Publications. Series on Testing and Assessment. No 184.

Oliver S, 2003. Functional genomics: All the king's horses and all the king's men can put humpty together again. Molecular Cell, 12, 1343-1344.Oliver SG, Winson MK, Kell DB and Baganz F, 1998. Systematic functional analysis of the yeast genome. Trends in Biotechnology, 16, 373-378.

Ozdemir V, Husereau D, Hyland S, Samper S and Salleh MZ, 2009. Personalized Medicine Beyond Genomics: New Technologies, Global Health Diplomacy and Anticipatory Governance. Current Pharmacogenomics Personalised Medicine, 7, 225-230.

Padhi BK, Pelletier G, Williams A, Berndt-Weis L, Yauk C, Bowers WJ and Chu I, 2008. Gene expression profiling in rat cerebellum following in utero and lactational exposure to mixtures of methylmercury, polychlorinated biphenyls and organochlorine pesticides. Toxicology Letters, 176, 93-103.

Paini A, Punt A, Scholz G, Gremaud E, Spenkelink B, Alink G, 2012. *In vivo* validation of DNA adduct formation by estragole in rats predicted by physiologically based biodynamic modelling. Mutagenesis, 27, 653-663.

Pan TL, Wang PW, Al-Suwayeh SA, Huang YJ and Fang JY, 2012. Toxicological effects of cationic nanobubbles on the liver and kidneys: biomarkers for predicting the risk. Food and Chemical Toxicology, 50, 3892-3901.

Pargent W, Heffner S, Schable KF, Soewarto D, Fuchs H and Hrabe de Angelis M, 2000. MouseNet database: digital management of a large-scale mutagenesis project. Mammalian Genome, 11, 590-593.

Patlewicz G, Jeliazkova N, Safford RJ, Worth AP and Aleksiev B, 2008. An evaluation of the implementation of the Cramer classification scheme in the Toxtree software. SAR and QSAR in Environmental Research, 19, 495-524.

Patti GJ, Tautenhahn R and Siuzdak G, 2012. Meta-analysis of untargeted metabolomic data from multiple profiling experiments. Nature Protocols, 7, 508-516.

Pelekis M and Emond C, 2009. Physiological modeling and derivation of the rat to human toxicokinetic uncertainty factor for the carbamate pesticide aldicarb. Environmental Toxicology and Pharmacology, 28, 179-191.

Peng S, Yan L, Zhang J, Wang Z, Tian M and Shen H, 2013. An integrated metabonomics and transcriptomics approach to understanding metabolic pathway disturbance induced by perfluorooctanoic acid. Journal of Pharmacy and Biomedical Analysis, 86, 56-64.

Perkins EJ, Ankley GT, Crofton KM, Garcia-Reyero N, LaLone CA, Johnson MS, Tietge JE and Villeneuve DL, 2013. Current perspectives on the use of alternative species in human health and ecological hazard assessments. Environmental Health Perspectives, 121, 1002-1010.

Péry AR, Brochot C, Hoet PH, Nemmar A and Bois FY, 2009. Development of a physiologically based kinetic model for 99m-technetium-labelled carbon nanoparticles inhaled by humans. Inhalation Toxicology, 21, 1099-1107.

Péry AR, Brochot C, Zeman FA, Mombelli E, Desmots S, Pavan M, Fioravanzo E and Zaldívar JM, 2013. Prediction of dose-hepatotoxic response in humans based on toxicokinetic/toxicodynamic modeling with or without *in vivo* data: a case study with acetaminophen. Toxicology Letters, 220, 26-34.

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

Pevsner and Jonathan, 2009. Bioinformatics and functional genomics (2nd ed.). Hoboken, N.J: Wiley-Blackwell. ISBN 9780470085851.

Pielaat A, Barker GC, Hendriksen P, Peijnenburg A and Kuile BH, 2013. A foresight study on emerging technologies: State of the art of omics technologies and potential applications in food and feed safety. REPORT 1: Review on the state of art of omics technologies in risk assessment related to food and feed safety. EFSA supporting publication 2013:EN-495, 126 pp. Available online: www.efsa.europa.eu/publications

Pietu G, Mariage-Samson R, Fayein NA, Matingou C, Eveno E, Houlgatte R, Decraene C, Vandenbrouck Y, Tahi F, Devignes MD, Wirkner U, Ansorge W, Cox D, Nagase T, Nomura N and Auffray C, 1999. The Genexpress IMAGE knowledge base of the human brain transcriptome: a prototype integrated resource for functional and computational genomics. Genome Research, 9, 195-209.

Poet TS, Kousba AA, Dennison SL and Timchalk C, 2004. Physiologically based pharmacokinetic/pharmacodynamic model for the organophosphorus pesticide diazinon. Neurotoxicology, 25, 1013-1130.

Potera C, 2014. Clues to Arsenic's Toxicity: Microbiome Alterations in the Mouse Gut. Environmental Health Perspectives, 122, A82.

Poulin P and Haddad S, 2013. Hepatocyte composition-based model as a mechanistic tool for predicting the cell suspension: aqueous phase partition coefficient of drugs in *in vitro* metabolic studies. Journal of Pharmaceutical Sciences, 102, 2806-2818.

