t⁴ Workshop Report

Optimizing Drug Discovery by Investigative Toxicology: Current and Future Trends

Mario Beilmann[1,§], Harrie Boonen[2,§], Andreas Czich[3,§], Gordon Dear[4,§], Philip Hewitt[5,§], Tomas Mow[6,§], Peter Newham[7,§], Teija Oinonen[8,§], Francois Pognan[9,§], Adrian Roth[10,§], Jean-Pierre Valentin[12,§], Freddy Van Goethem[13,§], Richard J. Weaver[14,§], Barbara Birk[15], Scott Boyer[16], Francesca Caloni[17], Alice E. Chen[18], Raffaella Corvi[19], Mark T. D. Cronin[20], Mardas Daneshian[21], Lorna C. Ewart[7], Rex E. FitzGerald[22], Geraldine A. Hamilton[23], Thomas Hartung[21,24], Joshua D. Kangas[25], Nynke I. Kramer[26], Marcel Leist[21], Uwe Marx[27], Sebastian Polak[28,29], Costanza Rovida[21], Emanuela Testai[30], Bob van de Water[31], Paul Vulto[32] and Thomas Steger-Hartmann[11,§]

[1]Nonclinical Drug Safety, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany; [2]H. Lundbeck A/S, Valby, Denmark; [3]Sanofi Deutschland GmbH, Frankfurt, Germany; [4]Platform Technology & Science, David Jack Centre for R&D, GSK, Hertfordshire, UK; [5]Non Clinical Safety, Merck KGaA, Darmstadt, Germany; [6]Global Discovery and Development Sciences, Novo Nordisk A/S, Maaloev, Denmark; [7]Drug Safety and Metabolism, Astra Zeneca, Cambridge, UK; [8]Investigative Toxicology and ADME, Orion Pharma, Espoo, Finland; [9]PreClinical Safety, Novartis Pharma, Basel, Switzerland; [10]Pharma Research and Early Development, Hoffmann-La Roche Ltd, Basel, Switzerland; [11]Research & Development, Pharmaceuticals, Investigational Toxicology, Bayer AG, Berlin, Germany; [12]Development Science, UCB Biopharma SPRL, Braine-l'Alleud, Belgium; [13]Mechanistic & Investigative Toxicology, Discovery Sciences, Janssen Research & Development, Beerse, Belgium; [14]Research & Biopharmacy, Servier, Suresnes Cedex, France; [15]Experimental Toxicology and Ecology, BASF, Ludwigshafen, Germany; [16]Computational Toxicology, Swedish Toxicological Sciences Research Center, Södertäljje, Sweden; [17]Department of Veterinary Medicine (DIMEVET), Università degli Studi di Milano, Milan, Italy; [18]Organovo, Inc., San Diego, USA; [19]EURL-ECVAM, Joint Research Center, European Commission, Ispra (VA), Italy; [20]School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Liverpool, England; [21]Center for Alternatives to Animal Testing (CAAT)-Europe, University of Konstanz, Konstanz, Germany; [22]Swiss Centre for Applied Human Toxicology, University of Basle, Basle, Switzerland; [23]Emulate Inc., Boston, MA, USA; [24]John Hopkins University, CAAT, Baltimore, USA; [25]Computational Biology Department, Carnegie Mellon University, Pittsburgh, USA; [26]Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands; [27]TissUse, GmbH, Berlin, Germany; [28]Certara UK (Simcyp), Sheffield, United Kingdom; [29]Jagiellonian University Medical College, Kraków, Poland; [30]Istituto Superiore di Sanità, Rome, Italy; [31]Division of Drug Discovery and Safety, Leiden Academic Centre for Drug Research (LACDR), Leiden University, Leiden, The Netherlands; [32]Mimetas, Leiden, The Netherlands

§ Members of the Investigative Toxicology Leaders Forum (ITLF)

Abstract

Investigative toxicology describes the de-risking and mechanistic elucidation of toxicities, supporting early safety decisions in the pharmaceutical industry. Recently, investigative toxicology has contributed to a shift in pharmaceutical toxicology, from a descriptive to an evidence-based, mechanistic discipline. This was triggered by high costs and low throughput of Good Laboratory Practice *in vivo* studies, and increasing demands for adherence to the 3R (Replacement, Reduction, and Refinement) principles of animal welfare. Outside the boundaries of regulatory toxicology, investigative toxicology has the flexibility to embrace new technologies, enhancing translational steps from *in silico, in vitro* to *in vivo* mechanistic understanding to eventually predict human response. One major goal of investigative toxicology is to improve pre-clinical decisions, which coincides with the concept of animal-free safety testing. Currently, compounds under preclinical development are being discarded owing to the use of inappropriate animal models. Progress in investigative toxicology could lead to humanized *in vitro* test systems and the development of medicines less reliant on animal tests. To advance this field, a group of 14 European-based leaders from the pharmaceutical industry founded the Investigative Toxicology Leaders Forum (ITLF), an open, non-exclusive, and pre-competitive group that shares knowledge and experience. The ITLF collaborated with the Centre for Alternatives to Animal Testing Europe (CAAT-Europe) to organize an "Investigative Toxicology Think Tank", which aimed to enhance interaction with experts from academia and regulatory bodies in the field. Summarizing the topics and discussion of the workshop, this article highlights investigative toxicology's position by identifying key challenges and perspectives.

1 Introduction

Tremendous progress in preclinical development across the pharmaceutical industry has been achieved over the past three decades. This pivotal phase, which prepares the transition into first-in-man trials, has been strongly harmonized under the umbrella of the International Conference for Harmonisation (ICH) of Technical Requirements for Pharmaceuticals for Human Use[1] (Ohno, 2002). The ICH has contributed to an internationally accepted set of submission-relevant guideline documents generally concerning *in vivo* drug safety studies, which are based on Organisation for Economic Co-operation and Development (OECD) test guidelines for the individual study conduct and are strongly connected to OECD documents for Good Laboratory Practice (GLP). The whole framework of harmonization has led to an increase of mutual acceptance of preclinical submission documents in the three regions involved (European Union, the United States, and Japan) and, as a consequence, to elimination of studies performed to satisfy the specific requirements of individual regulatory authorities (OECD, 2005).

