



Contents lists available at ScienceDirect

Toxicology in Vitro

journal homepage: www.elsevier.com/locate/toxinvit

Review

In vitro to in vivo extrapolation for high throughput prioritization and decision making



Shannon M. Bell^a, Xiaoqing Chang^a, John F. Wambaugh^b, David G. Allen^a, Mike Bartels^{c,1}, Kim L.R. Brouwer^d, Warren M. Casey^e, Neepa Choksi^a, Stephen S. Ferguson^f, Grazyna Fraczekiewicz^g, Annie M. Jarabek^b, Alice Ke^h, Annie Lumenⁱ, Scott G. Lynn^j, Alicia Paini^k, Paul S. Price^b, Caroline Ring^{l,2}, Ted W. Simon^m, Nisha S. Sipes^f, Catherine S. Sprankle^a, Judy Strickland^a, John Troutmanⁿ, Barbara A. Wetmore^{o,3}, Nicole C. Kleinstreuer^{e,*}

^a Integrated Laboratory Systems, Inc., P.O. Box 13501, Research Triangle Park, NC 27709, USA

^b U.S. Environmental Protection Agency, 109 T.W. Alexander Dr., Research Triangle Park, NC 27709, USA

^c Dow Chemical Company, Saginaw, MI 48603, USA

^d UNC Eshelman School of Pharmacy, The University of North Carolina at Chapel Hill, Campus Box 7569, Chapel Hill, NC 27599, USA

^e National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709, USA

^f National Toxicology Program, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709, USA

^g Simulations Plus, Inc., 42505 10th St. West, Lancaster, CA 93534, USA

^h Simcyp Limited (a Certara company), John Street, Sheffield, S2 4SU, United Kingdom

ⁱ National Center for Toxicological Research, U.S. Food and Drug Administration, 3900 NCTR Road, Jefferson, AR 72079, USA

^j U.S. Environmental Protection Agency, William Jefferson Clinton Building, 1200 Pennsylvania Ave. NW, Washington, DC 20460, USA

^k European Commission, Joint Research Centre, Directorate Health, Consumers and Reference Materials, Chemical Safety and Alternative Methods Unit incorporating EURL ECVAM, Via E. Fermi 2749, Ispra, Varese 20127, Italy

^l Oak Ridge Institute for Science and Education, P.O. Box 2008, Oak Ridge, TN 37831, USA

^m Ted Simon LLC, 4184 Johnston Road, Winston, GA 30187, USA

ⁿ Central Product Safety, The Procter & Gamble Company, Cincinnati, OH 45202, USA

^o ScitoVation LLC, 6 Davis Drive, Research Triangle Park, NC 27709, USA

ARTICLE INFO

Keywords:

IVIVE
High throughput testing
Toxicity testing
Toxicokinetic
Computational toxicology
Risk assessment

ABSTRACT

In vitro chemical safety testing methods offer the potential for efficient and economical tools to provide relevant assessments of human health risk. To realize this potential, methods are needed to relate in vitro effects to in vivo responses, i.e., in vitro to in vivo extrapolation (IVIVE). Currently available IVIVE approaches need to be refined before they can be utilized for regulatory decision-making. To explore the capabilities and limitations of IVIVE within this context, the U.S. Environmental Protection Agency Office of Research and Development and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods co-organized a workshop and webinar series. Here, we integrate content from the webinars and workshop to discuss activities and resources that would promote inclusion of IVIVE in regulatory decision-making. We discuss properties of models that successfully generate predictions of in vivo doses from effective in vitro concentration, including the experimental systems that provide input parameters for these models, areas of success, and areas for improvement to reduce model uncertainty. Finally, we provide case studies on the uses of IVIVE in safety

Abbreviations: ADME, chemical absorption, distribution, metabolism, and excretion; AOP, adverse outcome pathway; fu, fraction of chemical unbound; HTS, high throughput screening; IVIVE, in vitro to in vivo extrapolation; PBPK, physiologically based pharmacokinetic; QSAR, quantitative structure-activity relationship; TK, toxicokinetics

* Corresponding author at: National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709, USA.

E-mail addresses: SBell@ils-inc.com (S.M. Bell), XChang@ils-inc.com (X. Chang), wambaugh.john@epa.gov (J.F. Wambaugh), DAllen@ils-inc.com (D.G. Allen), mjbartels@toxmetrics.com (M. Bartels), kbrouwer@unc.edu (K.L.R. Brouwer), warren.casey@nih.gov (W.M. Casey), NChoksi@ils-inc.com (N. Choksi), Stephen.ferguson@nih.gov (S.S. Ferguson), grace@simulations-plus.com (G. Fraczekiewicz), jarabek.annie@epa.gov (A.M. Jarabek), alice.ke@certara.com (A. Ke), annie.lumen@fda.hhs.gov (A. Lumen), lynn.scott@epa.gov (S.G. Lynn), Alicia.paini@ec.europa.eu (A. Paini), price.pauls@epa.gov (P.S. Price), cring@toxstrategies.com (C. Ring), ted@tedsimon-toxicology.com (T.W. Simon), Nisha.sipes@nih.gov (N.S. Sipes), CSprankle@ils-inc.com (C.S. Sprankle), JStrickland@ils-inc.com (J. Strickland), troutman.ja@pg.com (J. Troutman), Wetmore.barbara@epa.gov (B.A. Wetmore), nicole.kleinstreuer@nih.gov (N.C. Kleinstreuer).

¹ ToxMetrics.com, 2407 Wilmington Drive, Midland, MI 48642 USA. Email: mjbartels@toxmetrics.com.

² ToxStrategies, Austin, TX (78759), USA. Email: cring@toxstrategies.com.

³ U.S. Environmental Protection Agency, 109 T.W. Alexander Dr., Research Triangle Park, NC 27709, USA. Email: Wetmore.barbara@epa.gov.

<https://doi.org/10.1016/j.tiv.2017.11.016>

Received 2 October 2017; Received in revised form 28 November 2017; Accepted 30 November 2017

Available online 05 December 2017

0887-2333/ Published by Elsevier Ltd.

assessments, which highlight the respective differences, information requirements, and outcomes across various approaches when applied for decision-making.

1. Introduction

As illustrated by environmental incidents such as the 2010 Deepwater Horizon oil spill (Judson et al., 2010) and the 2014 Elk River chemical spill (NTP, 2016), health hazard information is lacking for many commercial chemicals. To obtain this information, test guidelines currently accepted by regulatory authorities often use animal-based test methods. However, animal-based methods are expensive and time-consuming; using them to generate safety information for the tens of thousands of chemicals in commerce requiring such information is not practical. Therefore, there is strong interest in developing efficient and economical tools that use mechanistically anchored in vitro assays to predict human health effects.

In vitro to in vivo extrapolation (IVIVE) can be broadly defined as an approach utilizing in vitro experimental data to predict in vivo phenomena, including concentrations (e.g., toxicokinetics [TK]) and effects (e.g., toxicodynamics). Bioactivity data from in vitro high throughput screening (HTS) assays, used in programs such as Tox21 (Tice et al., 2013) and ToxCast (Kavlock et al., 2012), are potentially useful for predicting chemical effects. In the ten-year duration of the Tox21 and ToxCast programs, approximately 8000 chemicals have been screened using hundreds of assays. However, relating nominal concentrations that induce effects in in vitro HTS assays to exposures that cause adverse effects in humans (or other target organisms) is not straightforward and depends on TK. Consideration of TK is also needed to address the impact of protein binding, bioavailability, and clearance of the chemical and other aspects of chemical kinetics in vitro (Blaauboer, 2010; Coecke et al., 2013; Frazier, 1995; Leeson, 1995; Lipscomb et al., 1998). IVIVE approaches to predicting chemical TK can be used in conjunction with reverse dosimetry to estimate the in vivo dose required to achieve an in vitro bioactive concentration in the blood or target tissue. The predicted in vivo dose estimate can then be compared to human exposures (actual or estimated) to assess human health risk (Thomas et al., 2013; Wetmore et al., 2012, 2015).

Regulatory authorities are developing strategies to utilize in vitro data in human health risk assessment. In 2015, the European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) published its strategy to replace, reduce, and refine the use of animals in the assessment of TK and systemic toxicity (Bessemers et al., 2015). This document laid out four strategic aims (Fig. 1) to facilitate

the generation and use of non-animal data to assess human risk:

- Development and standardization of in vitro methods to measure chemical absorption, distribution, metabolism, and excretion (ADME);
- Establishment of good kinetic modeling practices and web-based kinetic modeling portals;
- Improved resources for collection, storage, and reporting of ADME/TK data that can be used for developing computational models or applied in the risk assessment process; and,
- Guidance on the use of human ADME/TK in integrated approaches for testing and assessment.

Similarly, the U.S. Environmental Protection Agency (EPA) has identified high throughput in vitro assays and computational tools to be used “as an alternative for some of the current assays” in its Endocrine Disruptor Screening Program (EPA, 2015). The application of IVIVE will allow these assays and tools to be used for prioritization of chemicals for further testing. The EPA efforts are complementary to activities proposed in the EURL ECVAM TK report and point to a growing momentum for the use of IVIVE in regulatory decision-making.

To further explore this trend and identify areas in need of additional data or research, the EPA Office of Research and Development and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) convened a workshop in February 2016. A preceding series of webinars focusing on the role of TK in IVIVE analyses laid the groundwork for the in-person event (NTP, 2017).

This report, based on the presentations and discussions at the February 2016 workshop, describes some considerations and resources needed to apply IVIVE to risk prioritization in regulatory contexts. We first describe new in vitro methods for obtaining data needed for model development, as well as computational platforms that use these data to characterize and decrease model uncertainty and address biological variability. We then describe a general workflow for IVIVE analysis and provide examples that highlight the application of IVIVE to prioritization and risk-based decision-making. We conclude by discussing how to increase confidence in use of IVIVE in regulatory contexts.

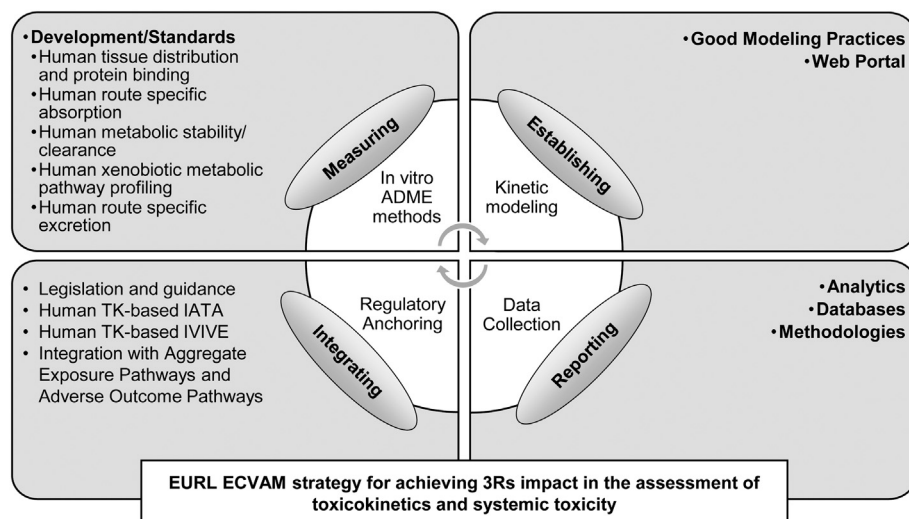


Fig. 1. Schematic overview of the EURL ECVAM strategy for promoting the use of non-animal approaches in the assessment of toxicokinetics and systemic toxicity. IATA, integrated approaches for testing and assessment.

2. Overview of model construction for IVIVE analyses and applications

Computational models used in IVIVE analyses consider two aspects of chemical action: TK and toxicodynamics. TK characterizes the ADME processes of a chemical after it enters the body, while toxicodynamics focuses on the dynamic interactions of a chemical with its target site.

TK can be distinguished from pharmacokinetics in that TK focuses on the relationship between exposure to a chemical and its potential toxicity while pharmacokinetics focuses on ADME for therapeutic compounds. Both TK and pharmacokinetic processes can be described using ordinary differential equations. A simplified description of chemical ADME can be obtained using a one-compartment model, while a detailed representation that captures these processes across multiple tissue compartments can be achieved using a physiologically based pharmacokinetic (PBPK) model.

While understanding of both TK and toxicodynamics is essential for using in vitro assay data to develop mechanistic knowledge and mechanism-based predictions of chemical action, TK processes are more relevant for the purposes of screening to estimate toxicity and identification of a margin of exposure for regulatory purposes. Accordingly, the February 2016 workshop focused on the TK modeling component of IVIVE analyses for the application of these analyses in current regulatory contexts.

2.1. Toxicokinetics and reverse dosimetry

IVIVE analysis requires (1) an in vitro system that provides an estimate of the active concentration of a test chemical and (2) a TK model to relate this active concentration to an in vivo dose that would produce an equivalent concentration in blood or tissue. Traditionally, TK models have been used to calculate tissue concentrations resulting from a known exposure, termed “forward dosimetry” (Caldwell et al., 2012). These same models can also be used for calculating “reverse dosimetry,” a concept central to IVIVE. Reverse dosimetry uses the measured or targeted internal chemical concentrations (e.g., plasma or tissue) as a starting point in a TK model to calculate a possible exposure dose (Tan et al., 2007).