Price PS, Hollnagel HM and Zabik JM, 2009. Characterizing the Noncancer Toxicity of Mixtures Using Concepts from the TTC and Quantitative Models of Uncertainty in Mixture Toxicity. Risk Analysis, 29, 1534-1548.

Punt A, Freidig AP, Delatour T, Scholz G, Boersma MG, Schilter B, van Bladeren PJ and Rietjens IM, 2008. A physiologically based biokinetic (PBBK) model for estragole bioactivation and detoxification in rat. Toxicology and Applied Pharmacology, 231, 248-259.

Punt A, Schiffelers MJWA, Horbach GJ, van de Sandt JJM, Groothuis GMM, Rietjens IMCM and Blaauboer BJ, 2011. Evaluation of research activities and research needs to increase the impact and applicability of alternative testing strategies in risk assessment practice. Regulatory Toxicology and Pharmacology, 61, 105-114.

Rabinowitz JR, Goldsmith MR, Little SB and Pasquinelli MA, 2008. Computational molecular modeling for evaluating the toxicity of environmental chemicals: prioritizing bioassay requirements. Environmental Health Perspectives, 116, 573-577.

Rakyan VK, Down TA, Balding DJ and Beck S, 2011. Epigenome-wide association studies for common human diseases. Nature Reviews. Genetics, 12, 529-541.

Ramirez T, Daneshian M, Kamp H, Bois FY, Clench MR, Coen M, Donley B, Fischer SM, Ekman DR, Fabian E, Guillou C, Heuer J, Hogberg HT, Jungnickel H, Keun HC, Krennrich G, Krupp E, Luch A, Noor F, Peter E, Riefke B, Seymour M, Skinner N, Smirnova L, Verheij E, Wagner S, Hartung T, van Ravenzwaay B and Leist M, 2013. Metabolomics in Toxicology and Preclinical Research. ALTEX-Alternatives to Animal Experimentation, 30, 209-225.

Rappaport SM and Smith MT, 2010. Epidemiology. Environment and disease risks. Science, 330, 460-461.

Reif DM, Martin MT, Tan SW, Houck KA, Judson RS, Richard AM, Knudsen TB, Dix DJ and Kavlock RJ, 2010. Endocrine profiling and prioritization of environmental chemicals using ToxCast data. Environmental Health Perspectives, 118, 1714-1720.

Reif DM, Sypa M, Lock EF, Wright FA, Wilson A, Cathey T, Judson RR and Rusyn I, 2013. ToxPi GUI: an interactive visualization tool for transparent integration of data from diverse sources of evidence. Bioinformatics, 29, 402-403.

Reinartz J, Bruyns E, Lin JZ, Burcham T, Brenner S, Bowen B, Kramer M and Woychik R, 2002. Massively parallel signature sequencing (MPSS) as a tool for in-depth quantitative gene expression profiling in all organisms. Briefings in Functional Genomics & Proteomics, 1, 95-104.

Robertson DG, Watkins PB and Reily MD, 2011. Metabolomics in Toxicology: Preclinical and Clinical Applications. Toxicological Sciences, 120, S146-S170.

Rodriguez-Suarez E and Whetton AD, 2013. The application of quantification techniques in proteomics for biomedical research. Mass Spectrometry Reviews, 32, 1-26.

Roncaglioni A, Toropov AA, Toropova AP and Benfenati E, 2013. *In silico* methods to predict drug toxicity. Current Opinion in Pharmacology, 13, 802-806.

Rostami-Hodjegan A and Tucker GT, 2007. Simulation and prediction of *in vivo* drug metabolism in human populations from *in vitro* data. Nature Reviews. Drug Discovery, 6, 140-148.

Rostkowski M, Spjuth O and Rydberg P, 2013. WhichCyp: prediction of cytochromes P450 inhibition. Bioinformatics, 29, 2051-2052.

Rotroff DM, Wetmore BA, Dix DJ, Ferguson SS, Clewell HJ, Houck KA, 2010. Incorporating human dosimetry and exposure into high-throughput *in vitro* toxicity screening. Toxicological Sciences, 117, 348-358.

Rowland M, 2013. Physiologically-Based Pharmacokinetic (PBPK) Modeling and Simulations Principles, Methods, and Applications in the Pharmaceutical Industry. CPT: Pharmacometrics & Systems Pharmacology, 2, e55.

Ruiz P, Mumtaz M, Osterloh J, Fisher J and Fowler BA, 2010. Interpreting NHANES biomonitoring data, cadmium. Toxicology Letters, 198, 44-48.

Rusyn I and Daston GP, 2010. Computational Toxicology: Realizing the Promise of the Toxicity Testing in the 21st Century. Environmental Health Perspectives, 118, 1047-1050.

Sabido E, Selevsek N and Aebersold R, 2012. Mass spectrometry-based proteomics for systems biology. Current Opinion in Biotechnology, 23, 591-597.

Santoriello C and Zon LI, 2012. Hooked! Modeling human disease in zebrafish. Journal of Clinical Investigation, 122, 2337-2343.