This positive development of streamlined testing strategies, however, also brought some challenges with it. Many companies are reluctant to integrate new technologies or assays next to the established safety assessment consisting of GLP *in vivo* studies due to the perception that non-GLP, or not fully validated assays, might compromise the pivotal studies and endanger the approval process. The high standard of harmonization and validation developed for the GLP studies is often requested for new technologies and assays. On the other hand, attrition due to safety reasons in preclinical and clinical drug development phases still represents a major factor for the overall loss of projects and therefore there is pressure to improve the predictive power of preclinical studies, including new screening strategies. In addition, the high costs and the rather low throughput of GLP *in vivo* studies and the intensifying demands to address the 3Rs has increased the push towards new screening strategies (Sewell et al., 2017).
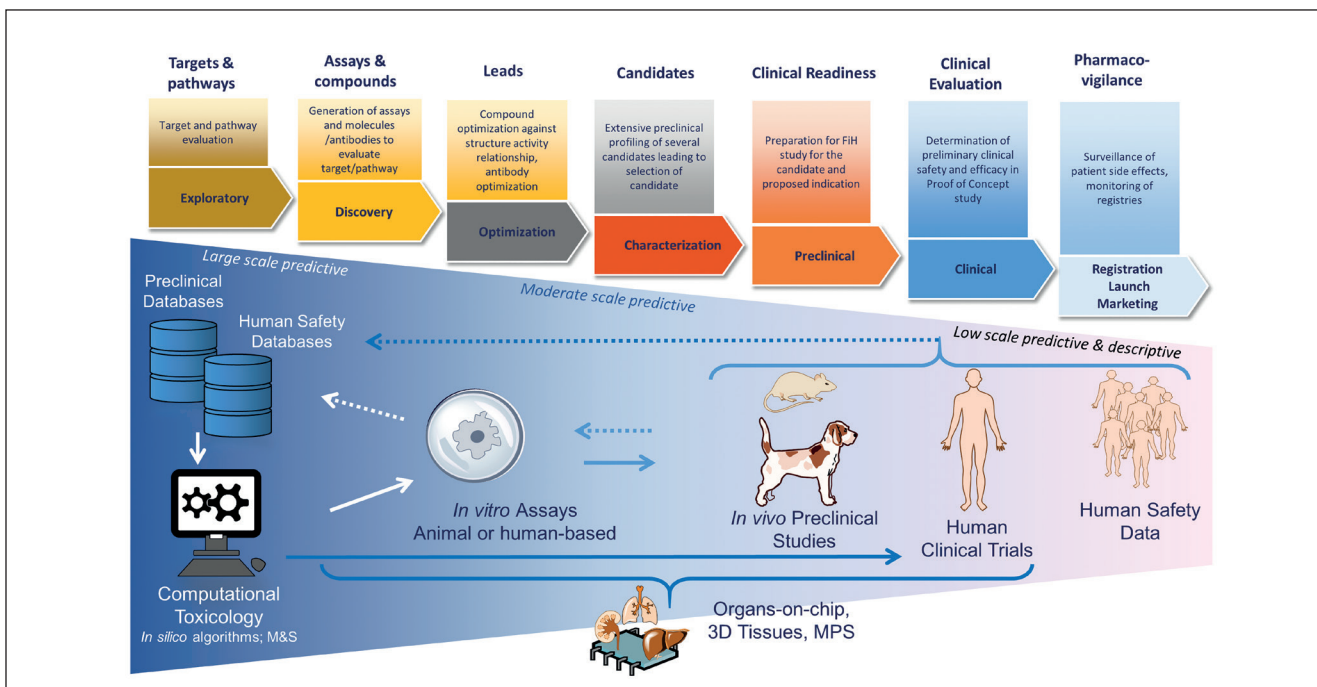
Fig. 1: A visual illustration of the continuum of investigative toxicology in the drug discovery and development pipeline
Plain arrows represent the forward feed information to move to the next step, while dashed arrows represent back feed of knowledge to improve predictivity. Organ-on-chip, 3D tissues and MPS have the potential to complement and, perhaps to some extent, replace certain steps of research and development. FiH, First-in-Human trial; MPS, microphysiological systems; M&S, modeling and simulation

---

[1] ICH Guideline M3 (R2). Non-clinical safety studies for the conduct of human clinical trials for pharmaceuticals. CPMP/ICH/286/95.
https://www.ema.europa.eu/documents/scientific-guideline/ich-m-3-r2-non-clinical-safety-studies-conduct-human-clinical-trials-marketing-authorization_en.pdf

---

Abbreviations

3D, three dimensional; 3Rs, refine, reduce, replace animal experimentation; ADR, adverse drug reaction; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AOP, adverse outcome pathway; AST, aspartate aminotransferase; BIL, bilirubin; DILI, drug-induced liver injury; EMA, European Medicines Agency; FDA, Food and Drug Administration; GLP, Good Laboratory Practice; IATA, integrated approaches to testing and assessment; ICH, International Conference for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; IMI, Innovative Medicines Initiative; iPSC, induced pluripotent stem cells; ITLF, Investigative Toxicology Leaders Forum; KE, key event; LLNA, local lymph node assay; MIE, molecular initiating event; MPS, microphysiological system; NBE, new biological entity; NCE, new chemical entity; NOAEL, no observed adverse effect level; OECD, Organisation for Economic Co-operation and Development; PBPK, physiology-based pharmacokinetic modeling; PoD, point of departure; PoT, pathways of toxicity; QIVIVE, quantitative *in vitro* to *in vivo* extrapolations; (Q)SAR, (quantitative) structure activity relationships; WoE, weight of evidence

As a consequence, most pharmaceutical companies have established specific toxicology functions, which complement the experimental GLP functions. Some companies have even gone so far as to fully outsource GLP activities and to focus in-house on preclinical safety activities on what is termed "discovery", "exploratory", "mechanistic", or "investigative" toxicology. While the tasks and organizational set-up of these functions differ from company to company, it has become evident that the value of these activities lies not only in screening assays preceding regulatory activities, but also in an enhanced understanding of the mechanism of toxicity, which is equally relevant for later phases of clinical development. In fact, this is shifting pharmaceutical toxicology from a purely descriptive to an evidence-based mechanistic discipline. For this reason, the authors of this publication prefer the term "investigative toxicology" over "discovery toxicology", since it avoids the perceptional limitation to serve only the early phases of safety assessment. The continuum of investigational toxicology in the drug development process is illustrated in Figure 1. Thus, from a phar-

maceutical industry point of view, the term "investigative toxicology" can be defined as a complementary effort to regulatory toxicology, encompassing both a prospective (screening for de-risking) and a retrospective approach (mechanistic investigations of adverse effects) (Moggs et al., 2012).