Fig. 2 (adapted from Judson et al., 2011) shows a basic workflow of IVIVE analysis, illustrating how TK models and in vitro data can be integrated to predict doses that cause in vivo effects. The TK model can be used to extrapolate between species and exposure routes by

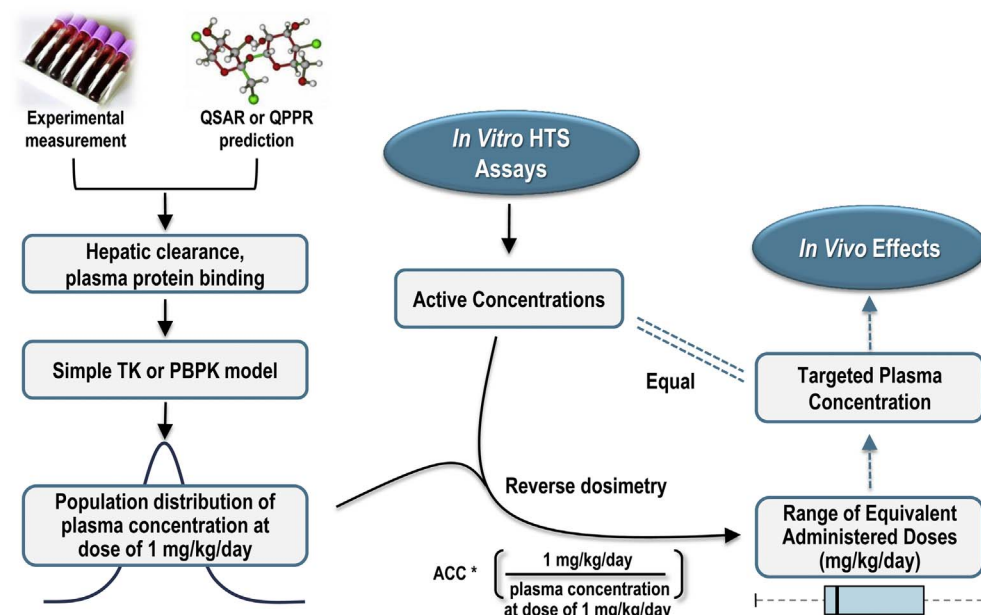


Fig. 2. Overview of in vitro to in vivo extrapolation. Parameters describing ADME processes of the chemical through the system (i.e., hepatic clearance, protein binding) may be obtained via experimental measurements or in silico predictions. These parameters are used to develop a one compartment TK or a PBPK model that can be used to predict the population distribution of plasma concentration from any given daily dose. Reverse dosimetry predicts administered doses equivalent to in vitro active concentration, which can be compared to the in vivo measurements.

incorporating species-specific physiology and route-specific parameters, which allows for evaluations of in vivo data in different test species as well as via multiple exposure routes. Depending on the TK outputs needed and the decision context, a range of cheminformatic data, experimental data, and/or modeling tools can be selected to predict desired chemical concentrations. These inputs and tools can be used in reverse dosimetry approaches to relate these internal concentrations to an external exposure metric.

PBPK models consist of multiple tissue compartments connected by blood flow. The tissue compartments are described by parameters defining tissue volumes, blood flows, tissue components, metabolism, and transport processes. As noted previously, reverse dosimetry can use either simple pharmacokinetic models consisting of one or two compartments that represent lumped tissues (O’Flaherty, 1981) or by PBPK models consisting of multiple tissue compartments (Campbell et al., 2012; Yoon et al., 2012). The PBPK models can also vary in complexity, ranging from models relying on only in vitro data (Jamei et al., 2009; Lukacova et al., 2009; Wambaugh et al., 2015; Wilkinson and Shand, 1975) to those supported by extensive in vivo data characterizing the distribution and disposition of chemicals among organs over space and time (Caldwell et al., 2012; Loizou and Hogg, 2011; McLanahan et al., 2012; Slob et al., 1997).

Traditional approaches to building PBPK models estimate model parameters using experimentally generated time-course data for the plasma and tissue concentrations of a chemical. However, obtaining these data is costly and time-consuming, often requiring animal measurements, the use of radiolabeled chemicals (Pellegatti, 2014), and precise analytical chemistry methods (Tolonen and Pelkonen, 2015). High throughput TK models based on chemical properties and in vitro high throughput assay data represent a more practical and efficient approach to PBPK modeling (Fig. 3). This approach has been used successfully by the pharmaceutical industry in preparation for human clinical trials (Jamei et al., 2009; Lukacova et al., 2009; Rowland et al., 2011; Wang, 2010) and by chemical manufacturers as an alternative to animal testing to assess potential hazards of environmental chemicals (Wetmore et al., 2012, 2015).

2.2. Using in vitro active concentration in TK models

As can be seen in Fig. 2, the active concentration in the in vitro assay is a key component relating the in vitro assay dose response data to plasma concentration. The active concentration can be expressed in a

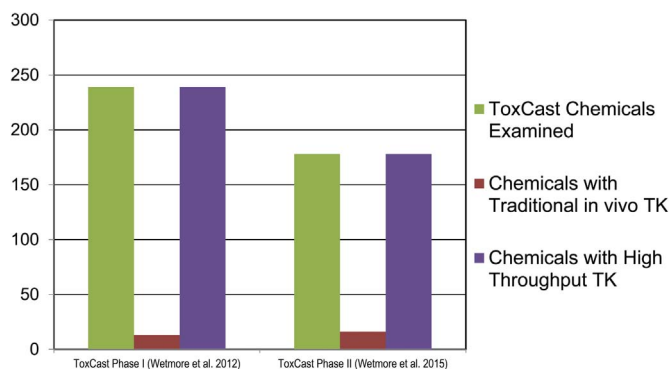


Fig. 3. Toxicokinetic information for ToxCast chemicals. Histogram compares the number of chemicals in Phases I and II of EPA's ToxCast high throughput screening project having traditional toxicokinetics (TK) data versus those with high throughput TK data.

variety of forms, including point of departure (the concentration where the response exceeds the assay-dependent noise threshold), activity concentration at cut-off (which employs curve-fitting models to identify an activity threshold for an assay response), or concentration of chemical that induces a half-maximal response (Filer et al., 2017). Many current approaches to IVIVE base estimates of the bioactive chemical concentration on the nominal concentration rather than considering or adjusting for non-specific binding affecting the amount of chemical available to elicit an effect in the assay. This is an important consideration, as several studies have demonstrated differential partitioning and availability of chemicals in in vitro toxicity assay systems (Armitage et al., 2014; Fischer et al., 2017; Groothuis et al., 2015; Kramer et al., 2015; Mundy et al., 2004; Zaldivar Comenges et al., 2017). Additionally, relating the magnitude of change in an in vitro assay response to prediction of a measurable response in vivo may require an informed understanding of the assay dose response behaviors in conjunction with consideration of chemical- or chemical group-specific assay responses.

2.3. Considerations for application of TK models to IVIVE

Pharmacokinetic models were primarily developed for understanding the fate of pharmaceuticals in biological systems. When applying TK models derived from pharmacokinetic models to environmentally relevant substances, it is important to consider the differences between therapeutic drugs and environmental chemicals. These differences can include concentration ranges assessed in toxicology vs. pharmacology studies, physicochemical properties, and ADME behaviors such as bioaccumulation potential. Adequate understanding of the rationale behind TK model assumptions applied for one space helps ensure appropriate decisions are made for its use for another space, where some of the assumptions may or may not be applicable.

A longstanding limitation to the adoption of in vitro assay data in chemical safety assessments has been the inability to relate nominal concentrations at which biological activity at a specific target is observed to a relevant in vivo exposure level. However, such a metric is provided with the incorporation of chemical kinetics (Rotroff et al., 2010; Thomas et al., 2013; Wambaugh et al., 2015). Additional computational tools can be employed to refine the parameterization of TK models and potentially expand or modify to consider relevant exposure scenarios (Dansirikul et al., 2005; Fouchécourt et al., 2001; Fujiwara et al., 2003; Votano et al., 2004). Again, appropriate consideration and assessment of uncertainty during the development and application of these TK models are needed to determine what data streams and modeling are appropriate for the relevant decision context (Fig. 4).

3. Experimental systems for obtaining input parameters

As noted above, several in vitro assays currently used to study drug metabolism can be used to generate input parameters for TK models. These assays vary in complexity, metabolic pathway coverage, and physiological relevance (Table 1), and range from inexpensive and high throughput to resource-intensive and low throughput. Unfortunately, the ability of an assay to represent the biological system of interest tends to be inversely related to simplicity and throughput.

The in vitro metabolizing system used to generate input data for a TK model should be appropriate for the research question at hand. For example, systems such as primary hepatocytes, liver S9, liver microsomes, and recombinant enzymes can be used to determine whether a chemical will be metabolized into alternate chemical structures by the liver (Chiba et al., 2009). Specific inhibitors of major drug metabolism pathways can be employed in these systems to identify important pathways involved in metabolic clearance (Harper and Brassil, 2008). Experiments using primary hepatocytes in culture, HepaRG cultures, or human liver microsomes can assess the potential for a chemical to induce or inhibit liver enzymes that alter drug or chemical clearance (Ferguson and Bonzo, 2014).

To accurately predict both systemic and tissue-level exposure to xenobiotics and their metabolites, it is important to consider hepatobiliary clearance as well as metabolism. The appropriate model to assess this process depends upon the specific element that needs to be predicted: hepatic uptake, hepatic clearance, biliary excretion, biliary clearance, hepatocyte accumulation, hepatotoxicity, or hepatic transporter activity and interactions. Several non-hepatocyte systems are commonly used to determine hepatobiliary clearance mechanisms for a particular chemical (Brouwer et al., 2013). Recombinant cell lines expressing uptake transporters are high throughput systems with low physiological relevance, but can be used to evaluate drug interactions with uptake transporters, determine substrate specificity, and identify uptake inhibitors. Polarized layers of recombinant cells expressing transporters with bidirectional transport functionality can be used to compare active transport with diffusion. Because these models only measure transport across a membrane, nonspecific binding is less of a concern. Membrane vesicle-based transporter assays are ideal for

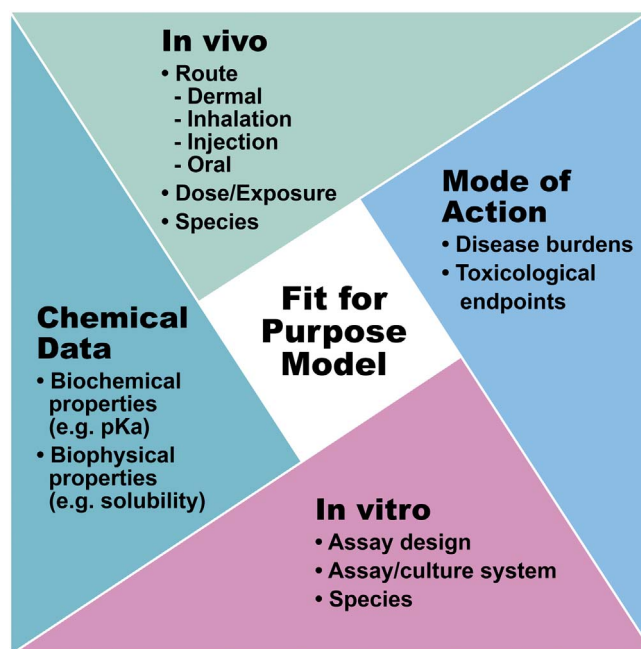


Fig. 4. Schematic of fit for purpose modeling. Building from Fig. 2, there are multiple components to be considered to assess if a model is appropriate and will provide the level of confidence needed for the intended use.

Table 1
Assay systems to predict metabolic clearance.

	Metabolic competency	Longevity	Barrier function	Physiological relevance	Costs	Throughput/ease of use
Isolated perfused organs	XXXX	X	XXXX	XXXX	XXXX	X
Tissue slices	XXX	X	X	XXX	XXXX	X
Primary cells (immediately isolated or cryopreserved from tissues)	XXX	X	X	X	XX	XXXX
Primary cells in culture	XX	XX	XXX	XXX	XX	XX
Differentiated cell lines (with barrier function)	N/A	N/A	XXX	N/A	X	XXX
Immortalized cell lines	X	XXXX	X	X	X	XXXX
Subcellular fractions (i.e., Microsomes/S9)	XXX	X	N/A	X	X	XXXX
Recombinant enzymes	XX	X	N/A	X	X	XXXX

Table cell contents represent a continuum of relevant characteristics, with “X” indicating low or minimal and “XXXX” indicating high or extensive. Use of each test system for a particular application should be considered within the context of the goals of the experiment.

identifying substrates for efflux transporters. These assays typically use membranes derived from recombinant cells expressing high levels of the transporter of interest. These systems are suitable for chemicals with low permeability and are not impacted by compounds that are cytotoxic. All three of these systems can be used in high throughput systems and are readily available, but are limited in their ability to predict how transport of a chemical might occur in vivo.

Suspension cultures of primary human hepatocytes have the capacity for Phase II metabolism and de novo cofactor synthesis, making them superior to liver microsomes or S9 systems for assessing metabolism. These cultures are also the gold standard for assessing species-specific active uptake and passive diffusion (Brouwer et al., 2013) and can be used to develop predictions of intrinsic clearance that are generally within two- to three-fold of experimental values for in vivo metabolic clearance for drug-like substances (Hallifax and Houston, 2012). Suspension hepatocyte cultures support a broad complement of metabolizing enzymes and active uptake transport, and maintain physiological levels of metabolic competence in appropriate proportions. Several metabolic pathways can be identified in these systems by monitoring loss of parent chemical.

However, suspension hepatocyte cultures have weaknesses for estimating both metabolism and hepatic clearance. These assay systems often do not include physiologically relevant levels of plasma proteins. For studies involving highly protein-bound compounds, this limitation may result in poor predictions of metabolic clearance due to inaccurate estimation of the amount of unbound compound available. Suspension hepatocytes can only be maintained for a few hours in detached form before metabolic activity and cell viability wane. The brief duration of metabolic competence limits the usefulness of these cultures for analysis of lower turnover compounds (Smith et al., 2012). Next-generation culture models begin to address this issue, but generally exhibit reduced levels of metabolic competence compared to primary human hepatocyte suspensions (Jackson et al., 2016). Suspension hepatocyte cultures also lack canalicular efflux transport capability (Bow et al., 2008), have limited basolateral efflux, and in general do not contain all the relevant transporters and other physiological machinery existing in living organisms. Furthermore, these systems may not accurately predict which proteins predominantly transport the chemical in the whole liver, and they have a limited ability to assess complex xenobiotic-transport interactions such as non-competitive mechanisms and metabolite interactions. Each of these limitations may lead to inaccurate estimates of metabolic outcomes, especially at lower substrate concentrations.

Spheroid cultures offer substantial improvements over suspensions of primary human hepatocytes (C. C. Bell et al., 2016; Lee et al., 2013). These cultures can be maintained longer than the typical four days for primary human hepatocyte cultures with a higher level of metabolic competence. Spheroid cultures also offer the potential to be effective models for integrating metabolism predictions measuring intrinsic clearance, metabolite formation/profiling, and transport/accumulation within a single tissue-like model system. Additional research is needed

to characterize the extent of the metabolic competence of these cultures and to anchor metabolic competency to metabolite profiles that accurately reflect stages of cell/tissue differentiation and development (e.g., neonatal hepatocytes).

Sandwich-cultured hepatocytes provide species-specific expression of uptake and efflux transporters that resembles in vivo expression. They also preserve functional metabolic enzymes and regulatory machinery (Swift et al., 2010). These holistic systems regain normal cell polarity and can be used to identify competitive and non-competitive transport inhibitors and inducers. Sandwich-cultured hepatocytes effectively replicate both human and laboratory animal biliary clearance (Ghibellini et al., 2007; Li et al., 2010; Liu et al., 1999). B-CLEAR® is a technology using sandwich-cultured hepatocytes that accurately measures hepatobiliary clearance for IVIVE along with intracellular concentrations of both test chemical and generated metabolites (Chu et al., 2013).