Squassina A, Manchia M, Manolopoulos VG, Artac M, Lappa-Manakou C, Karkabouna S, Mitropoulos K, Del Zompo M, and Patrinos GP, 2010. Realities and expectations of pharmacogenomics and personalized medicine: impact of translating genetic knowledge into clinical practice. Pharmacogenomics, 11, 1149-1167.

Sasso AF, Isukapalli SS and Georgopoulos PG, 2010. A generalized physiologically-based toxicokinetic modeling system for chemical mixtures containing metals. Theoretical Biology and Medical Modelling, 7:17.

Sasso AF, Georgopoulos PG, Isukapalli SS and Krishnan K, 2012. Bayesian Analysis of a Lipid-Based Physiologically Based Toxicokinetic Model for a Mixture of PCBs in Rats. Journal of Toxicology, 2012, 895391.

SCCS (Scientific Committee on Consumer Safety), SCENHIR (Scientific Committee on Emerging and Newly Identified Health Risks) and SCHER (Scientific Committee on Health and Environmental Risks), 2012. Toxicity and Assessment of Chemical Mixtures. 50 pp. Available online:http://ec.europa.eu/health/scientific_committees/environmental_risks/docs/scher_o_155.pdf

Schilter B, Benigni R, Boobis A, Chiodini A, Cockburn A, Cronin MT, Lo Piparo E, Modi S, Thiel A and Worth A, 2014. Establishing the level of safety concern for chemicals in food without the need for toxicity testing. Regulatory Toxicology and Pharmacology, 68, 275-296.

Schmidt CW, 2009. TOX 21: new dimensions of toxicity testing. Environmental Health Perspectives, 117(8), A348-353.

Schnackenberg LK, Sun JC, Pence LM, Bhattacharyya S, da Costa GG and Beger RD, 2012. Metabolomics evaluation of hydroxyproline as a potential marker of melamine and cyanuric acid nephrotoxicity in male and female Fischer F344 rats. Food and Chemical Toxicology, 50, 3978-3983.

Scholz S, 2013. Zebrafish embryos as an alternative model for screening of drug-induced organ toxicity. Archives of Toxicology, 87, 767-769.

Scholz S, Sela E, Blaha L, Braunbeck T, Galay-Burgos M, Garcia-Franco M, Guinea J, Kluver N, Schirmer K, Tanneberger K, Tobor-Kaplon M, Witters H, Belanger S, Benfenati E, Creton S, Cronin MT, Eggen RI, Embry M, Ekman D, Gourmelon A, Halder M, Hardy B, Hartung T, Hubesch B, Jungmann D, Lampi MA, Lee L, Leonard M, Kuster E, Lillicrap A, Luckenbach T, Murk AJ, Navas JM, Peijnenburg W, Repetto G, Salinas E, Schuurmann G, Spielmann H, Tollefsen KE, Walter-Rohde S, Whale G, Wheeler JR and Winter MJ, 2013. A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. Regulatory Toxicology and Pharmacology, 67, 506-530.

Setzer RW, Lau C, Mole ML, Copeland MF, Rogers JM and Kavlock RJ, 2001. Toward a biologically based dose-response model for developmental toxicity of 5-fluorouracil in the rat: A mathematical construct. Toxicological Sciences, 59, 49-58.

Shi Q, Yang X and Mendrick DL, 2013. Hopes and challenges in using miRNAs as translational biomarkers for drug-induced liver injury. Biomarkers in Medicine, 7, 307-15.

Shin KH, Choi MH, Lim KS, Yu KS, Jang IJ and Cho JY, 2013. Evaluation of Endogenous Metabolic Markers of Hepatic CYP3A Activity Using Metabolic Profiling and Midazolam Clearance. Clinical Pharmacology & Therapeutics, 94, 601-609.

Shuey DL, Lau C, Logsdon TR, Zucker RM, Elstein KH, Narotsky MG, Setzer RW, Kavlock RJ and Rogers JM, 1994. Biologically-Based Dose-Response Modeling in Developmental Toxicology - Biochemical and Cellular Sequelae of 5-Fluorouracil Exposure in the Developing Rat. Toxicology and Applied Pharmacology, 126, 129-144.

Shukla SJ, Huang RL, Austin CP and Xia MH, 2010. The future of toxicity testing: a focus on *in vitro* methods using a quantitative high-throughput screening platform. Drug Discovery Today, 15, 997-1007.

Sipes NS, Martin MT, Kothiya P, Reif DM, Judson RS, Richard AM, Houck KA, Dix DJ, Kavlock RJ and Knudsen TB, 2013. Profiling 976 ToxCast Chemicals across 331 Enzymatic and Receptor Signaling Assays. Chemical Research in Toxicology, 26, 878-895.

Sjogren E, Westergren J, Grant I, Hanisch G, Lindfors L, Lennernas H, Abrahamsson B and Tannergren C, 2013. *In silico* predictions of gastrointestinal drug absorption in pharmaceutical product development: application of the mechanistic absorption model GI-Sim. European Journal of Pharmacological Sciences, 49, 679-698.

Sommer F and Bäckhed F, 2013. The gut microbiota-masters of host development and physiology. Nature Reviews. Microbiology, 11, 227-238.

Song L and Florea L, 2013. CLASS: constrained transcript assembly of RNA-seq reads. BMC Bioinformatics, 14 Suppl 5, S14.