To foster awareness, development, and implementation of investigational toxicology and to share expertise, knowledge, and best practice in a pre-competitive space, a group of European-based investigative toxicology leaders from the pharmaceutical industry (see Fig. 2 for participating companies) founded the Investigative Toxicology Leaders Forum (ITLF) (Roth, 2017). This open, non-exclusive forum aims to enhance interaction with experts from academia and regulatory bodies in the field of investigative toxicology. The objective of the ITLF is to elaborate robust, reliable, and accepted investigative toxicology concepts and practices for decision-making related to early safety-related attrition, de-risking, and mechanistic elucidation of effects as shown in Figure 3. The figure illustrates how investigative toxicolo-

Fig. 2: Companies participating in the pre-competitive Investigative Toxicology Leaders Forum (ITLF) as of July 2018
Objectives of the ITLF are to jointly elaborate robust, reliable and accepted investigative toxicology concepts for decision-making for early safety-related attrition, de-risking, and mechanistic elucidation of safety-related effects to increase the understanding and improve the translation of *in vitro* to *in vivo* mechanistic data. Furthermore, the adoption of new technologies/platforms into the drug discovery back-bone is targeted by the forum to increase the knowledge and awareness of investigative toxicology as a discipline (e.g., through publications, meetings, and conferences).
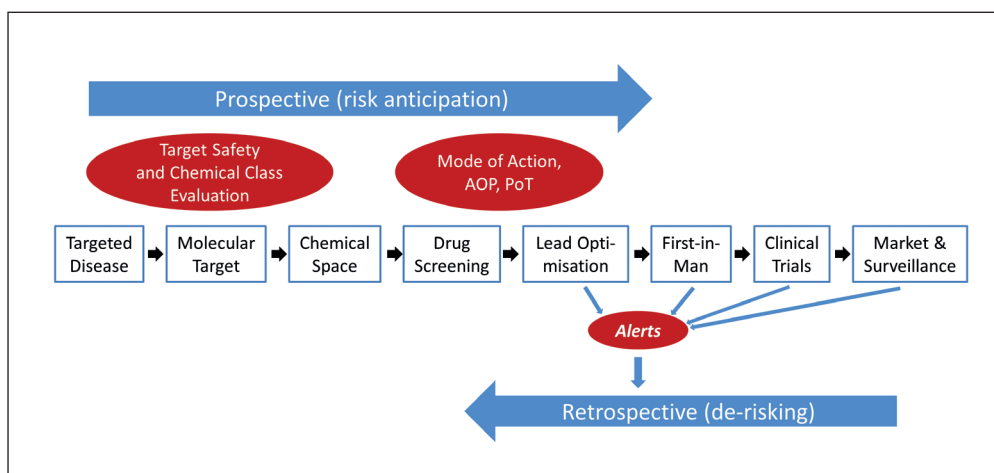
Fig. 3: Key objectives of investigative toxicology during drug discovery and development

gy adds to the traditional drug development process. Investigative toxicology supports the entire process by early assessments of target- and chemical class-related toxicological concerns and front-loading of assessments as prospective risk anticipation. Furthermore, alerts from later stages of development and market surveillance can trigger a retrospective de-risking process, which will typically include the elucidation of toxic mechanisms to assess their relevance to humans and possible mitigation strategies.

The goal of investigative toxicology is to improve preclinical decision-making, which coincides with the notion of animal-free safety testing. Currently, many compounds are ruled out based on results from animal models obtained during the preclinical phase without knowledge of how the compounds would behave in humans, i.e., the false-positive rate of animal studies cannot be assessed. In addition, significant attrition occurs in clinical phases due to safety issues that were not adequately identified during the preclinical phase (false-negative) (Waring et al., 2015). Progress in investigative toxicology towards humanized *in vitro* test systems promises a better rate of human-relevant predictions.

For this reason, the ITLF teamed up with CAAT-Europe to hold an "Investigative Toxicology Think Tank". in July 2017, which assembled 34 experts from academia, the pharmaceutical and other industries, regulatory authorities, and technology providers to develop a definition of "investigative toxicology" and to align academic and expert stakeholders with the needs for a predictive and mechanistic investigational toxicology. Although the focus of the meeting was on investigative toxicology in drug development, progress in this field may also influence safety assessment in other industry sectors (industrial, consumer or agro-chemical compounds). This report represents a position paper for investigative toxicology based on the topics of and discussions during the workshop. It starts with a gap analysis, followed by a critical assessment of new technologies, and finishes by summarizing challenges, and presenting perspectives and recommendations.

2 Gap analysis

The pharmaceutical industry has made substantial efforts towards the implementation of *in vitro* based models, which has improved the hazard identification and risk assessment of drug candidates prior to non-clinical development (Hornberg et al., 2014a,b; Goh et al., 2015). However, much remains to be accomplished to address the substantial gaps in our mechanistic understanding of adverse drug reactions (ADRs) and to support the development of biomedical tools that are truly predictive of inter-individual human susceptibility to ADRs. The rapid "design-make-test-analyze" cycle time in drug discovery also places greater emphasis to further the understanding of the mechanisms of toxicity and chemical liabilities, and to facilitate the decision-making processes on candidate selection and development of new chemical entities (NCEs) and new biological entities (NBEs). In addition, the increasing diversity of biopharmaceutics, which now include cell and gene therapies, chimeric antigen receptor T (CAR-T) cells and vectors, antibodies, and anti-sense oligonucleotides presents new and significant risks, such as cytokine release syndrome (CRS) and tissue cross-reactivity issues, with a variety of new challeng-

es to developing testing strategies (Brennan et al., 2015; Dixit and Boelsterli, 2007; Bussiere et al., 2009). The following outlines the perceived gaps in identifying NCE/NBE hazards for target-organ toxicities, limiting risk assessments and prediction of human safety, mitigation strategies to manage risk, and current governance for investigative toxicological sciences. The mechanisms of ADRs are extensively reviewed elsewhere (Atienzar et al., 2016; Hornberg et al., 2014a,b).