Essential components of a predictive hepatic model for human risk assessment include functional uptake and efflux transporters, metabolic capacity, and intact regulatory machinery. However, transport and metabolic proteins and their regulation may differ by species.

Finally, while current TK models typically do not account for metabolites, in reality characterization of both parent chemical and metabolites are often needed for accurate model parameterization. Emerging data indicate that some mechanisms of chemical-induced hepatotoxicity involve interaction of both parent chemical and/or metabolites with bile acids, which are important cell signaling molecules (Yang et al., 2014, 2015).

Quantitative systems pharmacology models that integrate species-specific physiological information and experimental data from sandwich-cultured hepatocytes and membrane vesicle systems can accurately predict the clinical incidence and time-course of drug-induced hepatotoxicity (Woodhead et al., 2014, 2017). Additional research is needed to characterize (1) the assay conditions that are most predictive for IVIVE, (2) the cellular regulatory mechanisms that impact hepatobiliary clearance, and (3) whether more complex models may improve the accuracy of the predictions.

4. Computational tools for developing TK and PBPK models

Many chemicals lack ADME data that are needed to construct a biologically relevant relationship between HTS bioactivity data and external effective dose. One way to address this gap is use predicted values as inputs for provisional PBPK models. Quantitative structure-activity relationship (QSAR) models may be used to generate the predicted values. The Swiss Institute of Bioinformatics has compiled a directory of software, web services, and databases available for in silico prediction of ADME parameters on its Click2Drug page (Swiss Institute of Bioinformatics, 2013). Multiple algorithms have also been developed for predicting ADME input parameters such as partition coefficients (Haddad et al., 2000; Peyret et al., 2010; Schmitt, 2008) or protein

binding (Zhu et al., 2013). Characterization of metabolism in this manner is thus far limited to cytochrome P450 (CYP) pathways; robust in silico approaches for predicting rates of non-CYP metabolism are not currently available.

A model's domain of applicability is limited to the chemical space used in model development, which for these QSAR models is most often pharmaceutical chemicals. The overlap between pharmaceutical and non-pharmaceutical chemicals in terms of chemical properties is unclear. As more data are collected for non-pharmaceuticals, new QSAR models that perform better for diverse chemical structures can be constructed. The following sections describe some case studies of approaches that use computational predictions of chemical properties as inputs for IVIVE analyses.

4.1. In silico screening of primary clearance mechanisms using the extended clearance classification system

As alluded to above, model development can be limited by the availability of high-quality data from in vitro and in vivo ADME studies for model parameterization, calibration, and evaluation. To address this limitation and integrate non-animal hazard data into emerging risk assessment frameworks, TK models have been developed for dermal exposure that reliably estimate internal dose from realistic external exposure scenarios using in silico and in vitro hazard data (Dancik et al., 2015). These scenarios consider skin surface area, duration and frequency of exposure, and the degree of occlusion of the area of skin that has been treated with chemical. Reverse dosimetry is then applied to obtain the relevant dose metrics for quantitative risk assessment.

A generic PBPK model was previously developed using in silico estimates of chemical-specific inputs to conservatively model tissue uptake and subsequent elimination by the liver and kidney (Dancik et al., 2013). However, the high throughput dosimetry models needed for screening thousands of chemicals may need a more accurate characterization of major elimination pathways, which as noted earlier involve several complex processes. This information on elimination pathway mechanisms is available through the Extended Clearance Classification System (Varma et al., 2015). This system, originally designed for drug studies, uses in silico predictions of physicochemical properties and passive membrane permeability to predict the rate-determining clearance mechanism for a chemical. These predictions can then be applied to model development. Further development will allow the Extended Clearance Classification System to be applied to cosmetics and consumer product ingredients. This will facilitate construction of TK models to identify information gaps in this sector and guide selection of relevant in vitro studies for screening-level predictions of internal dosimetry.

In addition to guiding study design, in vitro workflows can also inform prioritization of new chemicals for further development. Procter & Gamble has developed a workflow using commercially available models from ACD/Labs (Toronto, Canada) that use chemical structure to calculate passive permeability across Caco-2 cell monolayers and jejunum epithelium. The distribution coefficient and polarizability of the chemical are used to determine the role of transporters and passive diffusion in moving substances across cellular membranes. The Extended Clearance Classification System protocol then calculates chemical properties including molecular weight, permeability input and ionization state for qualitative predictions of primary clearance mechanisms. The prediction and exposure assumptions are used as inputs to multi-route PBPK models that are specific for species, sex, and age, and produce estimates of plasma concentrations of the target chemical. The screening-level information produced may also be useful for assessing the need for additional data generation when greater accuracy is required, and to inform decisions about what types of measured data would be most useful (e.g., in vitro metabolism studies or renal excretion mechanisms).

4.2. Simcyp (Certara USA, Inc.)

It is important to recognize variation in TK parameters among different populations. For example, the effect of age propagates variability to many system parameters, including body surface area, cardiac output, plasma protein, hematocrit, renal function, microsomal protein per gram of liver, etc. The commercially available Simcyp® Simulator (Jamei et al., 2009) provides built-in mechanistic, physiological multi-compartmental tissue models (e.g., brain, kidney, liver, lung, skin and intestine) and various databases to estimate intracellular tissue concentrations of test chemicals from in vitro and physicochemical properties. Moreover, this platform specifically accounts for population variability in parameters such as age, sex, genetics, and disease status, and predicts TK characteristics (e.g., CYP-, UGT-, or esterase-based enzyme metabolism) of compounds in various populations. Simcyp includes a full PBPK model together with extensive libraries on pediatric demography, developmental physiology, and biochemistry. As such, Simcyp is able to explicitly model the complexity of covariate effects as applied to the prediction of clearance and volume of distribution.

Fraction of chemical unbound (f_u), an important parameter for IVIVE analysis, is affected by factors that can vary highly among age groups such as dissociation constant of chemical to the binding protein and the plasma protein concentrations. These factors need to be mechanistically considered when, for example, using plasma protein binding characteristics of an adult population to estimate f_u for infants and children. Simcyp has recently updated the ontogeny functions of drug metabolizing enzymes (Salem et al., 2014), resulting in markedly improved performance in the prediction of age-related clearances for CYP3A4-metabolized drugs such as alfentanil and midazolam. The application of PBPK modeling using the Simcyp Simulator or similar tools to extrapolate efficacy and/or safety data to the pediatric population is encouraged by regulatory bodies (Wagner et al., 2015). Wetmore and colleagues successfully combined isozyme and physiologic differences to quantitate subpopulation pharmacokinetic variability and to estimate subpopulation-specific oral equivalent doses (Wetmore et al., 2014).

In addition to modeling in vivo systems in both healthy and specialized human populations, Simcyp also offers capabilities to model in vitro assay systems (Poirier et al., 2009) via the Simcyp In Vitro (Data) Analysis (SIVA) Toolkit. SIVA Toolkit includes various modules including a two-compartment model for incubations using suspension hepatocyte cultures to estimate chemical movement from incubation medium and intracellular space through passive diffusion and active uptake. For permeability or transport assays using a transwell system, SIVA Toolkit also provides three- and five-compartment models that can be used to describe chemical movement between the apical and basolateral membranes. While conventional in vitro analysis provides a single, global permeability value, the SIVA Toolkit generates parameters describing each of the physiological processes including passive diffusion, active transport, and metabolism.

4.3. GastroPlus and ADMET predictor (Simulations Plus, Inc.)

GastroPlus™ is a commercially available mechanistically based PBPK modeling software package that simulates the TK and toxicodynamics of chemical absorption via the intravenous, oral, oral cavity, ocular, pulmonary, and dermal routes in humans and animals. GastroPlus utilizes the ACAT™ (Advanced Compartmental Absorption and Transit) model to simulate intestinal absorption of chemicals along the gastrointestinal tract. The ADMET Predictor™ module of GastroPlus contains over 140 QSAR models to predict various physicochemical, biopharmaceutical, pharmacokinetic, and cytochrome P450 metabolism parameters required as inputs for PBPK models. ADMET Predictor also predicts sites of metabolism, metabolites and various toxicities.

Several studies document the prediction accuracy for specific endpoints using GastroPlus and ADMET Predictor. In one study using

properties measured in vitro for 18 chemicals, the predicted total chemical exposure over time for seven chemicals was within two-fold of values derived from published human in vivo TK data, and predicted values for the remaining 11 chemicals were within five-fold of in vivo data (Zhou et al., 2016). In another study conducted with 62 chemicals, human oral bioavailability estimated using QSAR and PBPK models was within two-fold of in vivo values for 70% of the chemicals (Lawless et al., 2015). GastroPlus also performs well at predicting volume of distribution if the correct physicochemical and biopharmaceutical properties are provided. Even with only in silico inputs, volume of distribution can be predicted within two-fold of in vivo values for chemicals undergoing only passive renal or hepatic cytochrome P450 clearance. However, there are still many areas for which accuracy could be improved. For example, it is difficult to predict rates of biliary clearance or transporter-based clearance in liver or kidney, which often results in under prediction of clearance in IVIVE analysis.

4.4. Resource needs for computational tools

A breakout group at the February 2016 workshop agreed that computational tools and approaches are critical components that facilitate the use of IVIVE to address current toxicological challenges. This group focused on identifying the resources needed to better utilize these tools and encourage their acceptance in regulatory applications.

Two key types of data are needed to develop and evaluate models: input parameters to develop the models and high-quality reference data to assess prediction accuracy. A database housing shared in vitro and in vivo TK data was identified as a key resource needed to support future development of more useful and broadly-accepted IVIVE analyses. Such a warehouse should support inclusion of data from peer-reviewed publications and individual curated submissions. The data and any models used should be machine-readable to increase transferability. An early activity in the development of a database will be to establish standards for data formatting and annotation required of housed data analogous to the MIAME standards for microarray reporting (Brazma et al., 2001). These standards specify the minimal amount of information needed for facilitating machine-readable collection, which will optimize data reuse. To increase reporting of negative data (data showing no toxicity), it was proposed that a digital object identifier (DOI) for deposited data could be provided to database contributors. Receipt of a DOI for submitted data would enable researchers to share and receive acknowledgement for their contribution to the field without having to publish in a peer-reviewed publication, which can be hard to accomplish with negative data.

Information on exposure route (e.g., oral dosing, dermal absorption, and inhalation) of in vivo studies needs to be complete and accurate to adequately assess and compare in vivo and in vitro datasets. Information on toxicokinetic and toxicodynamic variability, are needed to capture the typical range in response with both in vivo and in vitro. For example, if the endpoint of interest typically has a confidence interval of $\pm 5\%$ for oral but $\pm 20\%$ for dermal toxicity, the performance of an in vitro assay for in vivo extrapolation would differ depending on the route. Further, the acceptability of simplifying assumptions needs to be considered within the context of the particular approach being applied. For instance, when fraction of dose absorbed is assumed to be 100%, forward dosimetry models may greatly understate the potency of a chemical that causes toxicity. Finally, it is important to understand when certain predictive computational models may or may not incorporate underlying in vitro experimental data. For example, tools that assess membrane permeability typically rely on in vitro data due to difficulties in predicting these parameters in silico (e.g., active transport) (Bujard et al., 2015; Parrott and Lave, 2016).

Workshop attendees felt that a key hurdle to the acceptance of computational approaches by regulators is a lack of transparency. For external evaluation, a model's domain of applicability and limitations need to be adequately characterized, and any data or features used in

model development must be clearly described and readily available to the user. Model users need adequate training to ensure they understand underlying assumptions of the model and the impact that uncertainty surrounding a given parameter has on the overall model. User training along with transparency help ensure models are fit for their intended purpose.

5. Model evaluation and risk assessment applications

5.1. Current issues in construction and validation of TK models

Appropriate construction and validation of TK models (Clark et al., 2004; McLanahan et al., 2012) are essential to their use in regulatory applications. The assessment of a TK model for a particular application should include an evaluation of both parameter sensitivity and the extent to which the parameter represents the biological response of interest. Model complexity must reflect the needs of the risk decision. One-compartment models can be applied to quickly estimate chemical half-life in vivo, and are suitable for triaging chemicals into bins or ranking based on potential toxicities. These simple models are easy to understand and implement, and open-access versions of them are readily available (Chang et al., 2015; Pearce et al., 2017) which increases transparency and may help facilitate regulatory acceptance (Casey et al., 2015; Sullivan, 2016). If a chemical evaluated in a simple model appears to have a toxic dose close to the environmental exposure, it should then be further evaluated using a more complex model with additional structural parameters to refine dose estimates.

The European SEURAT-1 research initiative was established to develop pathway-based non-animal methods to predict human responses to repeated doses of chemicals. The goal of COSMOS (<http://www.cosmostox.eu/>; COSMOS, 2017), one of the projects within the SEURAT-1 initiative, was to develop computational tools to support these efforts. Tools developed under COSMOS include open-source PBPK models, ranging from a simple one-compartment model to predict bioaccumulation (Tonnelier et al., 2012) to more refined and chemical-specific PBPK models (Bois et al., 2017; Gajewska et al., 2014) and integrated biokinetic models to perform IVIVE (Gajewska et al., 2015), and a KNIME web-based platform (Sala Benito et al., 2017). These models were developed for specific chemicals, exposure routes, and species. Implementation of methods such as these for risk assessment purposes will require an evaluation of the necessary biological fidelity for a given context.

Factors to be considered during validation of TK models include how well the model describes the biological system, mode of action, chemical properties, and experimental design (Clark et al., 2004). To more accurately model a specific endpoint in an organ system, validation activities should be designed around a particular exposure scenario, such as a 90-day study, to ensure that the model is fit for purpose. Best practices for modeling the relevant in vivo and in vitro assays should be described, and the appropriate dose metric for each assay determined to define the relevant dose-effect relationship.

5.2. Approaches to evaluating models

Several key questions need to be considered when evaluating TK models for use in IVIVE for regulatory applications (McLanahan et al., 2012). These include model verifiability (adequacy of documentation, clearly stated model assumptions and parameters, and external reproducibility) and evaluation. Model evaluation depends on first establishing criteria appropriate for the model's application, and then using these criteria when considering the model's predictions relative to available data (Wambaugh et al., 2015). Developing fit-for-purpose IVIVE workflows depends upon careful consideration of parsimony, or the appropriateness of the model given the data, and domain of applicability, or the appropriateness of models given the chemicals used for developing and evaluating the model.