Steinbeck C, Conesa P, Haug K, Mahendraker T, Williams M, Maguire E, Rocca-Serra P, Sansone SA, Salek RM and Griffin JL, 2012. MetaboLights: towards a new COSMOS of metabolomics data management. Metabolomics, 8, 757-760.

Stierum R, Heijne W, Kienhuis A, van Ommen B and Groten J, 2005. Toxicogenomics concepts and applications to study hepatic effects of food additives and chemicals. Toxicology and Applied Pharmacology, 207, 179-188.

Sturla SJ, Boobis AR, Fitzgerald RE, Hoeng J, Kavlock RJ, Schirmer K, Whelan M, Wilks MF and Peitsch MC, 2014. Systems toxicology: from basic research to risk assessment. Chemical Research in Toxicology, 27, 314-329.

Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW, Fiehn O, Goodacre R, Griffin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriott P, Nicholls AW, Reily MD, Thaden JJ and Viant MR, 2007. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics, 3, 211-221.

Sun B, Utleg AG, Hu Z, Qin S, Keller A, Lorang C, Gray L, Brightman A, Lee D, Alexander VM, Ranish JA, Moritz RL and Hood L, 2013. Glycocapture-assisted global quantitative proteomics (gagQP) reveals multiorgan responses in serum toxicoproteome. Journal of Proteome Research, 12, 2034-2044.

Sun YJ, Wang HP, Liang YJ, Yang L, Li W and Wu YJ, 2012. An NMR-based metabonomic investigation of the subacute effects of melamine in rats. Journal of Proteome Research, 11, 2544-2550.

Supic G, Jagodic M and Magic Z, 2013. Epigenetics: a new link between nutrition and cancer. Nutrition and Cancer, 65, 781-792.

Sysi-Aho M, Katajamaa M, Yetukuri L and Oresic M, 2007. Normalization method for metabolomics data using optimal selection of multiple internal standards. BMC Bioinformatics, 8. doi: 10.1186/1471-2105-8-93

Taylor JA, vom Saal FS, Welshons WV, Drury B, Rottinghaus G, Hunt PA, Toutain PL, Laffont CM and VandeVoort CA, 2011. Similarity of Bisphenol A Pharmacokinetics in Rhesus Monkeys and Mice: Relevance for Human Exposure. Environmental Health Perspectives, 119, 422-430.

Teeguarden JG and Barton HA, 2004. Computational modeling of serum-binding proteins and clearance in extrapolations across life stages and species for endocrine active compounds. Risk Analysis, 24, 751-770.

Teschendorff AE and Widschwendter M, 2012. Differential variability improves the identification of cancer risk markers in DNA methylation studies profiling precursor cancer lesions. Bioinformatics, 28, 1487-1494.

Thelen K and Dressman JB, 2009. Cytochrome P450-mediated metabolism in the human gut wall. Journal of Pharmacy and Pharmacology, 61, 541-558.Thomas RS, Allen BC, Nong A, Yang L, Bermudez E, Clewell HJ, 3rd and Andersen ME, 2007. A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. Toxicological Sciences, 98, 240-248.

Thomas RS, Clewell HJ 3rd, Allen BC, Wesselkamper SC, Wang NC, Lambert JC, Hess-Wilson JK, Zhao QJ, Andersen ME, 2011.Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. Toxicological Sciences, 120,194-205.

Thomas RS, Clewell HJ, Allen BC, Yang LL, Healy E and Andersen ME, 2012. Integrating pathway-based transcriptomic data into quantitative chemical risk assessment: A five chemical case study. Mutation Research-Genetic Toxicology and Environmental Mutagenesis, 746, 135-143.

Thomas RS, Wesselkamper SC, Wang NCY, Zhao QJ, Petersen DD, Lambert JC, Cote I, Yang LL, Healy E, Black MB, Clewell HJ, Allen BC and Andersen ME, 2013a. Temporal Concordance between Apical and Transcriptional Points of Departure for Chemical Risk Assessment. Toxicological Sciences, 134, 180-194.

Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, Rowlands JC, Whelan MP, Hays SM, Andersen ME, Meek ME, Reiter LW, Lambert JC, Clewell HJ, Stephens ML, Zhao QJ, Wesselkamper SC, Flowers L, Carney EW, Pastoor TP, Petersen DD, Yauk CL and Nong A,

2013b. Incorporating New Technologies Into Toxicity Testing and Risk Assessment: Moving From 21st Century Vision to a Data-Driven Framework. Toxicological Sciences, 136, 4-18.

Tice R, 2011. Tox21: Transforming Environmental Health Protection LINCS Consortium Kick-off Meeting, Rockville, USA.

Tice RR, Austin CP, Kavlock RJ and Bucher JR, 2013. Improving the Human Hazard Characterization of Chemicals: A Tox21 Update. Environmental Health Perspectives, 121, 756-765.

Timchalk C and Poet TS, 2008. Development of a physiologically based pharmacokinetic and pharmacodynamic model to determine dosimetry and cholinesterase inhibition for a binary mixture of chlorpyrifos and diazinon in the rat. Neurotoxicology, 29, 428-443.

Tonnelier A, Coecke S and Zaldivar JM, 2012. Screening of chemicals for human bioaccumulative potential with a physiologically based toxicokinetic model. Archives of Toxicology, 86, 393-403.