2.1 Hazard identification risk assessment

2.1.1 Target organ toxicity models
Although 70% of human-relevant toxicities are detected in experimental species (Olson et al., 2000), the translational relevance of these toxicities is highly dependent on the affected target organ. Significant human ADRs are predominantly associated with liver, heart, and neurological organs (Cook et al., 2014; Olson et al., 2000; Sacks et al., 2014). The detection of dose-dependent drug hepatocellular cytotoxicity by *in vitro* cell-based models and animal studies is well accepted (Antoine et al., 2013; Ward et al., 2014), yet the multifactorial nature of drug induced liver injury (DILI) and known species differences are notable gaps that require the development of humanized models to detect liver injuries associated with immune or patient-specific susceptibilities. For the identification of cardiovascular drug liabilities, the concomitant use of both *in vitro* models and animal studies is well established (Laverty et al., 2011; Valentin et al., 2010). Whereas effects on ion channels can be easily identified by *in vitro* (e.g., patch clamp) models, the complex interplay between heart rate, ejection volume, and blood pressure eventually causing heart or kidney damage can currently only be assessed by *in vivo* models. Similarly, early *in vitro* prediction of neurological ADR is still challenging as many side effects can often only be detected in clinical trials since they are caused by interactions with rare targets or occur only after chronic administration, which is difficult to achieve in *in vitro* assays (Schmidt et al., 2017). Regarding other toxicities such as hematologic or hematopoietic disorders or carcinogenic risk, few *in vitro* models exist due to the nature and complexity of the underlying pathology (see Tab. 1). However, there is a surge in the development of organotypic and microphysiological systems (MPS) (Marx et al., 2016) including multiple organ systems. High expectations are placed in improved detection of drug liabilities for use in safety assessment by the use of these innovative three dimensional (3D) models (Hardwick et al., 2017; Lin et al., 2015; Mueller et al., 2014; Soldatow et al., 2013).

2.1.2 Disease models
Disease models are required to emulate organ-level functions and recapitulate key phenotypic features of human disease in cell or tissue-based as well as conventional and transgenic animal models. Disease status can impact considerably on the toxicity of substances and thus the target population of a novel drug candidate. Nevertheless, animal models established to reflect human disease often appear to have had limited success (Benam et al., 2015; Morgan et al., 2013) and likely contribute to the poor predictivity of efficacy and safety of drugs in later human clinical trials. The future incorporation and use of humanized *in vitro* disease models

Tab. 1: Categories of safety attrition challenges

Attrition	Organ		Discovery Cell & tissue-based models			Non-clinical (tox species)	Translational value of models & gaps
			Hazard	Risk assessment	Predictive		
High	Heart		●	●	●	● (1)	Relatively good for detection and prediction of functional changes but limited in relation to structural changes.
High	Liver	Hepatocellular	●		●	● (2)	Relatively good concordance between hepatocellular injury IVIVE (*in vitro-in vivo* extrapolation) and liver injury in humans.
		Cholestasis (acute)	●	x	●	x	Animal models are poorly predictive of cholestasis.
High	CNS		o	x	x	● (1,2)	Models available for behavior, seizures, drug abuse; no models for cognition, suicidal ideation.
Medium	Gastrointestinal		o	x	x		Limited to poor model detection and prediction
Medium	Kidney		o	x	x	● (1,2)	Translational biomarkers permit ID of injury in animal models.
Medium	Immune system		x	x	x	x	Hypersensitivity reactions – poor. No tests are yet available for testing *ab initio* in drug discovery or non-clinical testing.
Low	Lung		●	●	●	● (1,2)	Relatively good
Low	Hematology		x	x	x	● (1,2)	No *in vitro* model, yet good *in vivo* concordance animal - human
Low	Hemopoiesis		x	x	x	● (1,2)	No *in vitro* model, yet good *in vivo* concordance animal - human
Low	Skin (irritation/ sensitization)		●	●	o	● (2,3)	*In vitro* models available with *in vivo* confirmatory and/or studies
Medium	Reproductive organs & embryofetal development		x	x	x	● (2)	Relatively good models for detection and prediction of embryofetal toxicities, but limited in relation to reproductive organs.
Low	Genetic toxicity & Carcinogenicity		● / o	x	● / o	● (2)	*In vitro* models good for detection and prediction of genetic toxicity but poor for carcinogenicity risk identification

X, no model; o, models yet to be evaluated for application in drug R&D; ●, models routinely available/in use.
1, large species (dog or monkey); 2, rodent; 3, lagomorph

into toxicity assessments has the potential to concomitantly facilitate pharmacological discovery and safety evaluation of drugs (Hübner et al. 2018). This will improve the identification of safety margins with the potential to extrapolate phenotypic differences in patient populations and lead to mechanistically-driven safety margins in patient populations.

2.1.3 Safety margins
Dose limiting toxicity and the "no adverse effect level" (NOAEL) define safety margins[2] and toxicological profile for risk-benefit assessment of a drug candidate[3]. However, the NOAEL is often a subjective assessment of a biochemical or histopathological change on which to establish a safety margin (Dorato and Engelhardt, 2005). In the absence of well-characterized safety biomarkers[4] with clearly defined mechanistic and translational relevance to humans, preclinical findings in animals will, at best, only permit rough estimates of the safety margins. The pre-requisite for non-clinical safety testing is to select candidate drugs with large safety margins to improve the likelihood of clinical success. However, *in vivo* toxicological and clinical findings can result in unexpected and reduced safety margins in target organs during the clinical phase. Progress towards the identification of novel, sensitive biomarkers with mechanistic and translational relevance may help to improve the monitoring of drug safety profiles.

---

[2] In other, non-pharma sectors, the common expression is "margin of safety" (MoS). Instead of NOAEL, which is the highest experimental dose in an *in vivo* study that is without observable adverse effect, the Benchmark Dose Lower Confidence Limit based on benchmark dose modelling (BMDL) is more frequently used in these sectors.