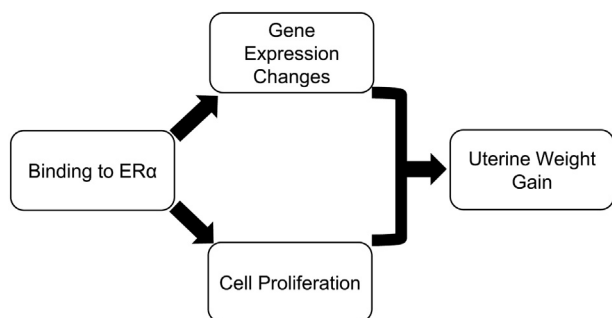


Fig. 5. Putative AOP for uterotrophy. Chemical binding to the estrogen receptor alpha (ERα) can lead to changes in gene expression and cell proliferation, both of which contribute to the adverse outcome of an increased uterine weight.

Even simple models can be challenging to evaluate. Users of IVIVE models must identify model biases and consider how these biases might impact the model's utility for generating predictions that will protect human health. Few published efforts that provide a systematic evaluation of TK modeling approaches currently exist (Wambaugh et al., 2015). Sensitivity analysis can be used to identify the parameters having the highest influence on the model and to focus model evaluation to those key parameters (Lumen et al., 2015; McNally et al., 2011). Evaluation of a model's domain of applicability requires comparing the model predictions for a set of chemicals to in vivo TK data for the same chemicals. Since many of the tools used for IVIVE analysis were originally developed for pharmaceuticals, their performance against the more diverse set of chemicals encountered in the environment needs to

be properly evaluated (Alves et al., 2017; Judson et al., 2008; Wambaugh et al., 2015).

A key question to consider is whether the data used by the model is appropriate for its proposed applications. Many current in vitro assays use human cells, but workshop participants agreed that corresponding human in vivo data needed to evaluate the in vitro assay results are often unavailable (Wambaugh et al., 2015; Yoon et al., 2014). On the in vivo side, there are more animal data than human data, so our understanding of animal toxicity is better than human toxicity. However, even in animal models, data gaps exist for both physiology and toxicology, which increase the uncertainty in IVIVE analysis. Some of these data gaps could be filled by incorporating TK endpoints (e.g., serial serum samples) into animal experiments intended to collect other (e.g., toxicological) data. In vitro methods that probe key events along the biological pathway leading to the adverse outcome of interest may help reduce uncertainty across species. Suggested methods to probe the biological event of interest from different angles include the use of orthogonal in vitro assays, that is, assays that use different approaches to measure the same biological event in the toxicity pathway.

If IVIVE is to be applied to risk-based prioritization, particular attention must be paid to uncertainty in model predictions and the ability to address biological variability, especially with respect to identifying sensitive populations (Barton et al., 2007; Hines et al., 2010). Sensitive populations require assessment of intra-individual variability in TK parameters (Wetmore et al., 2012, 2014, 2015). The level of confidence needed when evaluating a model for risk-based prioritization also depends upon the acceptable margins of exposure. If the IVIVE estimated dose range is close to or overlaps the range of potential hazard, a higher predictive accuracy is needed than if a wider margin exists between

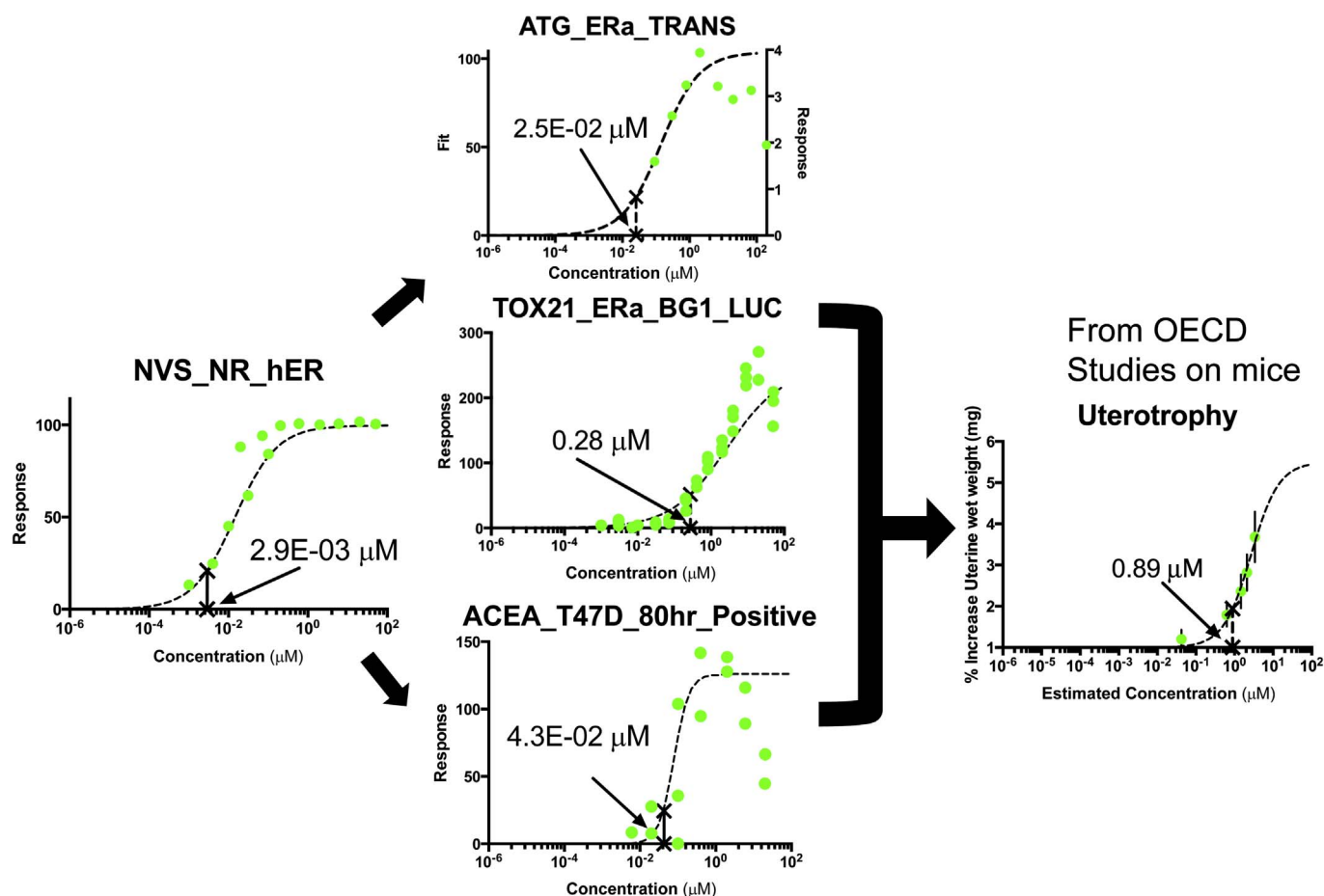


Fig. 6. The example uses the putative adverse outcome pathway for uterotrophy (Fig. 5) to illustrate how high throughput assays can be mapped to each key event of an adverse outcome pathway. NVS_NR_hER, cell free human estrogen receptor.

those ranges. If a conservative, provisional model indicates an extremely large margin between these ranges for a given chemical (such as eight orders of magnitude), this may indicate that the chemical needs little to no further evaluation compared with those with relatively smaller margin indicating there may be are higher priority targets (Wetmore et al., 2015).

The following cases studies drawn from the 2016 workshop illustrate how IVIVE may be evaluated and applied to risk assessment:

5.3. Case study: using in vitro data in quantitative risk assessment and adverse outcome pathway frameworks

Current risk assessment systems that have been codified in regulatory systems worldwide were developed with regard to the strengths and weaknesses of in vivo data. In those instances where there is existing knowledge of the limitations of assessments, uncertainty or safety factors are applied to account for any extrapolation issues and to ensure human health protection (National Research Council, 1983, 1994). As the use of in vitro and in silico methods increases to fill data gaps and inform data interpretation, methods used to assess data quality, uncertainty, and variability will need to be modified to ensure adequate consideration and use of these new data streams in decision-making.

Adverse outcome pathways (AOPs; Villeneuve et al., 2014a, 2014b) describe sets of key events that begin with an initial interaction of a chemical with the biological system through all the necessary molecular and biological steps, or key events, required to cause an adverse outcome. AOPs can be used to develop a testing strategy by identifying assays that measure the key events in the pathway leading to the adverse outcome. These key events often encompass different levels of biological organization within a living system, and the in vivo, in vitro, and in silico assays measuring each key event may report outcomes in different units of measure. Dosimetry models can help in resolving these differences, and work is ongoing to improve these models (Phillips et al., 2016; Teeguarden et al., 2016).

A 2014 workshop on AOPs highlighted the necessity for incorporating IVIVE in interpretation of AOP-based predictions of toxicity based on in vitro tests (Kleinstreuer et al., 2016b). An example of this was the application of IVIVE using a simple dosimetry model tied to an AOP framework to contextualize the results of an in vitro replacement for an in vivo assay for estrogenic activity. The uterotrophic assay is an in vivo rodent assay that assesses increase in uterine weight as a measure of potential estrogenic activity. A putative AOP for uterotrophy (Fig. 5) was developed, and in vitro assays from the ToxCast high throughput screening program were associated with key events in the AOP (Fig. 6). Using data from these assays as inputs, IVIVE analysis was applied to predict in vivo doses that would yield internal blood concentrations equivalent to the in vitro HTS assay activity concentrations. Results of the analysis were then compared to results of guideline-like uterotrophic studies from the literature. It was found that the range of equivalent doses from the in vitro assays provided good approximations of uterotrophic lowest effective levels for the majority of chemicals for both injection and oral administration routes (Kleinstreuer et al., 2016a).

Similarly, a set of human in vitro HTS assays measuring various endpoints in vasculogenesis or angiogenesis was used to develop an AOP-based predictive signature of developmental toxicity mediated by disruption of embryonic vascular development (Kleinstreuer et al., 2011; Knudsen and Kleinstreuer, 2011). Recent work has applied life-stage specific PBPK models to estimate maternal exposures yielding fetal blood concentrations equivalent to chemical concentrations inducing activity in these vascular disruption AOP-based assays (El-Masri et al., 2016). The resulting predictions were compared to realistic exposure estimates derived from biomonitoring data to derive AOP-based margins of exposure.

5.4. Case study: modeling population variability and sensitive populations

High throughput risk prioritization compares potential chemical hazard to potential exposure for a range of chemicals and ranks those based on the degree of overlap between the two observations (Ring et al., 2017; Thomas et al., 2013; Wetmore et al., 2015). While exposure can be estimated using high throughput model frameworks, such as the EPA's ExpoCast (Wambaugh et al., 2014), chemical hazard needs to be related to in vivo risk using TK models. In addition to modeling risk for a generic population, including variability within specific populations enables risk estimates for particularly susceptible groups (e.g., children).

IVIVE approaches (as illustrated in Fig. 2) developed by Wetmore and collaborators (Wetmore et al., 2012, 2015) simulated population distributions of TK parameters such as hepatic blood flow and body weight. These produced a range of predicted steady-state concentrations resulting from the same fixed exposure (e.g., 1 mg/kg body-weight/day). The upper 95th percentile of this range was then used to represent sensitive populations for which the same exposure would result in higher plasma concentrations (Wetmore et al., 2012, 2015). Recent efforts at the EPA have extended this approach to produce htk, an open-source correlated Monte Carlo tool for simulating the U.S. population (Pearce et al., 2017; Ring et al., 2017). The HTK-Pop module within htk samples physiological model parameters based on data from the U.S. Centers for Disease Control National Health and Nutrition Examination Survey (NHANES). These data provide estimates of variation within specific demographic groups for parameters such as tissue masses, tissue blood flows, kidney function, and hepatocellularity. These calculations of population variability, combined with HTS data and high throughput exposure inferences, allow users to conduct high throughput prioritization based on activity-exposure ratios for potentially sensitive subpopulations. Population-based PBPK modeling can be conducted in a similar manner with commercial programs such as Simcyp and GastroPlus.

In addition to variability in physiologic parameters, differences in xenobiotic metabolizing enzymes may contribute significantly to TK variability observed across different life stages and populations. Developmental differences in the abundances of more than 14 metabolic isozymes have been extensively characterized in newborns and children (Hines, 2007; Kearns et al., 2003). Genetic polymorphisms in metabolic isozymes are also well documented (Hiratsuka, 2012). A recent case study measured isozyme-specific clearance rates for 12 chemicals using recombinant-expressed isozymes and incorporated this information along with physiologic parameters to predict in vivo steady-state plasma concentrations across a range of life-stages and populations (Wetmore et al., 2014). In this study, use of pre-parameterized population libraries that captured differing isozyme abundances across the various life-stages and ethnic populations was instrumental in quantifying the anticipated TK variability across the different populations for each chemical, and provided clues as to what may be key metabolic drivers in TK variability.

Sensitive life stages, including fetuses, infants, and pregnant or nursing women, present unique challenges for TK modeling. For example, disturbances in the hypothalamus–pituitary–thyroid axis during pregnancy have been associated with development of adverse neuro-developmental effects (Lumen and George, 2017). One major cause of these disturbances is iodide deficiency, the effects of which may predispose individuals to additional thyroid homeostasis alterations after exposure to thyroid active chemicals (e.g., perchlorate) (Horton et al., 2015).

Scientists at the U.S. Food and Drug Administration developed a deterministic, biologically based dose-response model for the hypothalamus–pituitary–thyroid axis to evaluate the effects of varying iodide intake and perchlorate exposure on thyroid hormone levels in the near-term pregnant woman and fetus (Lumen et al., 2013). Global sensitivity analysis (Lumen et al., 2015) was used to evaluate the effects

of variability and uncertainty in the model input parameters on the predicted maternal free thyroxine levels. A population-based pregnancy model was developed to predict serum thyroid hormone level distribution given the current iodine nutritional status specific for late-gestation pregnant women in the U.S. (Lumen and George, 2017). The model predictions were in line with ranges of maternal thyroid hormone levels observed in several large U.S. and international biomarker studies. Availability of these data and the deterministic and population-based modeling framework allow researchers and risk assessors to ask relevant questions about the impact of exposure of pregnant women to thyroid-active chemicals. For example, the model can be used to address questions such as how much perchlorate exposure (from food and drinking water) is necessary to cause a significant alteration in the thyroid hormone levels for pregnant women, given the current iodine nutritional status of the population. This information plays an important role when making regulatory decisions in these data-sparse sensitive subpopulations. Such models and modeling approaches are increasingly being considered for regulatory applications (EPA, 2016b).

5.5. Case study: using computational tools to compress the product development timeline

The Dow Chemical Company evaluated GastroPlus in the context of developing risk assessments for a broad range of non-pharmaceutical chemicals and formulations. To support their existing TK assessment applications, Dow was seeking modeling software that would support multiple exposure routes and regimens; include both simple compartmental TK models and complex PBPK models; address multiple species and life stages; and require minimal or no coding.