Tornero-Velez R, Mirfazaelian A, Kim KB, Anand SS, Kim HJ, Haines WT, Bruckner JV and Fisher JW, 2010. Evaluation of deltamethrin kinetics and dosimetry in the maturing rat using a PBPK model. Toxicology and Applied Pharmacology, 244, 208-217.

Truong L, Reif DM, St Mary L, Geier MC, Truong HD and Tanguay RL, 2014. Multidimensional *in vivo* hazard assessment using zebrafish. Toxicological Sciences, 137, 212-233.

Tseng YJ, Hopfinger AJ and Esposito EX. 2012. The great descriptor melting pot: mixing descriptors for the common good of QSAR models. Journal of Computer-Aided Molecular Design, 26,39-43.

Turco L, Catone T, Caloni F, Di Consiglio E, Testai E and Stammati A, 2011. Caco-2/TC7 cell line characterization for intestinal absorption: how reliable is this *in vitro* model for the prediction of the oral dose fraction absorbed in human? Toxicology in Vitro, 25, 13-20.

US-EPA (US Environmental Protection Agency). 1996. Proposed guidelines for carcinogen risk assessment. Fed. Reg. 61: 17960–18011.

US-EPA (U.S. Environmental Protection Agency), 2002. Science Policy Council. Interim Policy on Genomics. Available online: http://www.epa.gov/OSP/spc/genomics.pdf

US-EPA (U.S. Environmental Protection Agency), 2004. Potential Implications of Genomics for Regulatory and Risk Assessment Applications at EPA. 70pp. Available online: http://www.epa.gov/osa/pdfs/EPA-Genomics-White-Paper.pdf

US EPA (U.S. Environmental Protection Agency). 2005. Guidelines for Carcinogen Risk Assessment.US EPA, Washington, DC. Available at:http://www.epa.gov/cancerguidelines/

US-EPA (U.S. Environmental Protection Agency), 2009. An Effects-based Expert System to Predict Estrogen Receptor Binding Affinity for Food Use Inert Ingredients and Antimicrobial Pesticides: Application in a Prioritization Scheme for Endocrine Disrupting Screening [Draft]. Office of Pesticide Programs, US EPA, Washington, DC. Available at http://www.regulations.gov/#!documentDetail;D = EPA-HQ-OPP-2009-0322-0002

US-EPA (U.S. Environmental Protection Agency), 2013. Next Generation Risk Assessment: Incorporation of Recent Advances in Molecular, Computational, and Systems Biology. Available online: http://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=259936#Download

Valcke M and Krishnan K, 2014. Characterization of the human kinetic adjustment factor for the health risk assessment of environmental contaminants. Journal of Applied Toxicology, 34(3): 227-240.

Van den Berg SJ, Klaus V, Alhusainy W and Rietjens IM, 2013. Matrix-derived combination effect and risk assessment for estragole from basil-containing plant food supplements (PFS). Food and Chemical Toxicology, 62, 32-40.

Van der Merwe D, Brooks JD, Gehring R, Baynes RE, Monteiro-Riviere NA and Riviere JE, 2006. A physiologically based pharmacokinetic model of organophosphate dermal absorption. Toxicological Sciences, 89, 188-204.

Van Ravenzwaay B, Herold M, Kamp H, Kapp MD, Fabian E, Looser R, Krennrich G, Mellert W, Prokoudine A, Strauss V, Walk T and Wiemer J, 2012. Metabolomics: a tool for early detection of toxicological effects and an opportunity for biology based grouping of chemicals-from QSAR to QBAR. Mutation Research, 746, 144-150.

Van Summeren A, Renes J, van Delft JH, Kleinjans JC and Mariman EC, 2012. Proteomics in the search for mechanisms and biomarkers of drug-induced hepatotoxicity. Toxicology in Vitro, 26, 373-385.

Vasdev N, Dorff PN, O'Neil JP, Chin FT, Hanrahan S and VanBrocklin HF, 2011. Metabolic stability of 6,7-dialkoxy-4-(2-, 3- and 4-[18F]fluoroanilino) quinazolines, potential EGFR imaging probes. Bioorganic and Medicinal Chemistry, 19, 2959–2965.

Vedani A and Dobler M, 2000. Multi-dimensional QSAR in drug research. Predicting binding affinities, toxicity and pharmacokinetic parameters. Progress in Drug Research, 55, 105-135.

Vedani A, Dobler M and Lill MA, 2006. The challenge of predicting drug toxicity *in silico*. Basic & Clinical Pharmacology & Toxicology, 99, 195-208.

Verner MA, Ayotte P, Muckle G, Charbonneau M and Haddad S, 2009. A Physiologically Based Pharmacokinetic Model for the Assessment of Infant Exposure to Persistent Organic Pollutants in Epidemiologic Studies. Environmental Health Perspectives, 117, 481-487.

Verner MA, McDougall R and Johanson G, 2012. Using population physiologically based pharmacokinetic modeling to determine optimal sampling times and to interpret biological exposure markers: The example of occupational exposure to styrene. Toxicology Letters, 213, 299-304.