[3] http://www.fda.gov/downloads/drugs/guidances/ucm078932.pdf

[4] A biomarker is a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions. Molecular, histologic, radiographic, or physiologic characteristics are types of biomarkers. Safety biomarkers are applied to indicate the likelihood, presence, or extent of toxicity as an adverse effect (for definitions see: https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/BiomarkerQualificationProgram/default.htm)

A better understanding of mechanistic toxicokinetics and toxicodynamics (TKTD) relationships in combination with pharmacokinetics and pharmacodynamics (PKPD) should establish improved quantitative monitoring of safety margins in non-clinical and clinical research.

2.2 Predicting human safety with mechanistic insight

2.2.1 Adverse outcome pathways and pathways of toxicity

The organization of mechanistic knowledge into temporal events includes pathways of toxicity (PoT) (Kleensang, 2014), mode of action (MoA), and adverse outcome pathways (AOPs) (Ankley et al., 2010; Burden et al., 2015; Villeneuve et al., 2014). An AOP describes a sequential chain of causally linked events at different levels of biological organization that lead to an adverse effect on human health. AOPs best define the qualitative organization of information, whilst PoT relates more to quantitative, dynamic, and molecularly defined systems. The application of AOPs with an understanding of mechanisms can help adopt novel biomarkers for use in the identification and monitoring of safety signals. Nevertheless, these are of limited value unless signals identified in *in vitro* and animal models can be linked to human ADRs through either target-based or phenotypic-based testing as weight of evidence (WoE) to facilitate improved risk assessment of human target organ toxicities.

2.2.2 Idiosyncratic and hypersensitivity reactions

Difficulties in the detection of hypersensitivity reactions and idiosyncratic toxicities arise due to ADR events that often occur already at low therapeutic dose levels in only small numbers of individuals during clinical development or post-registration (Pallardy and Bechara, 2017; Park et al., 2000; Uetrecht, 2013). The "non-existence" of relevant humanized pre-clinical models for early testing of drug candidates, coupled with the absence of clear dose-related toxicities and the complex dimensionality of immune-drug response necessitates urgent research to establish innovative diagnostic assays for drug discovery and continued efforts towards understanding mechanisms to support research and development of safer drugs. A successful example of such research is the specific case of hypersensitivity, namely skin sensitization, where the application of the AOP concept has led to a series of approved *in vitro* assays replacing animal studies (OECD, 2014).

2.2.3 Translational gap

Significant gaps remain on the path to achieving fully integrated and characterized humanized organ-specific panel(s) of *in vitro* models. Use of such models will require arrays of qualified mechanistic translational safety biomarkers, while dose (exposure) dependent toxicities will continue to rely on observational or phenomenological-based endpoint tests and WoE approaches to assess human drug safety (see above). Therefore, much work remains to be done towards establishing the next generation of *in vitro* models for target organ safety testing. This includes the identification of novel safety biomarkers with translational value from early *in vitro* safety assessment to non-clinical as well as clinical safety assessment. More recently, gene editing technologies such as CRISPR/Cas9 have allowed refined engineering of animal disease models. For example, pig models have been established for cystic fibrosis research, which are reported to be superior compared to the existing mouse models regarding their similarity to the human phenotype (Klymiuk et al., 2016). These technologies have the potential to bridge the gap between proof-of-concept studies in animals and clinical trials in patients, thus supporting translational medicine.

2.2.4 Monitoring safety signals

Biomarkers permit identification and monitoring of potential safety signals (see Section 3.2) by employing a broad spectrum of technologies. Biomarkers detected by imaging and molecular techniques have advanced in recent years for on-target and off-target assessments. Despite a rapid rise in the use of omics technologies, several challenges remain before their routine adoption and application can be achieved (Khan et al., 2014). For example, genomic and pharmacogenomic screening have found use in clinical trial enrolment for an indirect assessment of drug metabolizing enzyme activity, yet the measurement of enzyme activity and drug-drug interaction (DDI) (Ward et al., 2014) only results in predictive values of around 40% when relying on protein and transcriptomic data alone (Weaver, 2001). In contrast, transcriptomics has yielded more success in the detection of organ-specific or selective pathologies (Chen et al., 2012), but nevertheless only appears to share similar sensitivity to that of established biomarkers (Zhang et al., 2012). A further limitation to the use and implementation of omics is that they currently require invasive biopsies.

2.2.5 Data transparency

The conduct and design of experimental studies has often drawn criticism due to the incompleteness of published data and the lack of reproducibility of results. In addition, the lack of data standards, definitions, and ontologies represents a major hurdle for modelling and simulation exercises. However, the reuse and sharing of available public and private data, both within and across organizations, is progressively recognized as a valuable source of information for read-across, hazard identification, and risk mitigation. The described hurdles are increasingly addressed through data governance frameworks. These efforts towards harmonization of study design, data curation, and controls are more widely applied with public and public-private data repositories (Steger-Hartmann and Pognan, 2018). Pharmaceutical companies' decision-making processes increasingly rely on these data repositories to help support and complement internal research programs. EFPIA (European Federation of Pharmaceutical Industries and Associations) activities to facilitate data-sharing across companies will equally encourage high-value projects for cooperative data-sharing, which in turn are likely drivers towards greater harmonization of operating protocols and use and re-use of data in support of public health and drug research.
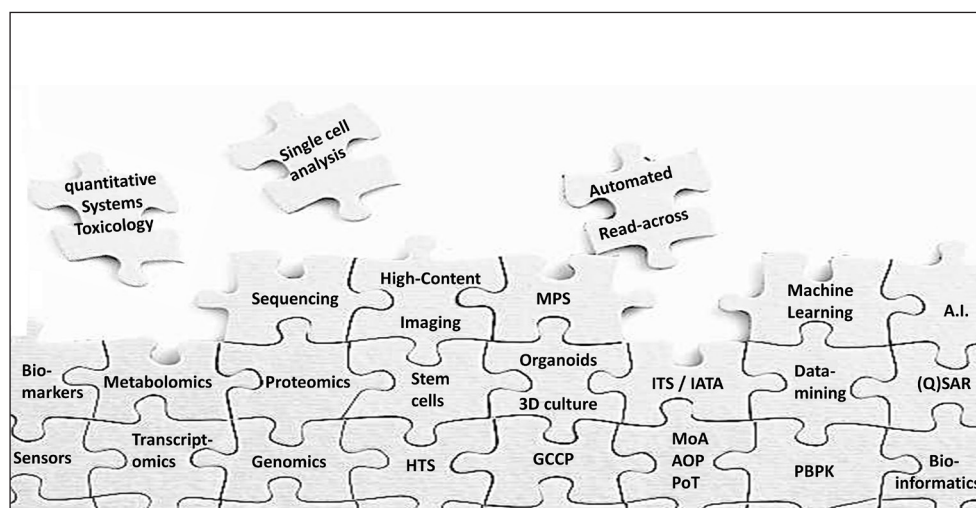
Fig. 4: Enabling technologies for investigative toxicology
AI, artificial intelligence; AOP, adverse outcome pathway; GCCP, Good Cell Culture Practice; HTS, high-throughput screening; IATA, integrated approaches to testing and assessment; ITS, integrated testing strategy; MoA, mode of action; MPS, microphysiological systems; PBPK, physiology-based pharmacokinetic modeling; PoT, pathways of toxicity; (Q)SAR, (quantitative) structure activity relationships