Dow established a multistep validation plan to evaluate the accuracy of GastroPlus' predictions. Predicted values for acid dissociation constant from GastroPlus correlated well with literature data and improved upon predictions by another commercially available application, Pipeline Pilot. Likewise, octanol-water partition coefficient values predicted by GastroPlus correlated well with literature data and were similar to values predicted by an EPA-developed application, EPI Suite (EPA, 2016a). Clearance estimates were generally within an order of magnitude of empirical data, while f_u estimates were within 30% of empirical data (Wambaugh et al., 2015). Predicted TK values also correlated well with literature values; predictions of two key values, maximum plasma concentration and area under the plasma concentration vs. time curve were within an order of magnitude for much of the available in vivo data. Steady-state blood concentration predictions from GastroPlus were consistent with those obtained using Simcyp and were generally conservative in comparison to reference data.

Ongoing investigations at Dow will use a test set of about 60 chemicals to evaluate an in-house exposure model to predict systemic blood concentrations for oral, inhalation, and dermal exposure routes. Formulation types and exposure scenarios will be chosen to provide conservative blood concentration predictions. Goals of future research include refining model predictions with empirical intrinsic clearance and f_u values, and deriving correlations for pulmonary clearance of unmetabolized volatile compounds.

5.6. Case study: ecological applications

A conservative estimate is that there are over 3000 vertebrate wildlife species in the United States, including approximately 450 species listed as endangered (U.S. Fish and Wildlife Service, 2015). Because toxicity testing for every chemical on every wildlife species is not feasible, EPA utilizes harmonized test guidelines that rely on the use of surrogate species to test for or predict adverse effects (EPA, 2016c). Thus, only a relatively few species have been commonly used to represent broad taxonomic groups in toxicity testing. IVIVE can potentially be used to provide estimates of chemical toxicity to a large range of wildlife species for which traditional toxicity testing is not feasible.

However, a further limitation is imposed by the fact that most HTS in vitro assays are based on human or rat cell lines or proteins (Kavlock et al., 2012; LaLone et al., 2016). An AOP framework was used to demonstrate that pathway-based analysis can allow data from human-focused in vitro assays to predict outcomes in non-mammalian species (Perkins et al., 2013). Also, a cross-species extrapolation approach can be used to investigate the structural and functional conservation of critical pathway components to help determine biological susceptibility (Ankley et al., 2016; LaLone et al., 2016). IVIVE coupled with species extrapolation of mammalian in vitro test data allows comparison of relative species internal doses, which will enable prioritization of data gaps for wildlife testing and assessment, potentially providing greater confidence in ecological risk assessments for chemical safety.

Fish species provide examples of challenges involved in developing PBPK models for IVIVE in ecological applications. PBPK models have been developed for a number of fish species over the past two decades: rainbow trout (Nichols et al., 1990), lake trout (Lien et al., 2001), fathead minnow (Lien et al., 1994), crucian carp (F. Yang et al., 2013), zebrafish (Péry et al., 2014) and medaka (Parhizgari and Li, 2014). Toxicity testing in fish has primarily relied on three of these species (fathead minnow, medaka, and zebrafish) because their small size facilitates in vivo testing (Ankley and Johnson, 2004). However, the size of their organs (about 75 mg for the liver and 20–60 μ l of plasma for the fathead minnow) make it difficult to obtain the large quantities of biological material needed to measure hepatic clearance and plasma protein binding, parameters needed for IVIVE models (Fig. 2; Jensen et al., 2001; Watanabe et al., 2007). A study that compared evaluation of estrogenic potency of environmental chemicals using both in vivo and in vitro zebrafish assay found that, without consideration of TK or employing PBPK models, the in vitro assays poorly estimated in vivo data and tended to overestimate potency of estrogenic activity (Segner et al., 2003). Species such as rainbow trout would be more useful for ecological studies using IVIVE. Rainbow trout are biologically good surrogates for other economically important sport fish, and their relatively large body and thus organ size would facilitate measurement of input parameters for IVIVE models. However, few traditional ecotoxicity tests have been done in this species, so little data are available. Bridging this data and species gap is important in future IVIVE applications for aquatic toxicology, especially as there are challenges associated with applying rainbow trout plasma binding and hepatic clearance empirical data to a different fish species (Strobel et al., 2015).

6. Promoting IVIVE for specific regulatory applications

6.1. Challenges facing regulatory application of IVIVE

Gaining regulatory acceptance is key to the success of any alternative approach to animal testing. Potential regulatory applications for IVIVE include prioritization, screening, and risk assessment. Relevant stakeholders for these applications include regulatory agencies, industry, nongovernmental organizations, and consumers. However, regulatory agencies that have testing requirements for hazard identification and assessment are central to any discussion of how best to implement new test methods and IVIVE approaches.

Regulators are faced with difficult questions when considering the predictive performance of a model. They may need to establish confidence criteria and define the applicability domains that are appropriate for the decision-making need (e.g., prioritization, screening, classification and labeling). While many agencies have similar mandates to protect public and environmental health, specific requirements differ globally and among and even within agencies in the U.S. This lack of harmonization presents practical challenges for companies who may need to generate both animal and non-animal data for the same endpoint to satisfy the various regulatory bodies. Ultimately, until all regulatory agencies reject animal tests, use of these tests is likely to continue to ensure that all global requirements are fulfilled.

A frequently overlooked challenge is the potential for legal implications associated with data generated using an alternative approach. In some cases, stakeholders will hesitate to use non-animal data in regulatory applications due to litigation concerns. The Daubert standard (National Research Council, 2011) provides criteria for determining if scientific testimony is based on reasoning or methodology that is scientifically valid and can be applied properly. These criteria include peer review, the presence of standards and controls along with associated error rates, and widespread acceptance within the scientific community. Regulatory agencies develop processes to assure that the standards they use for required data comply with these criteria. For example, EPA has multiple comprehensive peer review processes to provide this information, including an external Board of Scientific Counselors. IVIVE methods developed for regulatory uses will need to be well-documented and transparent. Only then will users have the confidence to use such method to inform a regulatory decision that is legally defensible.

Before they make formal submissions to regulatory agencies, companies routinely make product development decisions based on non-animal screening results. However, these same companies may hesitate to include non-animal data in formal regulatory submissions because of uncertainty of how the data will be interpreted or accepted by regulators. Consensus in the use of predictive in vitro systems for regulatory applications can be achieved through increasing the number of examples where high-quality data are applied to practical regulatory testing. One idea suggested by workshop participants was the concept of a “safe harbor” within regulatory agencies whereby sponsors could submit non-animal data for evaluation without fear of its use in a contradictory manner. The biomarker qualification process at the U.S. Food and Drug Administration serves as a precedent for this effort (U.S. Food and Drug Administration, 2017). Broad participation in such a program will require both transparency in how these data will be used and revisions or clarifications to legislation that prohibit such uses (such as EPA regulatory requirements under the statutory authority of the Federal Insecticide, Fungicide, and Rodenticide Act and the Toxic Substance Control Act to report any indication of toxicity). Lack of global harmonization also applies in this context. For example, with pesticides and biocides the EU system is hazard-based (European Union, 2009, 2012), while the U.S. system is risk-based (EPA, 2017c). Therefore, companies may be hesitant to release data that suggest a hazard due to the implications that suggestion of hazard may have in the EU regulatory context, regardless of whether the tested concentrations are relevant to human exposure scenarios.

6.2. Opportunities for regulatory applications of IVIVE

Despite these barriers, participants at the 2016 workshop suggested that IVIVE could be applied in the near term to several areas of regulatory use. One application is initial screening for toxicity. EPA has proposed accepting data from in vitro HTS assays and an associated computational model for detecting estrogen receptor bioactivity as an alternative for three current EPA Endocrine Disruptor Screening Program Tier 1 estrogen agonist screening assays, including the rodent uterotrophic test (EPA, 2017a). In this approach, results from the HTS data are contextualized using IVIVE to calculate an integrated bioactivity exposure relationship for each chemical, constituting perhaps the most advanced example of current regulatory use (Chang et al., 2015; EPA, 2015). Another potential opportunity for immediate application of IVIVE could be replacement of the use of the maximum tolerated dose to determine the appropriate dose range and spacing for traditional toxicology studies. Using in vitro test results in this way could improve the quality and relevance of data from animal tests, reduce animal use, and improve animal welfare by avoiding highly toxic doses.

In vitro testing approaches incorporating IVIVE could also be applied to address the lack of sufficient chemical hazard data in many sectors, providing at least approximations of toxicity for regulators and

public health workers who need to rapidly make decisions in situations involving chemicals for which there are currently little or no data. IVIVE allows for better risk prediction and communication by putting effects in the context of actual exposure, so that the “worst case” scenario can be identified and translated into testable concentrations.

Another potential use of IVIVE in the regulatory setting is in developing data-driven uncertainty factors. Currently, uncertainty factors are calculated to account for missing information, inter- and intra-species considerations, calculated or reported lowest or no-observed-effect levels, and extrapolating from subchronic to chronic effects. It is critical to determine both where the uncertainty factors are best applied as well as what the uncertainty factors are intended to address. For example, data-driven uncertainty factors might be determined by applying high throughput exposure model predictions (e.g., ExpoCast, (EPA, 2017b; Ring et al., 2017; Wetmore et al., 2015)).

In the long term, IVIVE will likely become a component of a scientific confidence framework for in vitro assays and used routinely to put results into proper exposure context. IVIVE is necessary for dose-response assessment of in vitro data for risk assessment applications, and thus, it should become part of the toxicologist's and risk assessor's toolbox. IVIVE is also expected to become an integral component of the AOP framework as the biological pathway altering dose approach (Judson et al., 2011; Patlewicz et al., 2015) becomes more utilized and chemical-specific exposure information is generated to inform quantitative AOPs.

6.3. Increasing buy-in

While addressing the needs of government regulators is important, the needs of other stakeholders—industry, nongovernmental organizations, and consumers—should also be considered. As in regulatory applications, transparent communication with these stakeholders will be the key to establishing a sufficient level of trust in testing results based on alternative methods. IVIVE offers the promise of predicting effects based on in vitro results and projecting them to human exposures arising from specific scenarios, such as a chemical spill, to establish whether there is cause for concern. For communications to the public, these results should be contextualized such that nonscientists can understand the implications of the testing results. Journal editors, reviewers, and science journalists should be appropriately educated to ensure that publications and communications to the popular press have proper biological context to minimize the spread of misinformation.

With an ever-increasing number of non-animal assays and approaches being developed, a process is needed to establish how best to evaluate and then promote use of the most promising IVIVE approaches. Ideally, new methods will first be evaluated through fit-for-purpose validation processes focusing on specific regulatory needs, and then used as the basis for development of performance-based test guidelines (OECD, 2016) to address those regulatory needs. These test guidelines can then be used as the basis for evaluation of modifications or improvements of these methods.

The academic community is at the cutting edge of test method development, but to be most useful, methods under development need to be properly aligned with regulatory testing needs. To address this concern, avenues should be explored to facilitate interaction between the academic and regulatory communities, who could collaborate in developing guidance to support this alignment. Appropriate funding mechanisms should support new research and development efforts and provide for validation studies. Funds for such validation studies are available through National Institutes of Health Small Business Innovation and Research Phase 2B grants that provide funding specifically for the validation of mature test methods (NIEHS, 2017).

7. Conclusions and recommendations

A consensus opinion from the workshop was that there is a growing

momentum for the use of IVIVE for regulatory decision-making. Workshop participants identified the following specific activities and resources that could support this use:

- Guidance for model development to help users determine the complexity of the model needed for a specific application
- Guidance for validation of IVIVE analyses and TK models, including consideration of differences between human and animal physiology
- Better characterization of the differences in properties between pharmaceuticals and non-pharmaceutical chemicals, and associated applicability domains of models developed for each
- Definition of effective practices for collection of TK model input data, recording of data on model predictions, and reporting of data and model code
- Creation of a database that could house all shared in vitro and in vivo TK data, and identification of actions to be taken to encourage sharing of existing data
- International harmonization of data requirements by regulators

Workshop participants agreed that IVIVE may currently be used to:

- Screen data-poor chemicals to rapidly provide information on potential toxicity
- Improve dose selection and spacing in animal studies
- Developing data-driven uncertainty factors
- Support development of aggregate exposure pathway and adverse outcome pathway based integrated testing strategies using in silico and in vitro methods

Transparency document

The [Transparency document](#) associated with this article can be found, in online version.

Acknowledgements

We would like to thank the EPA for hosting the workshop, EPA and ILS staff for providing logistical support, and conference participants for their contributions to discussion during the workshop.

This project was funded in part with federal funds from the National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH) under Contract No. HHSN273201500010C to ILS in support of NICEATM.

Kim L.R. Brouwer receives support from the National Institute of General Medical Sciences through award numbers R01GM041935 and R35GM122576, and from industry (Otsuka Pharmaceutical Development and Commercialization, Inc., and Intercept Pharmaceuticals). K.L.R.B. is a coinventor of the sandwich-cultured hepatocyte technology for quantification of biliary excretion (B-CLEAR) and related technologies, which have been licensed exclusively to Qualyst Transporter Solutions, LLC.

Disclaimers

This manuscript has been reviewed in accordance with the policy of the Office of Research and Development, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views or policy of the Agency nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

This manuscript does not necessarily reflect the views of the U.S. Food and Drug Administration.

This manuscript may be the work product of an employee or group of employees of NIEHS, NIH, or other organizations. However, the statements, opinions or conclusions contained therein do not necessarily represent the statements, opinions, or conclusions of NIEHS, NIH,

U.S. government, or other organizations. ILS staff do not represent NIEHS, the National Toxicology Program, or the official positions of any federal agency.