Verner MA, McDougall R, Glynn A, Andersen ME, Clewell HJ and Longnecker MP, 2013. Is the Relationship between Prenatal Exposure to PCB-153 and Decreased Birth Weight Attributable to Pharmacokinetics? Environmental Health Perspectives, 121, 1219-1224.

Villeneuve D, Volz DC, Embry MR, Ankley GT, Belanger SE, Leonard M, Schirmer K, Tanguay R, Truong L and Wehmas L, 2014. Investigating alternatives to the fish early-life stage test: a strategy for discovering and annotating adverse outcome pathways for early fish development. Environmental Toxicology and Chemistry, 33, 158-169.

Vinken M, Landesmann B, Goumenou M, Vinken S, Shah I, Jaeschke H, Willett C, Whelan M and Rogiers V, 2013. Development of an Adverse Outcome Pathway from Drug-Mediated Bile Salt Export Pump Inhibition to Cholestatic Liver Injury. Toxicological Sciences, 136, 97-106.

Wang Z, Gerstein M and Snyder M, 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nature Review Genetics, 10, 57-63.

Washburn MP, Ulaszek RR and Yates JR 3rd, 2003. Reproducibility of quantitative proteomic analyses of complex biological mixtures by multidimensional protein identification technology. Analytical Chemistry, 75, 5054-5061.

West PR, Weir AM, Smith AM, Donley ELR and Cezar GG, 2010. Predicting human developmental toxicity of pharmaceuticals using human embryonic stem cells and metabolomics. Toxicology and Applied Pharmacology, 247, 18-27.

Wetmore BA, Wambaugh JF, Ferguson SS, Sochaski MA, Rotroff DM and Freeman K, 2012. Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. Toxicological Sciences, 125, 157-174.

Wetmore BA, Wambaugh JF, Ferguson SS, Li L, Clewell HJ 3rd, Judson RS, Freeman K, Bao W, Sochaski MA, Chu TM, Black MB, Healy E, Allen B, Andersen ME, Wolfinger RD and Thomas RS, 2013. Relative impact of incorporating pharmacokinetics on predicting *in vivo* hazard and mode of action from high-throughput *in vitro* toxicity assays. Toxicological Sciences, 132, 327-346.

WHO (World Health Organization), 2002. Genomics and World Health: Report of the Advisory Committee on Health research, Geneva.

WHO (World Health Organization), 2005. International Programme on Chemical Safety: Chemical-specific Adjustment Factors for Interspecies Differences and Human Variability: Guidance Document for Use of Data in Dose/concentration Response Assessment. World Health Organization: Geneva. Available online: http://www.who.int/ipcs/methods/harmonization/areas/uncertainty/en/index.html.

WHO (World Health Organization), 2010. Characterization and application of physiologically based pharmacokinetic models in risk assessment. Available online: http://www.who.int/ipcs/methods/harmonization/areas/pbpk_models.pdf?ua=1

Wiklund S, Johansson E, Sjostrom L, Mellerowicz EJ, Edlund U, Shockcor JP, Gottfries J, Moritz T and Trygg J, 2008. Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. Analytical Chemistry, 80, 15-122.

Wild CP, 2005. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiology, Biomarkers and Prevention, 14, 1847-1850.

Wild CP, 2012. The exposome: from concept to utility. International Journal of Epidemiology, 41, 24-32.

Wild CP, Scalbert A and Herceg Z, 2013. Measuring the exposome: a powerful basis for evaluating environmental exposures and cancer risk. Environmental Molecular Mutagenesis, 54, 480-499.

Wilkins MR, Sanchez JC, Gooley AA, Appel RD, HumpherySmith I, Hochstrasser DF and Williams KL, 1996. Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. Biotechnology and Genetic Engineering Reviews, 13, 19-50.

Williams TD, Mirbahai L and Chipman JK, 2014. The toxicological application of transcriptomics and epigenomics in zebrafish and other teleosts. Briefings in Functional Genomics. doi: 10.1093/bfgp/elt053

Wilmes A, Limonciel A, Aschauer L, Moenks K, Bielow C, Leonard MO, Hamon J, Carpi D, Ruzek S, Handler A, Schmal O, Herrgen K, Bellwon P, Burek C, Truisi GL, Hewitt P, Di Consiglio E, Testai E, Blaauboer BJ, Guillou C, Huber CG, Lukas A, Pfaller W, Mueller SO, Bois FY, Dekant W and Jennings P, 2013. Application of integrated transcriptomic, proteomic and metabolomic profiling for the delineation of mechanisms of drug induced cell stress. Journal of Proteomics, 79, 180-194.

Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, Djoumbou Y, Mandal R, Aziat F, Dong E, Bouatra S, Sinelnikov I, Arndt D, Xia J, Liu P, Yallou F, Bjorndahl T, Perez-Pineiro R, Eisner R, Allen F, Neveu V, Greiner R and Scalbert A, 2013. HMDB 3.0 - The Human Metabolome Database in 2013. Nucleic Acids Research, 41, D801-807.

Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, Cheng D, Jewell K, Arndt D, Sawhney S, Fung C, Nikolai L, Lewis M, Coutouly MA, Forsythe I, Tang P, Shrivastava S, Jeroncic K, Stothard P, Amegbey G, Block D, Hau DD, Wagner J, Miniaci J, Clements M, Gebremedhin M, Guo N, Zhang Y, Duggan GE, Macinnis GD, Weljie AM, Dowlatabadi R, Bamforth F, Clive D, Greiner R, Li L, Marrie T, Sykes BD, Vogel HJ and Querengesser L, 2007. HMDB: the Human Metabolome Database. Nucleic Acids Research, 35, D521-526.

Worth A, Fuart-Gatnik M, Lapenna S, Lo Piparo E, Mostrag-Szlichtyng A and Serafimova R, 2011. The use of computational methods in the toxicological assessment of chemicals in food: current status and future prospects. JRC report EUR 24748 EN. Publications Office of the European Union, Luxembourg. Available online: http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/publications/

Wu BJ, Dong D, Hu M and Zhang SX, 2013. Quantitative Prediction of Glucuronidation in Humans Using the *In Vitro-In Vivo* Extrapolation Approach. Current Topics in Medicinal Chemistry, 13, 1343-1352.

Wu SD, Blackburn K, Amburgey J, Jaworska J and Federle T, 2010. A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. Regulatory Toxicology and Pharmacology, 56, 67-81.

Xie G, Zheng X, Qi X, Cao Y, Chi Y, Su M, Ni Y, Qiu Y, Liu Y, Li H, Zhao A and Jia W, 2010. Metabonomic evaluation of melamine-induced acute renal toxicity in rats. Journal of Proteome Research, 9, 125-133.

Yeo KR, Kenny JR and Rostami-Hodjegan A, 2013. Application of *in vitro-in vivo* extrapolation (IVIVE) and physiologically based pharmacokinetic (PBPK) modelling to investigate the impact of the CYP2C8 polymorphism on rosiglitazone exposure. European Journal of Clinical Pharmacology, 69, 1311-1320.

Yokoi T and Nakajima M, 2013. microRNAs as mediators of drug toxicity. Annual Review of Pharmacology and Toxicology, 53, 377-400.

Yoon M, Campbell JL, Andersen ME and Clewell HJ, 2012. Quantitative *in vitro* to *in vivo* extrapolation of cell-based toxicity assay results. Critical Reviews in Toxicology, 42, 633-652.

Yoon M, Efremenko A, Blaauboer BJ and Clewell HJ, 2014. Evaluation of simple *in vitro* to *in vivo* extrapolation approaches for environmental compounds. Toxicology In Vitro, 28, 164-170.

Young JF, Luecke RH and Doerge DR, 2007. Physiologically based pharmacokinetic/pharmacodynamic model for acrylamide and its metabolites in mice, rats, and humans. Chemical Research in Toxicology, 20, 388-399.

Yu LR, 2011. Pharmacoproteomics and toxicoproteomics: The field of dreams. Journal of Proteomics, 74, 2549-2553.

Zaldivar Comenges JM, Menecozzi M, Macko P, Rodriguez R, Bouhifd M and Baraibar Fentanes J, 2011. A Biology-Based Dynamic Approach for the Modelling of Toxicity in Cell Assays: Part II: Models for Cell Population Growth and Toxicity. JRC Scientific and Technical Report EUR 24898 EN. Publications Office of the European Union, Luxembourg. Available from: http://publications.jrc.ec.europa.eu/repository/.

Zaldívar JM, Mennecozzi M, Marcelino Rodrigues R and Bouhifd M, 2010. A biology-based dynamic approach for the modelling of toxicity in cell-based assays. Part I: Fate modelling. EUR 24374 EN.

Zhang X, Tsang AM, Okino MS, Power FW, Knaak JB, Harrison LS and Dary CC, 2007. A physiologically based pharmacokinetic/pharmacodynamic model for carbofuran in Sprague-Dawley rats using the exposure-related dose estimating model. Toxicological Sciences, 100, 345-359.

Zhu H, Martin TM, Young DM and Tropsha A, 2009. Combinatorial QSAR Modeling of Rat Acute Toxicity by Oral Exposure. Chemical Research in Toxicology, 22, 1913-1921.

Zhu XW, Sedykh A, Zhu H, Liu SS and Tropsha A, 2013. The Use of Pseudo-Equilibrium Constant Affords Improved QSAR Models of Human Plasma Protein Binding. Pharmaceutical Research, 30, 1790-1798.