3 New technologies

The development of improved, innovative models for the detection of toxicity of drugs, industrial or consumer chemical products is crucial to efficiently bring new products safely to the market in a cost-effective and timely manner. Figure 4 illustrates some of the modern technologies going into investigative toxicology.

This non-exhaustive list of technologies – especially in combinations – encompasses a strong toolbox for mechanism-based human-relevant investigative toxicology approaches.

3.1 *In silico* tools and modeling

3.1.1 Overview
The prediction of mutagenic activity of new chemical entities (NCEs) based on their structure and potential reactivity towards DNA has been used for some decades, and *in silico* tools are now accepted for regulatory decision-making in the area of genotoxicity of drug candidates and impurities in pharmaceuticals (Amberg et al., 2014). Beside this, a lot of effort has been put into the prediction of organ toxicities, such as DILI, using different computational models (Kotsampasakou et al., 2017; Mulliner et al., 2016a), which achieved accuracies in the range of 70 and 80%.

New perspectives for *in silico*, read-across, and modeling approaches are resulting from the emerging availability of "big data" in toxicology (Hartung, 2016; Clark and Steger-Hartmann, 2018). One opportunity to push investigative toxicologists to embrace the 3Rs principles lies in developing new *in silico* approaches, and also in effectively integrating existing *in silico* tools with *in vitro* technologies, as well as with preclinical and clinical databases (Rovida et al., 2015), conceivably within an AOP-like framework (Tollefsen et al., 2014). For example, the *in silico* prediction of on/off-target liabilities was, in part, addressed

in the Innovative Medicines Initiative (IMI) project, eTOX[5]. Predictive models were built to cover hundreds of clinical safety events linked to drugs and their pharmacological properties (Garcia-Serna et al., 2015). Linking different data sources (toxicity, on-target and off-target, drug metabolism and pharmacokinetics (DMPK)) using computational methods will allow toxicologists to go beyond traditional structural alerts and move towards an understanding of toxicity cascades. This could eventually contribute to AOP development and validation and ultimately to the interpretation of the underlying mechanism(s). The multifactorial origin of drug toxicity can thus be analyzed by combined approaches or network models to identify the causality of a toxic effect, ultimately shedding light on the likely mechanisms by which NCEs generate a safety risk.

3.1.2 Machine learning and artificial intelligence
A number of recent developments in quantitative pharmacology modeling have the potential to further embed these tools into an *in silico* drug development framework, thus contributing to an early assessment of drug candidates regarding the differentiation between on-target or off-target related liabilities (Murphy, 2011). The standardization and automation of the development of quantitative pharmacology models, together with their validation and reporting, will facilitate the acceptance and uptake of QSARs (Kausar and Falcao, 2018).

As a compliment to the traditional QSAR models relating a chemical to a biological property, molecular docking models allow the rapid calculation of the binding potential of drugs to a target protein. Studies assessing the performance of commonly used molecular docking programs (e.g., Glide, GOLD, FlexX, eHiTS, PDBbind database) indicate that these programs can perform precise protein conformation, but their scoring functions are still too inaccurate for a reliable prediction across a variety of targets (Plewczynski and Klingström, 2011).

---

[5] http://www.etoxproject.eu
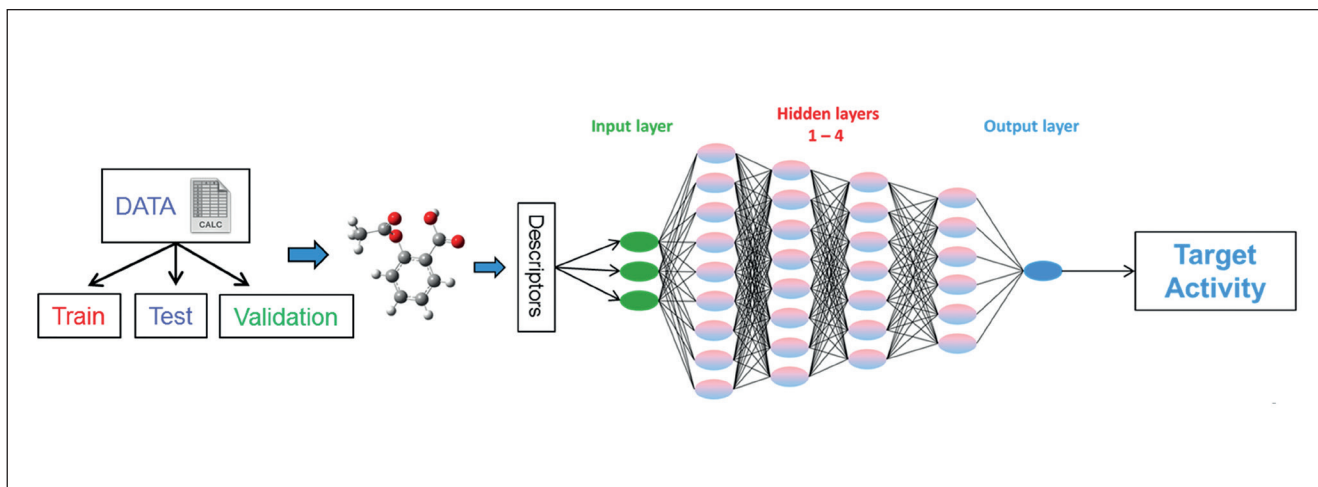
Fig. 5: Enabling technologies for investigative toxicology: Application of machine learning / artificial intelligence for the prediction of target activity