References

- Alves, V.M., Muratov, E.N., Zakharov, A., Muratov, N.N., Andrade, C.H., Tropsha, A., 2017. Chemical toxicity prediction for major classes of industrial chemicals: is it possible to develop universal models covering cosmetics, drugs, and pesticides? *Food Chem. Toxicol.* <http://dx.doi.org/10.1016/j.fct.2017.04.008>. (Epub ahead of print).
- Ankley, G.T., Johnson, R.D., 2004. Small fish models for identifying and assessing the effects of endocrine-disrupting chemicals. *ILAR J.* 45, 469–483. <http://dx.doi.org/10.1093/ilar.45.4.469>.
- Ankley, G.T., LaLone, C.A., Gray, L.E., Villeneuve, D.L., Hornung, M.W., 2016. Evaluation of the scientific underpinnings for identifying estrogenic chemicals in nonmammalian taxa using mammalian test systems. *Environ. Toxicol. Chem.* 35, 2806–2816. <http://dx.doi.org/10.1002/etc.3456>.
- Armitage, J.M., Wania, F., Arnot, J.A., 2014. Application of mass balance models and the chemical activity concept to facilitate the use of in vitro toxicity data for risk assessment. *Environ. Sci. Technol.* 48, 9770–9779.
- Barton, H.A., Chiu, W.A., Woodrow, S.R., Andersen, M.E., Bailer, A.J., Bois, F.Y., DeWoskin, R.S., Hays, S., Johanson, G., Jones, N., Loizou, G., MacPhail, R.C., Portier, C.J., Spendiff, M., Tan, Y.-M., 2007. Characterizing uncertainty and variability in physiologically based pharmacokinetic models: state of the science and needs for research and implementation. *Toxicol. Sci.* 99, 395–402. <http://dx.doi.org/10.1093/toxsci/kfm100>.
- Bell, C.C., Hendriks, D.F.G., Moro, S.M.L., Ellis, E., Walsh, J., Renblom, A., Fredriksson Puigvert, L., Dankers, A.C.A., Jacobs, F., Snoeys, J., Sison-Young, R.L., Jenkins, R.E., Nordling, Å., Mkrtchian, S., Park, B.K., Kitteringham, N.R., Goldring, C.E.P., Lauschke, V.M., Ingelman-Sundberg, M., 2016. Characterization of primary human hepatocyte spheroids as a model system for drug-induced liver injury, liver function and disease. *Sci. Rep.* 6, 25187. <http://dx.doi.org/10.1038/srep25187>.
- Bessemis, J., Coecke, S., Gouliarmou, V., Whelan, M., Worth, A., 2015. EURL ECVAM Strategy for Achieving 3Rs Impact in the Assessment of Toxicokinetics and Systemic Toxicity. Publications Office of the European Union, Luxembourg.
- Blauboer, B.J., 2010. Biokinetic modeling and in vitro-in vivo extrapolations. *J. Toxicol. Environ. Health* 13, 242–252.
- Bois, Y.F., Diaz, J.G., Gjewaska, M., Kovarich, S., Mauch, K., Paini, A., Pery, A., Sala Benito, J.V., Teng, S., Worth, A., 2017. Multiscale modelling approaches for assessing cosmetic ingredients safety. *Toxicology (0300-483X)* 392, 130–139. <http://dx.doi.org/10.1016/j.tox.2016.05.026>.
- Bow, D.A.J., Perry, J.L., Miller, D.S., Pritchard, J.B., Brouwer, K.L.R., 2008. Localization of P-gp (Abcb1) and Mrp2 (Abcc2) in freshly isolated rat hepatocytes. *Drug Metab. Dispos.* 36, 198–202. <http://dx.doi.org/10.1124/dmd.107.018200>.
- Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J., Vingron, M., 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.* 29, 365–371. <http://dx.doi.org/10.1038/ng1201-365>.
- Brouwer, K.L.R., Keppler, D., Hoffmaster, K.A., Bow, D.A., Cheng, Y., Lai, Y., Palm, J.E., Stieger, B., Evers, R., International Transporter Consortium, 2013. In vitro methods to support transporter evaluation in drug discovery and development. *Clin. Pharmacol. Ther.* 94, 95–112. <http://dx.doi.org/10.1038/clpt.2013.81>.
- Bujard, A., Voirol, H., Carrupt, P.-A., Schappler, J., 2015. Modification of a PAMPA model to predict passive gastrointestinal absorption and plasma protein binding. *Eur. J. Pharm. Sci.* 77, 273–278. <http://dx.doi.org/10.1016/j.ejps.2015.06.023>.
- Caldwell, J.C., Evans, M.V., Krishnan, K., 2012. Cutting edge PBPK models and analyses: providing the basis for future modeling efforts and bridges to emerging toxicology paradigms. *J. Toxicol.* 2012, 1–10. <http://dx.doi.org/10.1155/2012/852384>.
- Campbell, J.L., Clewell, R.A., Gentry, P.R., Andersen, M.E., Clewell, H.J., 2012. Physiologically based pharmacokinetic/toxicokinetic modeling. *Methods Mol. Biol.* 929, 439–499. http://dx.doi.org/10.1007/978-1-62703-050-2_18.
- Casey, W., Jacobs, A., Maull, E., Matheson, J., Clarke, C., Lowit, A., 2015. A new path forward: the interagency coordinating committee on the validation of alternative methods (ICCVAM) and National Toxicology Program's interagency Center for the Evaluation of alternative toxicological methods (NICEATM). *J. Am. Assoc. Lab. Anim. Sci.* 54, 170–173.
- Chang, X., Kleinstreuer, N., Ceger, P., Hsieh, J.-H., Allen, D., Casey, W., 2015. Application of reverse dosimetry to compare in vitro and in vivo estrogen receptor activity. *Appl. Vitro Toxicol.* 1, 33–44. <http://dx.doi.org/10.1089/avto.2014.0005>.
- Chiba, M., Ishii, Y., Sugiyama, Y., 2009. Prediction of hepatic clearance in human from in vitro data for successful drug development. *AAPS J.* 11. <http://dx.doi.org/10.1208/s12248-009-9103-6>.
- Chu, X., Korzekwa, K., Elsby, R., Fenner, K., Galetin, A., Lai, Y., Matsson, P., Moss, A., Nagar, S., Rosania, G.R., Bai, J.P.F., Polli, J.W., Sugiyama, Y., Brouwer, K.L.R., International Transporter Consortium, 2013. Intracellular drug concentrations and transporters: measurement, modeling, and implications for the liver. *Clin. Pharmacol. Ther.* 94, 126–141. <http://dx.doi.org/10.1038/clpt.2013.78>.
- Clark, L.H., Setzer, R.W., Barton, H.A., 2004. Framework for evaluation of physiologically-based pharmacokinetic models for use in safety or risk assessment. *Risk Anal.* 24, 1697–1717. <http://dx.doi.org/10.1111/j.0272-4332.2004.00561.x>.
- Coecke, S., Pelkonen, O., Leite, S.B., Bernauer, U., Bessemis, J.G., Bois, F.Y., Gundert-Remy, U., Loizou, G., Testai, E., Zaldivar, J.-M., 2013. Toxicokinetics as a key to the

- integrated toxicity risk assessment based primarily on non-animal approaches. *Toxicol. in Vitro* 27, 1570–1577. <http://dx.doi.org/10.1016/j.tiv.2012.06.012>.
- COSMOS, 2017. COSMOS KNIME Webportal. <https://knimewebportal.cosmostox.eu/>, Accessed date: 20 July 2017.
- Dancik, Y., Troutman, J.A., Jaworska, J., 2013. A framework incorporating the impact of exposure scenarios and application conditions on risk assessment of chemicals applied to skin. *In Silico Pharmacol.* 1, 10. <http://dx.doi.org/10.1186/2193-9616-1-10>.
- Dancik, Y., Troutman, J.A., Jaworska, J., 2015. Estimation of in vivo dose of dermally applied chemicals leading to estrogen/androgen receptor-mediated toxicity from in vitro data—illustration with four reproductive toxicants. *Reprod. Toxicol.* 55, 50–63. <http://dx.doi.org/10.1016/j.reprotox.2015.01.002>.
- Dansirikul, C., Choi, M., Duffull, S.B., 2005. Estimation of pharmacokinetic parameters from non-compartmental variables using Microsoft Excel. *Comput. Biol. Med.* 35, 389–403. <http://dx.doi.org/10.1016/j.combiomed.2004.02.008>.
- El-Masri, H., Kleinstreuer, N., Hines, R.N., Adams, L., Tal, T., Isaacs, K., Wetmore, B.A., Tan, Y.-M., 2016. Integration of life-stage physiologically based pharmacokinetic models with adverse outcome pathways and environmental exposure models to screen for environmental hazards. *Toxicol. Sci.* 152, 230–243. <http://dx.doi.org/10.1093/toxsci/kfw082>.
- EPA, 2015. Use of high throughput assays and computational tools; endocrine disruptor screening program; notice of availability and opportunity for comment. *Fed. Regist.* 80, 35350–35355. <http://dx.doi.org/10.1016/j.fedreg.2015.07.008>.
- EPA, 2016a. EPI Suite™ Estimation Programs Interface. <https://www.epa.gov/tsca-screening-tools/epi-suite-estimation-program-interface>, Accessed date: 20 July 2017.
- EPA, 2016b. Request for public comments to be sent to EPA on peer review materials to inform the safe drinking water act decision making on perchlorate. *Fed. Regist.* 81, 67350–67351.
- EPA, 2016c. Final Test Guidelines for Pesticides and Toxic Substances. <https://www.epa.gov/test-guidelines-pesticides-and-toxic-substances/final-test-guidelines-pesticides-and-toxic>, Accessed date: 20 July 2017.
- EPA, 2017a. Use of High Throughput Assays and Computational Tools in the Endocrine Disruptor Screening Program. <https://www.epa.gov/endocrine-disruption/use-high-throughput-assays-and-computational-tools-endocrine-disruptor>, Accessed date: 20 July 2017.
- EPA, 2017b. Rapid Chemical Exposure and Dose Research. <https://www.epa.gov/chemical-research/rapid-chemical-exposure-and-dose-research>, Accessed date: 20 July 2017.
- EPA, 2017c. Summary of the Federal Insecticide, Fungicide, and Rodenticide Act. <https://www.epa.gov/laws-regulations/summary-federal-insecticide-fungicide-and-rodenticide-act>, Accessed date: 26 July 2017.
- European Union, 2009. Regulation (EC) No. 1107/2009 of the European parliament and of the council concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. *Off. J. Eur. Union* 309, 1–50 11 Nov 2009.
- European Union, 2012. Regulation (EU) No 528/2012 of the European parliament and of the council of 22 May 2012 concerning the making available on the market and use of biocidal products (Text with EEA relevance). *Off. J. Eur. Union* 167, 1–123 (27 June 2012).
- Ferguson, S.S., Bonzo, J.A., 2014. In vitro approaches to study drug–drug interactions. In: Kirchner, J. (Ed.), *Drug Metabolism Prediction*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, pp. 441–484.
- Filer, D.L., Kothiyar, P., Setzer, R.W., Judson, R.S., Martin, M.T., 2017. tcpl: The ToxCast pipeline for high-throughput screening data. *Bioinformatics* 33, 618–620. <http://dx.doi.org/10.1093/bioinformatics/btw680>.
- Fischer, F.C., Henneberger, L., König, M., Bittermann, K., Linden, L., Goss, K.U., Escher, B.I., 2017. Modeling exposure in the Tox21 in vitro bioassays. *Chem. Res. Toxicol.* 30, 1197–1208.
- Fouchécourt, M.O., Bélièveau, M., Krishnan, K., 2001. Quantitative structure-pharmacokinetic relationship modelling. *Sci. Total Environ.* 274, 125–135.
- Frazier, J.M., 1995. Interdisciplinary approach to toxicity test development and validation. *Toxicol. in Vitro* 9, 845–849. <http://dx.doi.org/10.1021/acs.chemrestox.7b00023>.
- Fujiwara, S.-I., Yamashita, F., Hashida, M., 2003. QSAR analysis of interstudy variable skin permeability based on the “latent membrane permeability” concept. *J. Pharm. Sci.* 92, 1939–1946. <http://dx.doi.org/10.1002/jps.10462>.
- Gajewska, M., Worth, A., Urani, C., Briesen, H., Schramm, K.-W., 2014. Application of physiologically-based toxicokinetic modelling in oral-to-dermal extrapolation of threshold doses of cosmetic ingredients. *Toxicol. Lett.* 227, 189–202. <http://dx.doi.org/10.1016/j.toxlet.2014.03.013>.
- Gajewska, M., Paini, A., Sala Benito, J.V., Burton, J., Worth, A., Urani, C., Briesen, H., Schramm, K.-W., 2015. In vitro-to-in vivo correlation of the skin penetration, liver clearance and hepatotoxicity of caffeine. *Food Chem. Toxicol.* 75, 39–49. <http://dx.doi.org/10.1016/j.fct.2014.10.017>.
- Ghibellini, G., Vassist, L.S., Leslie, E.M., Heizer, W.D., Kowalsky, R.J., Calvo, B.F., Brouwer, K.L.R., 2007. In vitro-in vivo correlation of hepatobiliary drug clearance in humans. *Clin. Pharmacol. Ther.* 81, 406–413. <http://dx.doi.org/10.1038/sj.clpt.6100059>.
- Groothuis, F.A., Heringa, M.B., Nicol, B., Hermens, J.L., Blaauboer, B.J., Kramer, N.I., 2015. Dose metric considerations in in vitro assays to improve quantitative in vitro-in vivo dose extrapolations. *Toxicology* 332, 30–40. <http://dx.doi.org/10.1016/j.tox.2013.08.012>.
- Haddad, S., Poulin, P., Krishnan, K., 2000. Relative lipid content as the sole mechanistic determinant of the adipose tissue: blood partition coefficients of highly lipophilic organic chemicals. *Chemosphere* 40, 839–843.
- Hallifax, D., Houston, J.B., 2012. Evaluation of hepatic clearance prediction using in vitro data: emphasis on fraction unbound in plasma and drug ionisation using a database of 107 drugs. *J. Pharm. Sci.* 101, 2645–2652. <http://dx.doi.org/10.1002/jps.23202>.
- Harper, T.W., Brassil, P.J., 2008. Reaction phenotyping: current industry efforts to identify enzymes responsible for metabolizing drug candidates. *AAPS J.* 10, 200–207. <http://dx.doi.org/10.1208/s12248-008-9019-6>.
- Hines, R.N., 2007. Ontogeny of human hepatic cytochromes P450. *J. Biochem. Mol. Toxicol.* 21, 169–175.
- Hines, R.N., Sargent, D., Autrup, H., Birnbaum, L.S., Brent, R.L., Doerr, N.G., Cohen Hubal, E.A., Juberg, D.R., Laurent, C., Luebke, R., Olejniczak, K., Portier, C.J., Slikker, W., 2010. Approaches for assessing risks to sensitive populations: lessons learned from evaluating risks in the pediatric population. *Toxicol. Sci.* 113, 4–26. <http://dx.doi.org/10.1093/toxsci/kfp217>.
- Hiratsuka, M., 2012. In vitro assessment of the allelic variants of cytochrome P450. *Drug Metab. Pharmacokinet.* 27, 68–84.
- Horton, M.K., Blout, B.C., Valentin-Blasini, L., Wapner, R., Whyatt, R., Gennings, C., Factor-Litvak, P., 2015. Co-occurring exposure to perchlorate, nitrate and thiocyanate alters thyroid function in healthy pregnant women. *Environ. Res.* 143, 1–9.
- Jackson, J.P., Li, L., Chamberlain, E.D., Wang, H., Ferguson, S.S., 2016. Contextualizing hepatocyte functionality of cryopreserved HepaRG cell cultures. *Drug Metab. Dispos.* 44, 1463–1479. <http://dx.doi.org/10.1124/dmd.116.069831>.
- Jamei, M., Marciniak, S., Feng, K., Barnett, A., Tucker, G., Rostami-Hodjegan, A., 2009. The Simcyp® population-based ADME simulator. *Expert Opin. Drug Metab. Toxicol.* 5, 211–223. <http://dx.doi.org/10.1517/17425250802691074>.
- Jensen, K.M., Korte, J.J., Kahl, M.D., Pasha, M.S., Ankley, G.T., 2001. Aspects of basic reproductive biology and endocrinology in the fathead minnow (*Pimephales promelas*). *Comp. Biochem. Physiol. C: Toxicol. Pharmacol.* 128, 127–141.
- Judson, R., Richard, A., Dix, D., Houck, K., Elloumi, F., Martin, M., Cathey, T., Transue, T.R., Spencer, R., Wolf, M., 2008. ACToR—aggregated computational toxicology resource. *Toxicol. Appl. Pharmacol.* 233, 7–13. <http://dx.doi.org/10.1016/j.taap.2007.12.037>.
- Judson, R.S., Martin, M.T., Reif, D.M., Houck, K.A., Knudsen, T.B., Rotroff, D.M., Xia, M., Sakamuru, S., Huang, R., Shinn, P., Austin, C.P., Kavlock, R.J., Dix, D.J., 2010. Analysis of eight oil spill dispersants using rapid, in vitro tests for endocrine and other biological activity. *Environ. Sci. Technol.* 44, 5979–5985. <http://dx.doi.org/10.1021/es102150z>.
- Judson, R.S., Kavlock, R.J., Setzer, R.W., Hubal, E.A.C., Martin, M.T., Knudsen, T.B., Houck, K.A., Thomas, R.S., Wetmore, B.A., Dix, D.J., 2011. Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. *Chem. Res. Toxicol.* 24, 451–462. <http://dx.doi.org/10.1021/tx100428e>.
- Kavlock, R., Chandler, K., Houck, K., Hunter, S., Judson, R., Kleinstreuer, N., Knudsen, T., Martin, M., Padilla, S., Reif, D., Richard, A., Rotroff, D., Sipes, N., Dix, D., 2012. Update on EPA's ToxCast program: providing high throughput decision support tools for chemical risk management. *Chem. Res. Toxicol.* 25, 1287–1302. <http://dx.doi.org/10.1021/tx3000939>.
- Kearns, G.L., Abdel-Rahman, S.M., Alander, S.W., Blowey, D.L., Leeder, J.S., Kauffman, R.E., 2003. Developmental pharmacology—drug disposition, action, and therapy in infants and children. *N. Engl. J. Med.* 349, 1157–1167. <http://dx.doi.org/10.1056/NEJMr030592>.
- Kleinstreuer, N.C., Judson, R.S., Reif, D.M., Sipes, N.S., Singh, A.V., Chandler, K.J., Dewoskin, R., Dix, D.J., Kavlock, R.J., Knudsen, T.B., 2011. Environmental impact on vascular development predicted by high-throughput screening. *Environ. Health Perspect.* 119, 1596–1603. <http://dx.doi.org/10.1289/ehp.1103412>.
- Kleinstreuer, N.C., Ceger, P.C., Allen, D.G., Strickland, J., Chang, X., Hamm, J.T., Casey, W.M., 2016a. A curated database of rodent uterotrophic bioactivity. *Environ. Health Perspect.* 124, 556–562. <http://dx.doi.org/10.1289/ehp.1510183>.
- Kleinstreuer, N.C., Sullivan, K., Allen, D., Edwards, S., Mendrick, D.L., Embry, M., Matheson, J., Rowlands, J.C., Munn, S., Maull, E., Casey, W., 2016b. Adverse outcome pathways: from research to regulation scientific workshop report. *Regul. Toxicol. Pharmacol.* 76, 39–50. <http://dx.doi.org/10.1016/j.yrtph.2016.01.007>.
- Knudsen, T.B., Kleinstreuer, N.C., 2011. Disruption of embryonic vascular development in predictive toxicology. *Birth Defects Res. C Embryo Today* 93, 312–323. <http://dx.doi.org/10.1002/bdrc.20223>.
- Kramer, N.I., Di Consiglio, E., Blaauboer, B.J., Testai, E., 2015. Biokinetics in repeated-dosing in vitro drug toxicity studies. *Toxicol. in Vitro* 30, 217–224. <http://dx.doi.org/10.1016/j.tiv.2015.09.005>.
- LaLone, C.A., Villeneuve, D.L., Lyons, D., Helgen, H.W., Robinson, S.L., Swintek, J.A., Saari, T.W., Ankley, G.T., 2016. Editor's highlight: sequence alignment to predict across species susceptibility (SeqAPASS): a web-based tool for addressing the challenges of cross-species extrapolation of chemical toxicity. *Toxicol. Sci.* 153, 228–245. <http://dx.doi.org/10.1093/toxsci/kfw119>.
- Lawless, M., DiBella, J., Bolger, M.B., Clark, R.D., Huehn, E., Waldman, M., Zhang, J., Lukacova, V., 2015. Prediction of oral bioavailability in silico. In: *ISSX Annual Meeting*.
- Lee, S.-A., No, D.-Y., Kang, E., Ju, J., Kim, D.-S., Lee, S.-H., 2013. Spheroid-based three-dimensional liver-on-a-chip to investigate hepatocyte-hepatic stellate cell interactions and flow effects. *Lab Chip* 13, 3529–3537. <http://dx.doi.org/10.1039/c3lc50197c>.
- Leeson, L.J., 1995. In vitro/in vivo correlations. *Drug Inf. J.* 29, 903–915. <http://dx.doi.org/10.1177/009286159502900312>.
- Li, N., Singh, P., Mandrell, K.M., Lai, Y., 2010. Improved extrapolation of hepatobiliary clearance from in vitro sandwich cultured rat hepatocytes through absolute quantification of hepatobiliary transporters. *Mol. Pharm.* 7, 630–641. <http://dx.doi.org/10.1021/mp9001574>.
- Lien, G.J., Nichols, J.W., McKim, J.M., Galliani, C.A., 1994. Modeling the accumulation of three waterborne chlorinated ethanes in fathead minnows (*Pimephales promelas*): a physiologically based approach. *Environ. Toxicol. Chem.* 13, 1195–1205. <http://dx.doi.org/10.1002/etc.1195>.