ABBREVIATIONS

2-DE	Two-dimensional gel electrophoresis
3-MCPD dipalmitate	3-chloropropane-1,2-dipalmitate
AA	Acrylamide
AA-GS	Glutathione conjugate of acrylamide
ABC	ATP binding cassette
ACAT	Advance Compartmental Absorption Transit
AchE	Acetylcholinesterase
ADAM	Advanced Dissolution model
ADI	Acceptable daily intake
ADME	Absorption, Distribution in the body, Metabolism and Excretion
ADMET	Absorption, Distribution, Metabolism, Excretion and Toxicity
AhR	Aryl hydrocarbon receptor
AOP	Adverse outcome Pathways
AOP-KB	AOP Wiki/Effectopedia Knowledge Base
ArfD	Acute Reference Dose
ATP	Adenosine triphosphate
ATSDR	Agency for Toxic Substances and Disease Registry
B2M	Beta-2 microglobulin
BBDR	Biologically-based dose-response
BMDL	Benchmark Dose Limit
BPA	Bisphenol A
b.w.	Body weight
CIMR	Core Information for Metabolomics Reporting
CLi	Intrinsic clearance
COSMOS	Coordination of Standards in MetabOlomicS
CPF	Chlorpyrifos
CSAF	Chemical-Specific Adjustment Factor
CYP	Cytochrome P-450
DDT	1,1,1-trichloro-2,2-bis(p-chlorophenyl)ethane
DG RTD	Directorate General for Research and Innovation
DNEL	Derived No-Effect Level
DZ	Diazinon
ECHA	European Chemicals Agency
EFSA	European Food Safety Authority
ESI-MS	Electrospray ionisation mass spectrometry
ESTs	Expressed Sequences Tags
FAO	Food and Agriculture Organization of the United Nations
FDA	U.S. Food and Drug Administration
GA	Glycidamide
GA-GS	Glutathione conjugate of glycidamide
GI	Gastrointestinal
GO	Gene Ontology
GWAS	Genome-wide association studies
Hb	Haemoglobin
HBGV	Health-based guidance value
HED	Human equivalent dose
HI	Hazard Index
HMDB	Human Metabolome Database
HMP	Human Microbiome Project
HTS	Highthroughput Screening
HTVMD	High-throughput Virtual Molecular Docking
IATA	Integrated Approach on Testing and Assessment
IE	Initiating Event

IKE	Intermediate Key Events
IPCS	International Programme on Chemical Safety
ITS	Integrated Testing Strategies
IVIVE	*In vitro* to *In vivo* extrapolation
JECFA	Joint Expert Committee on Food Additives
JRC	Joint Research Centre
KE	Key events
Km	Michealis Menten constant
LC–MS/MS	Liquid chromatography tandem mass spectrometry
LOAEL	Lowest-Observed-Adverse-Effect Level
MALDI	Matrix assisted laser desorption ionisation
MALDI-MS	Matrix-assisted laser desorption/ionization mass spectrometry
MeHg	Methylmercury
MIE	Molecular Initiating Event
MoA	Mode of Action
MOE	Margin of Exposure
MRM	Multiple Reaction Monitoring
MRPs	Multidrug resistance proteins
MSI	Metabolomics Standards Initiative
NCBI	National Center for Biotechnology Information
NCGC	National Chemical Genomics Center
NCM	Northern Contaminant Mixture
NGS	Next Generation Sequencing
NIH	National Institutes of Health
NMR	Nuclear Magnetic Resonance
NOAEL	No-Observed-Adverse-Effect Level
NRC	National Research Council
NTP	National Toxicology Program
OATPs	Organic anion transporting polypeptides
OCs	Organochlorine pesticides
OCTs	Organic Cation Transporters
OECD	Organisation for Economic Co-Operation and Development
OM	Oligonucleotide Microarrays
OPLS-DA	Orthogonal PLS-Discriminant Analysis
PB-PK	Physiologically-based pharmacokinetic models
PB-PK-PD	Physiologically-based pharmacokinetic pharmacodynamic models
PB-TK	Physiologically-based toxicokinetic models
PB-TK-TD	Physiologically-based toxicokinetic toxicodynamic models
PCA	Principal Component Analysis
PCBs	Polychlorinated Biphenyls
PCR	Polymerase Chain Reaction
PCIS	Precision Cut Slice Model
PFOA	Perfluorooctanoic acid
PFOS	Perfluorosulphonate
PLS	Partial Least Squares
PLS-DA	Partial least square discriminant analysis
PND	Post-natal day
POD	Point of departure
PON-1	Paraoxonase-1
POP	Persistent Organic Pollutant
Q	Quadrupole
QBAR	Quantitative Biological Activity Relationship
qPCR	Quantitative PCR
QSAR	Quantitative Structure Activity Relationships
RfC	Reference concentration

RfD	Reference dose
RP	Reference Point
RPF	Relative Potency Factor
RTK	Receptor Tyrosine Kinase
RVs	Reference Values
SAR	Structure Activity Relationships
SCCS	Scientific Committee on Consumer Safety
SCENIHR	Scientific Committee on Emerging and Newly Identified Health Risks
SCF	Scientific Committee on Food
SCHER	Scientific Committee on Health and Environmental Risks
SEURAT	Safety Evaluation Ultimately Replacing Animal Testing
SHH	Sonic Hedgehog
SLC	Solute carrier
SRM	Selective Reaction Monitoring
TCDD	Tetrachlorodibenzo-p-dioxin
TD	Toxicodynamic
TDI	Tolerable Daily Intake
TEF	Toxic equivalency Factors
TEQ	Toxic equivalent
TGF	Tumour Growth Factor
TK	Toxicokinetic
TOXPi	Toxicological Prioritisation index
TTC	Threshold of Toxicological Concern
TTD	Target-organ Toxicity Dose
TWI	Tolerable Weekly Intake
US-EPA	US-Environmental Protection Agency
Vmax	Maximum rate of catalysis
VSD	Virtual Safe Dose
WHO	World Health Organization
Wnt	Wingless-related integration site
WoE	Weight of evidence