There have been significant advances in machine or deep learning technology in recent years. Although deep learning approaches have been shown to yield accurate predictions (Mayr et al., 2016), they require large, costly datasets. When it is practical to generate a relatively small dataset, researchers often seek to test a diverse set of compounds in their assay. Because of the complexity of compound space as well as the assay results within that space, diversity selection of compounds does not always yield an optimally predictive model. One solution to this problem is the use of transfer learning. With this approach, data from biologically similar assays can be used to predict one another. This allows for the effective expansion of chemical space for toxicities for which data are more limited (Kangas et al., 2014). The second solution to the problem of generating data for learning predictive models is the use of active machine learning (Murphy, 2011). In essence, a machine learning algorithm can be used to identify which tests will yield the most informative data. By focusing experimentation primarily on the informative experiments that yield the best data, far fewer experiments are needed to learn an accurate predictive model. In practice, these active machine learning approaches can significantly reduce *in vitro* and *in vivo* experimentation, while also increasing prediction accuracy, and they are not strictly limited in application to investigative toxicology.

The power of machine-learning approaches in drug discovery lies in its integration with network modeling (Fig. 5). A well-curated, comprehensive molecular interaction network can reveal causes and effects of protein interactions in signaling and metabolic pathways, thus allowing network-based screening to systematically identify target proteins of a drug and their impact (Hsin et al., 2013).

3.1.3 AOPs and their role in network models
To date, AOPs have been applied in the safety assessment of chemicals but less so in drug discovery. AOPs serve as a mostly linear concept to identify measurable key events (KEs). Howev-er, AOPs, especially quantitative AOPs, may also prove beneficial as a framework to build *in silico* tools and *in vitro* testing batteries for drug discovery (Hartung, 2017b). AOP networks based on shared KEs are in active development (Knapen et al., 2018). Systems biology models, such as neural networks, have been a focus in drug development but require comprehensive, complex tools for their quantification (Hartung et al., 2012, 2017). Although not formally applied thus far, toxicokinetic-toxicodynamic (TKTD) modelling (Tsaioun et al., 2016; Kretschmann et al., 2012) may prove to be a useful tool to quantify KEs. These models simulate processes leading to toxicity in organisms over time, where (a) uptake and elimination rate constants for an NCE/NBE in an organism are determined to estimate the time course of a toxicant at a target (e.g., molecular initiating events (MIEs)) and (b) damage accrual and recovery rate constants for an effect across biological scales are determined to estimate the time course of an effect.

3.1.4 QIVIVE and PBPK/PD
A quantitative understanding of the progression of biological events from MIEs to adverse outcomes allows us to derive tissue-specific points of departure (PoD) from organ-specific *in vitro* assays assessing perturbations of relevant KEs. The PoD is used to mark the beginning of extrapolation to determine the risk associated with expected human exposures. Quantitative AOPs will help answer what level of *in vitro* perturbation should be used as a PoD for quantitative *in vitro* to *in vivo* extrapolations (QIVIVE) (Hartung, 2017a). An understanding of the activity threshold is required that pushes the toxicity pathway onward from one molecular event in this pathway to the next and the internal dose of the drug or toxicant that affects the probability and severity of an event perturbation.

PBPK modelling is becoming indispensable for QIVIVE (Basketter et al., 2012; Leist et al., 2014). Specifically, reverse dosimetry PBPK is being used to estimate human exposures that lead to concentration-time profiles that are equivalent to sufficiently

active concentration-time profiles *in vitro* (Louisse et al., 2017). Recent efforts in the US EPA ToxCast program[6] illustrate the integration of *in vitro* activity concentrations with reverse dosimetry PBPK for risk assessment. *In vitro* determined hepatic clearance and plasma protein binding parameterized a TK model to predict the chemical steady-state concentrations (C_{ss}) in plasma resulting from repeated daily exposure. Reverse dosimetry PBPK tools were subsequently used to estimate human equivalent doses (in mg/kg/day) required to achieve blood C_{ss} levels identical to *in vitro* bioactive concentrations.

3.1.5 Big data

Besides the omics applications and the concomitant pathway analysis, future use of big data in safety science will encompass two fields, early compound (drug candidate) assessment and translation concordance analysis. On the one hand, mining of large preclinical data sets will result in automated read-across procedures (Hartung, 2016), which will enable the assessment of new chemical structures, including structural moieties, for their potential toxicity liabilities. Such tools will enable medicinal chemists to guide their hit-to-lead search, not only for criteria of pharmacophore, drug metabolism and pharmacokinetics (DMPK), and physico-chemical properties, but also for specific safety aspects, also termed "green toxicology" (Maertens et al., 2014; Crawford et al., 2017; Maertens and Hartung, 2018). An example of how such read-across approaches might be applied for optimizing drug candidate selection to reduce toxicity liabilities in early phases has recently been published (Steger-Hartmann and Pognan, 2018).

The other area of interest is the automated analysis of animal-human translation or concordance. Questions such as, "*Tell me how an n-fold decrease in white blood cell count in species x at dose y corresponds with effects in humans?*" with all subsequent ramifications (e.g. "*Can results be grouped according to preclinical species, pharmacology, mode of action, …?*") or "*What is the most sensitive preclinical species for a specific organ toxicity?*" can be approached by analyses of big data sets (Clark and Steger-Hartmann, 2018).

Big data analyses however require:
− accessibility of large preclinical and clinical data sets, while safeguarding aspects of intellectual property and personal data protection
− automated procedures for data curation
− integration of controlled vocabularies and ontologies to enable cross-analyzing data
− quality control of data by scientific experts

Such efforts can only be achieved via consortia approaches and should be run in parallel to data sharing guidelines and principles. Examples of such initiatives are DruSafe[7] (Monticello, 2015), eTransafe[8], or the initiative to make data "fair" (= "*findable, accessible, interoperable, reusable*"[9]).

3.2 Safety biomarkers

3.2.1 Application and classification of safety biomarkers

Safety biomarkers for use in investigative toxicology fall largely into the category of "response biomarkers" (Amur et al., 2015). A drug liability identified early in discovery, which shall be monitored and ideally de-risked during non-clinical and clinical research, requires robust and reliable safety biomarkers that are of translational relevance to humans. The same biomarkers would conceivably also support monitoring of patient populations and positively impact therapeutic safety margins. As biomarkers provide valuable information on drug safety, they are increasingly integrated as part of drug discovery and non-clinical development. Establishing the use of novel safety biomarkers with target organ specificity and mechanistic insight for use in non-clinical (Blaauboer et al., 2012) and clinical studies nevertheless remains challenging.