- doi.org/10.1002/etc.5620130721.
- Lien, G.J., McKim, J.M., Hoffman, A.D., Jenson, C.T., 2001. A physiologically based toxicokinetic model for lake trout (*Salvelinus namaycush*). *Aquat. Toxicol.* 51, 335–350.
- Lipscomb, J.C., Fisher, J.W., Confer, P.D., Byczkowski, J.Z., 1998. In vitro to in vivo extrapolation for trichloroethylene metabolism in humans. *Toxicol. Appl. Pharmacol.* 152, 376–387. <http://dx.doi.org/10.1006/taap.1998.8485>.
- Liu, X., Chism, J.P., LeCluyse, E.L., Brouwer, K.R., Brouwer, K.L.R., 1999. Correlation of biliary excretion in sandwich-cultured rat hepatocytes and in vivo in rats. *Drug Metab. Dispos.* 27, 637–644.
- Loizou, G.D., Hogg, A., 2011. MEGen: a physiologically based pharmacokinetic model generator. *Front. Pharmacol.* 2, 56. <http://dx.doi.org/10.3389/fphar.2011.00056>.
- Lukacova, V., Woltoz, W.S., Bolger, M.B., 2009. Prediction of modified release pharmacokinetics and pharmacodynamics from in vitro, immediate release, and intravenous data. *AAPS J.* 11, 323–334. <http://dx.doi.org/10.1208/s12248-009-9107-2>.
- Lumen, A., George, N.I., 2017. Estimation of iodine nutrition and thyroid function status in late-gestation pregnant women in the United States: development and application of a population-based pregnancy model. *Toxicol. Appl. Pharmacol.* 314, 24–38. <http://dx.doi.org/10.1016/j.taap.2016.10.026>.
- Lumen, A., Mattie, D.R., Fisher, J.W., 2013. Evaluation of perturbations in serum thyroid hormones during human pregnancy due to dietary iodide and perchlorate exposure using a biologically based dose-response model. *Toxicol. Sci.* 133, 320–341. <http://dx.doi.org/10.1093/toxsci/kft078>.
- Lumen, A., McNally, K., George, N., Fisher, J.W., Loizou, G.D., 2015. Quantitative global sensitivity analysis of a biologically based dose-response pregnancy model for the thyroid endocrine system. *Front. Pharmacol.* 6, 107. <http://dx.doi.org/10.3389/fphar.2015.00107>.
- McLanahan, E.D., El-Masri, H.A., Sweeney, L.M., Kopylev, L.Y., Clewell, H.J., Wambaugh, J.F., Schlosser, P.M., 2012. Physiologically based pharmacokinetic model use in risk assessment—why being published is not enough. *Toxicol. Sci.* 126, 5–15. <http://dx.doi.org/10.1093/toxsci/kfr295>.
- McNally, K., Cotton, R., Loizou, G.D., 2011. A workflow for global sensitivity analysis of PBPK models. *Predict. Toxicol.* 2, 31. <http://dx.doi.org/10.3389/fphar.2011.00031>.
- Mundy, W.R., Freudenrich, T.M., Crofton, K.M., DeVito, M.J., 2004. Accumulation of PBDE-47 in primary cultures of rat neocortical cells. *Toxicol. Sci.* 82, 164–169.
- National Research Council, 2011. Reference Manual on Scientific Evidence. The National Academies Press, third edition.
- National Research Council (US), 1983. Risk Assessment in the Federal Government: Managing the Process. National Academies Press, Washington, DC.
- National Research Council (US), 1994. Science and Judgment in Risk Assessment. National Academies Press, Washington, DC.
- Nichols, J.W., McKim, J.M., Andersen, M.E., Gargas, M.L., Clewell, H.J., Erickson, R.J., 1990. A physiologically based toxicokinetic model for the uptake and disposition of waterborne organic chemicals in fish. *Toxicol. Appl. Pharmacol.* 106, 433–447.
- NIEHS, 2017. Small Business Programs (SBIR/STTR). <https://www.niehs.nih.gov/research/supported/translational/sbir/index.cfm>, Accessed date: 21 July 2017.
- NTP, 2016. West Virginia Chemical Spill. <https://ntp.niehs.nih.gov/results/areas/wvspill/index.html>, Accessed date: 21 July 2017.
- NTP, 2017. In Vitro to in Vivo Extrapolation for High Throughput Prioritization and Decision Making. <https://ntp.niehs.nih.gov/pubhealth/evalatm/3rs-meetings/past-meetings/ivive-2016/ivive-2016.html>, Accessed date: 21 July 2017.
- OECD, 2016. Test No. 455: Performance-based Test Guideline for Stably Transfected Transactivation in Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists. OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264265295-en>.
- O'Flaherty, E.J., 1981. Toxicants and Drugs: Kinetics and Dynamics. Wiley, New York.
- Parhizgari, Z., Li, J., 2014. A physiologically-based pharmacokinetic model for disposition of 2,3,7,8-TCDD in fathead minnow and medaka. *Environ. Toxicol. Chem.* 33, 1064–1071. <http://dx.doi.org/10.1002/etc.2504>.
- Parrott, N., Lave, T., 2016. Computer models for predicting drug absorption. In: Dressman, J.B., Reppas, C. (Eds.). Oral Drug Absorption: Prediction and Assessment, second edition. Drugs and the Pharmaceutical Sciences CRC Press, Boca Raton, FL, pp. 338–350.
- Patlewicz, G., Simon, T.W., Rowlands, J.C., Budinsky, R.A., Becker, R.A., 2015. Proposing a scientific confidence framework to help support the application of adverse outcome pathways for regulatory purposes. *Regul. Toxicol. Pharmacol.* 71, 463–477. <http://dx.doi.org/10.1016/j.yrtph.2015.02.011>.
- Pearce, R.G., Setzer, R.W., Strope, C.L., Sipes, N.S., Wambaugh, J.F., 2017. R package for high-throughput toxicokinetics. *J. Stat. Softw.* 79 (1), 1–26.
- Pellegati, M., 2014. The debate on animal ADME studies in drug development: an update. *Expert Opin. Drug Metab. Toxicol.* 10, 1615–1620. <http://dx.doi.org/10.1517/1742525.2015.979152>.
- Perkins, E.J., Ankley, G.T., Crofton, K.M., Garcia-Reyero, N., LaLone, C.A., Johnson, M.S., Tietge, J.E., Villeneuve, D.L., 2013. Current perspectives on the use of alternative species in human health and ecological hazard assessments. *Environ. Health Perspect.* 121, 1002–1010. <http://dx.doi.org/10.1289/ehp.1306638>.
- Péry, A.R.R., Devillers, J., Brochet, C., Mombelli, E., Palluel, O., Piccini, B., Brion, F., Beaudouin, R., 2014. A physiologically based toxicokinetic model for the zebrafish *Danio rerio*. *Environ. Sci. Technol.* 48, 781–790. <http://dx.doi.org/10.1021/es404301q>.
- Peyret, T., Poulin, P., Krishnan, K., 2010. A unified algorithm for predicting partition coefficients for PBPK modeling of drugs and environmental chemicals. *Toxicol. Appl. Pharmacol.* 249, 197–207. <http://dx.doi.org/10.1016/j.taap.2010.09.010>.
- Phillips, M.B., Leonard, J.A., Grulke, C.M., Chang, D.T., Edwards, S.W., Brooks, R., Goldsmith, M.-R., El-Masri, H., Tan, Y.-M., 2016. A workflow to investigate exposure and pharmacokinetic influences on high-throughput in vitro chemical screening based on adverse outcome pathways. *Environ. Health Perspect.* 124, 53–60. <http://dx.doi.org/10.1289/ehp.1409450>.
- Poirier, A., Funk, C., Scherrmann, J.-M., Lavé, T., 2009. Mechanistic modeling of hepatic transport from cells to whole body: application to napsagatran and fexofenadine. *Mol. Pharm.* 6, 1716–1733. <http://dx.doi.org/10.1021/mp8002495>.
- Ring, C.L., Pearce, R.G., Setzer, R.W., Wetmore, B.A., Wambaugh, J.F., 2017. Identifying populations sensitive to environmental chemicals by simulating toxicokinetic variability. *Environ. Int.* 106, 105–118.
- Rotroff, D.M., Wetmore, B.A., Dix, D.J., Ferguson, S.S., Clewell, H.J., Houck, K.A., Lecluyse, E.L., Andersen, M.E., Judson, R.S., Smith, C.M., Sochaski, M.A., Kavlock, R.J., Boellmann, F., Martin, M.T., Reif, D.M., Wambaugh, J.F., Thomas, R.S., 2010. Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol. Sci.* 117, 348–358. <http://dx.doi.org/10.1093/toxsci/kqf220>.
- Rowland, M., Peck, C., Tucker, G., 2011. Physiologically-based pharmacokinetics in drug development and regulatory science. *Annu. Rev. Pharmacol. Toxicol.* 51, 45–73. <http://dx.doi.org/10.1146/annurev-pharmtox-010510-100540>.
- Sala Benito, J.V., Pains, A., Richarz, A.N., Meinl, T., Berthold, M.R., Cronin, M., Worth, A.P., 2017. Automated workflows for modelling chemical fate, kinetics and toxicity [published online ahead of print 18 March 2017]. *Toxicol. in Vitro*. <http://dx.doi.org/10.1016/j.tiv.2017.03.004>.
- Salem, F., Johnson, T.N., Abduljalil, K., Tucker, G.T., Rostami-Hodjegan, A., 2014. A re-evaluation and validation of ontogeny functions for cytochrome P450 1A2 and 3A4 based on in vivo data. *Clin. Pharmacokinet.* 53, 625–636. <http://dx.doi.org/10.1007/s40262-014-0140-7>.
- Schmitt, W., 2008. General approach for the calculation of tissue to plasma partition coefficients. *Toxicol. in Vitro* 22, 457–467. <http://dx.doi.org/10.1016/j.tiv.2007.09.010>.
- Segner, H., Navas, J.M., Schäfers, C., Wenzel, A., 2003. Potencies of estrogenic compounds in in vitro screening assays and in life cycle tests with zebrafish in vivo. *Ecotoxicol. Environ. Saf.* 54, 315–322.
- Slob, W., Janssen, P.H.M., van den Hof, J.M., 1997. Structural identifiability of PBPK models: practical consequences for modeling strategies and study designs. *Crit. Rev. Toxicol.* 27, 261–272. <http://dx.doi.org/10.3109/10408449709089895>.
- Smith, C.M., Nolan, C.K., Edwards, M.A., Hatfield, J.B., Stewart, T.W., Ferguson, S.S., Lecluyse, E.L., Sahi, J., 2012. A comprehensive evaluation of metabolic activity and intrinsic clearance in suspensions and monolayer cultures of cryopreserved primary human hepatocytes. *J. Pharm. Sci.* 101, 3989–4002. <http://dx.doi.org/10.1002/jps.23262>.
- Strobel, A., Burkhardt-Holm, P., Schmid, P., Segner, H., 2015. Benzo(a)pyrene metabolism and EROD and GST biotransformation activity in the liver of red- and white-blooded Antarctic fish. *Environ. Sci. Technol.* 49, 8022–8032. <http://dx.doi.org/10.1021/acs.est.5b00176>.
- Sullivan, K., 2016. It takes a village: stakeholder participation is essential to transforming science. *Altern. Lab. Anim* 44, 411–415.
- Swift, B., Pfeifer, N.D., Brouwer, K.L.R., 2010. Sandwich-cultured hepatocytes: an in vitro model to evaluate hepatobiliary transporter-based drug interactions and hepatotoxicity. *Drug Metab. Rev.* 42, 446–471. <http://dx.doi.org/10.3109/03602530903491881>.
- Swiss Institute of Bioinformatics, 2013. Click2Drug. <http://www.click2drug.org/index.html>, Accessed date: 21 July 2017.
- Tan, Y.-M., Liao, K.H., Clewell, H.J., 2007. Reverse dosimetry: interpreting trihalo-methanes biomonitoring data using physiologically based pharmacokinetic modeling. *J. Expo. Sci. Environ. Epidemiol.* 17, 591–603. <http://dx.doi.org/10.1038/sj.jes.7500540>.
- Teeguarden, J.G., Tan, Y.-M., Edwards, S.W., Leonard, J.A., Anderson, K.A., Corley, R.A., Kile, M.L., Simonich, S.M., Stone, D., Tanguay, R.L., Waters, K.M., Harper, S.L., Williams, D.E., 2016. Completing the link between exposure science and toxicology for improved environmental health decision making: the aggregate exposure pathway framework. *Environ. Sci. Technol.* 50, 4579–4586. <http://dx.doi.org/10.1021/acs.est.5b05311>.
- Thomas, R.S., Philbert, M.A., Auerbach, S.S., Wetmore, B.A., Devito, M.J., Cote, I., Rowlands, J.C., Whelan, M.P., Hays, S.M., Andersen, M.E., Meek, M.E.B., Reiter, L.W., Lambert, J.C., Clewell, H.J., Stephens, M.L., Zhao, Q.J., Wesselkamper, S.C., Flowers, L., Carney, E.W., Pastoor, T.P., Petersen, D.D., Yauk, C.L., Nong, A., 2013. Incorporating new technologies into toxicity testing and risk assessment: moving from 21st century vision to a data-driven framework. *Toxicol. Sci.* 136, 4–18. <http://dx.doi.org/10.1093/toxsci/kft178>.
- Tice, R.R., Austin, C.P., Kavlock, R.J., Bucher, J.R., 2013. Improving the human hazard characterization of chemicals: a Tox21 update. *Environ. Health Perspect.* 121, 756–765. <http://dx.doi.org/10.1289/ehp.1205784>.
- Tolonen, A., Pelkonen, O., 2015. Analytical challenges for conducting rapid metabolism characterization for QIVIVE. *Toxicology* 332, 20–29. <http://dx.doi.org/10.1016/j.tox.2013.08.010>.
- Tonneller, A., Coecke, S., Zaldívar, J.-M., 2012. Screening of chemicals for human bioaccumulation potential with a physiologically based toxicokinetic model. *Arch. Toxicol.* 86, 393–403. <http://dx.doi.org/10.1007/s00204-011-0768-0>.
- U.S. Fish and Wildlife Service, 2015. Endangered Species. <https://www.fws.gov/endangered/species/us-species.html>, Accessed date: 21 July 2017.
- U.S. Food and Drug Administration, 2017. Biomarker Qualification Program. <https://www.fda.gov/drugs/developmentapprovalprocess/drugdevelopmenttoolsqualificationprogram/biomarkerqualificationprogram/default.htm>, Accessed date: 21 July 2017.
- Varma, M.V., Steyn, S.J., Allerton, C., El-Kattan, A.F., 2015. Predicting clearance mechanism in drug discovery: extended clearance classification system (ECCS). *Pharm. Res.* 32, 3785–3802. <http://dx.doi.org/10.1007/s11095-015-1749-4>.

- Villeneuve, D.L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T.H., LaLone, C.A., Landesmann, B., Lettieri, T., Munn, S., Nepelska, M., Ottinger, M.A., Vergauwen, L., Whelan, M., 2014a. Adverse outcome pathway (AOP) development I: strategies and principles. *Toxicol. Sci.* 142, 312–320. <http://dx.doi.org/10.1093/toxsci/kfu199>.
- Villeneuve, D.L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T.H., LaLone, C.A., Landesmann, B., Lettieri, T., Munn, S., Nepelska, M., Ottinger, M.A., Vergauwen, L., Whelan, M., 2014b. Adverse outcome pathway development II: best practices. *Toxicol. Sci.* 142, 321–330. <http://dx.doi.org/10.1093/toxsci/kfu200>.
- Votano, J.R., Parham, M., Hall, L.H., Kier, L.B., 2004. New predictors for several ADME/tox properties: aqueous solubility, human oral absorption, and Ames genotoxicity using topological descriptors. *Mol. Divers.* 8, 379–391.
- Wagner, C., Zhao, P., Pan, Y., Hsu, V., Grillo, J., Huang, S., Sinha, V., 2015. Application of physiologically based pharmacokinetic (PBPK) modeling to support dose selection: report of an FDA public workshop on PBPK. *CPT Pharmacomet. Syst. Pharmacol.* 4, 226–230. <http://dx.doi.org/10.1002/psp4.33>.
- Wambaugh, J.F., Wang, A., Dionisio, K.L., Frame, A., Egeghy, P., Judson, R., Setzer, R.W., 2014. High throughput heuristics for prioritizing human exposure to environmental chemicals. *Environ. Sci. Technol.* 48, 12760–12767. <http://dx.doi.org/10.1021/es503583j>.
- Wambaugh, J.F., Wetmore, B.A., Pearce, R., Strobe, C., Goldsmith, R., Sluka, J.P., Sedykh, A., Tropsha, A., Bosgra, S., Shah, I., Judson, R., Thomas, R.S., Setzer, R.W., 2015. Toxicokinetic triage for environmental chemicals. *Toxicol. Sci.* 147, 55–67. <http://dx.doi.org/10.1093/toxsci/kfv118>.
- Wang, Y.-H., 2010. Confidence assessment of the Simcyp time-based approach and a static mathematical model in predicting clinical drug-drug interactions for mechanism-based CYP3A inhibitors. *Drug Metab. Dispos.* 38, 1094–1104. <http://dx.doi.org/10.1124/dmd.110.032177>.
- Watanabe, K.H., Jensen, K.M., Orlando, E.F., Ankley, G.T., 2007. What is normal? A characterization of the values and variability in reproductive endpoints of the fathead minnow, *Pimephales promelas*. *Comp. Biochem. Physiol. C: Toxicol. Pharmacol.* 146, 348–356. <http://dx.doi.org/10.1016/j.cbpc.2007.04.015>.
- Wetmore, B.A., Wambaugh, J.F., Ferguson, S.S., Sochaski, M.A., Rotroff, D.M., Freeman, K., Clewell, H.J., Dix, D.J., Andersen, M.E., Houck, K.A., Allen, B., Judson, R.S., Singh, R., Kavlock, R.J., Richard, A.M., Thomas, R.S., 2012. Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol. Sci.* 125, 157–174. <http://dx.doi.org/10.1093/toxsci/kfr254>.
- Wetmore, B.A., Allen, B., Clewell, H.J., Parker, T., Wambaugh, J.F., Almond, L.M., Sochaski, M.A., Thomas, R.S., 2014. Incorporating population variability and susceptible subpopulations into dosimetry for high-throughput toxicity testing. *Toxicol. Sci.* 142, 210–224. <http://dx.doi.org/10.1093/toxsci/kfu169>.
- Wetmore, B.A., Wambaugh, J.F., Allen, B., Ferguson, S.S., Sochaski, M.A., Setzer, R.W., Houck, K.A., Strobe, C.L., Cantwell, K., Judson, R.S., LeCluyse, E., Clewell, H.J., Thomas, R.S., Andersen, M.E., 2015. Incorporating high-throughput exposure predictions with dosimetry-adjusted in vitro bioactivity to inform chemical toxicity testing. *Toxicol. Sci.* 148, 121–136. <http://dx.doi.org/10.1093/toxsci/kfv171>.
- Wilkinson, G.R., Shand, D.G., 1975. Commentary: a physiological approach to hepatic drug clearance. *Clin. Pharmacol. Ther.* 18, 377–390.
- Woodhead, J.L., Yang, K., Siler, S.Q., Watkins, P.B., Brouwer, K.L.R., Barton, H.A., Howell, B.A., 2014. Exploring BSEP inhibition-mediated toxicity with a mechanistic model of drug-induced liver injury. *Front. Pharmacol.* 5, 240. <http://dx.doi.org/10.3389/fphar.2014.00240>.
- Woodhead, J.L., Brock, W.J., Roth, S.E., Shoaf, S.E., Brouwer, K.L.R., Church, R., Grammatopoulos, T.N., Stiles, L., Siler, S.Q., Howell, B.A., Mosedale, M., Watkins, P.B., Shoda, L.K.M., 2017. Application of a mechanistic model to evaluate putative mechanisms of tolvaftan drug-induced liver injury and identify patient susceptibility factors. *Toxicol. Sci.* 155, 61–74. <http://dx.doi.org/10.1093/toxsci/kfw193>.
- Yang, F., Sun, N., Sun, Y.X., Shan, Q., Zhao, H.Y., Zeng, D.P., Zeng, Z.L., 2013. A physiologically based pharmacokinetics model for florfenicol in crucian carp and oral-to-intramuscular extrapolation. *J. Vet. Pharmacol. Ther.* 36, 192–200. <http://dx.doi.org/10.1111/j.1365-2885.2012.01419.x>.
- Yang, K., Woodhead, J.L., Watkins, P.B., Howell, B.A., Brouwer, K.L.R., 2014. Systems pharmacology modeling predicts delayed presentation and species differences in bile acid-mediated troglitazone hepatotoxicity. *Clin. Pharmacol. Ther.* 96, 589–598. <http://dx.doi.org/10.1038/clpt.2014.158>.
- Yang, K., Pfeifer, N.D., Köck, K., Brouwer, K.L.R., 2015. Species differences in hepatobiliary disposition of taurocholic acid in human and rat sandwich-cultured hepatocytes: implications for drug-induced liver injury. *J. Pharmacol. Exp. Ther.* 353, 415–423. <http://dx.doi.org/10.1124/jpet.114.221564>.
- Yoon, M., Kedderis, G.L., Yang, Y., Allen, B.C., Yan, G.Z., Clewell, H.J., 2012. Use of in vitro data in PBPK models: an example of in vitro to in vivo extrapolation with carbaryl. In: Knaak, J.B., Timchalk, C.A., Tornero-Velez, R. (Eds.), *Parameters for Pesticide QSAR and PBPK/PD Models for Human Risk Assessment*. ACS Symposium Series American Chemical Society, Washington, DC, pp. 323–338. <http://dx.doi.org/10.1021/bk-2012-1099.ch020>.
- Yoon, M., Efremenko, A., Blaauboer, B., Clewell, H.J., 2014. Evaluation of simple in vitro to in vivo extrapolation approaches for environmental compounds. *Toxicol. in Vitro* 28, 164–170.
- Zaldivar Comenges, J.M., Joossens, E., Sala Benito, J.V., Worth, A., Paini, A., 2017. Theoretical and mathematical foundation of the virtual cell based assay - a review. *Toxicol. in Vitro* 45, 209–221. <http://dx.doi.org/10.1016/j.tiv.2016.07.013>.
- Zhou, H., Fraczekiewicz, G., Lawless, M., 2016. Using physiologically based pharmacokinetic modeling for in vitro in vivo extrapolation to predict chemical exposure [Poster]. In: *In Vitro to In Vivo Extrapolation for High Throughput Prioritization and Decision Making*; Durham, NC, (17–18 February 2016).
- Zhu, X., Sedykh, A., Zhu, H., Liu, S., Tropsha, A., 2013. The use of pseudo-equilibrium constant affords improved QSAR models of human plasma protein binding. *Pharm. Res.* 30, 1790–1798. <http://dx.doi.org/10.1007/s11095-013-1023-6>.