Biomarkers include messenger and micro RNAs, proteins, metabolites, clinical chemistry (Brooks et al., 2017) as single end-point measurements or as multiplexed processes in microarray and microfluidic platforms. Whatever the biomarker selected, preclinical confirmation on the comparative molecular biology, translational relevance of the mechanism of toxicity, target organ, and time course with known histopathology in humans is required for later qualification (Matheis et al., 2011). The classification of biomarkers as exploratory, probably valid biomarkers and valid biomarkers defines how biomarkers are applied in research and development (Chau et al., 2008). Classification of the increasing numbers of qualified biomarkers helps define how emerging and future biomarkers can be used to support decision-making and their acceptance by regulatory authorities (Edwards et al., 2016).

3.2.2 Safety biomarkers for the three key target organs

The development of safety biomarkers for the organs that contribute to the highest attrition, i.e., heart, liver, and CNS (central nervous system), has been pivotal owing to both the severity and occurrence of these target organ toxicities across many classes of drugs (Marrer and Dieterle, 2010). The progress towards the development of biomarkers among these three target-organ toxicities is highlighted below.

Heart (cardiovascular toxicity)
Cardiovascular toxicities accounting for ADRs, drug attrition, and withdrawal relate to all components of the cardiovascular system (Laverty et al., 2011; Valentin et al., 2010) and can be broadly categorized into i) structural damage, ii) functional deficits with or without histopathological correlates, and iii) altered cell or tissue homeostasis in the absence of obvious structural or functional deficits (Wallace et al., 2004). The diversity of

---

ADRs necessitates a range of biomarkers to detect, predict, and monitor ADRs in non-clinical and clinical testing. Biomarkers of hemodynamic effects include monitoring of blood pressure, heart rate, and ejection fraction using semi-invasive approaches. Cardiac electrophysiological effects such as QTc prolongation or shortening, QRS widening, PR prolongation, arrhythmias such as torsade de pointes and ventricular fibrillation are detectable via an electrocardiogram. For some of these endpoints, predictive *in vitro* screens are well established, there is a good relationship between free plasma concentration associated with significant QT prolongation and torsade de pointes in the clinic and *in vitro* Ikr IC_{50} values (Webster et al., 2002). More recently, safety testing in stem cell derived cardiomyocytes has been suggested as part of a new integrated risk assessment of pro-arrhythmic liability (Sager et al., 2014). Degenerative or inflammatory lesions can be monitored via body fluid sampling and measurement of N-terminal pro-brain natriuretic peptide (NT-proBNP), miRNAs, creatine kinase (CK), aspartate aminotransferase (AST), troponin, and pro-atrial natriuretic peptide (pro-ANP) / brain natriuretic peptide (BNP) in both non-clinical species and humans. Although, numerous biomarkers of drug-induced cardiotoxicity have been proposed and are being used, some lack sensitivity and/or specificity, therefore the quest for mechanism-based cardiotoxicity biomarkers is continuing.

Liver (hepatotoxicity)
DILI represents one of the most significant ADRs. Attrition of promising drug candidates due to DILI occurs in preclinical and clinical development (Clarke et al., 2016; Pognan, 2018). DILI is classified as either intrinsic, with clear dose-dependent hepatocellular injury (Corsini et al., 2012), or idiosyncratic with low incidence rates in humans, which cannot be predicted with current *in vitro* and *in vivo* tests. The phenotypic assessment of DILI in patients relies on measurement of alanine aminotransferase (ALT), AST, alkaline phosphatase (ALP), and bilirubin (BIL). Despite wide acceptance, ALT, AST, and ALP are not specific measures of liver injury and detection of BIL occurs after extensive liver injury has occurred (Church et al., 2018). In non-clinical testing, detection of DILI relies substantially on histology (Weaver et al., 2017).

The use of ALP, ALT, and BIL as biomarkers is insufficient for the detection of human DILI and predicting outcome. Efforts to improve upon these liver safety biomarkers have yielded promising, novel biomarkers with additional mechanistic information: High mobility group box 1 (HMGB1) for detection of necrosis (Scaffidi et al., 2002) and its acetylated form in immune DILI (Lu et al., 2012). The value of these and other novel biomarkers, such as keratin-18 and miR122, are presented in detail elsewhere (Antoine et al., 2013; Clarke et al., 2016; Ward et al., 2014). These novel biomarkers are best defined as "response biomarkers" and further work is encouraged to extend knowledge towards their translational and predictive value as qualified biomarkers of DILI (Matheis et al., 2011). The prospect of translationally relevant safety biomarkers for use in the prognosis of DILI outcomes in patients is encouraging (Ozer et al., 2008).

CNS (Neurotoxicity)
There is a need for more sensitive and specific biomarkers that can help diagnose and predict neurotoxicity that is relevant across animal models and can be translated to the clinic (Schmidt et al., 2017). Some traditional functional biomarkers with established non-clinical to clinical translational value, e.g., electroencephalogram, electroretinogram, and brainstem auditory evoked potential, can be used. Fluid-based biomarkers, such as miRNAs, F2-isoprostanes, translocator protein, glial fibrillary acidic protein, ubiquitin C-terminal hydrolase L1, myelin basic protein, microtubule-associated protein-2, and total tau, hold great potential due to the relative ease of sampling. However, some of these biomarkers (such as those in the cerebrospinal fluid) require invasive sampling or are specific to one disease such as Alzheimer's or Parkinson's disease, while others require further validation. In addition, neuroimaging methodologies may also provide potential biomarkers and, coupled with functional, genetic- and protein-based biomarker assessments, offer an exciting way forward to predict, detect, and monitor drug-induced neurotoxicity.

3.2.3 Future perspectives
Continued efforts towards the discovery and characterization of novel, sensitive and relevant biomarkers to effectively bridge *in vitro* to preclinical to clinical testing would strengthen our ability to predict, detect, and monitor drug-induced organ injuries (Park et al., 2000). The principal challenges ahead include the identification and qualification of these biomarkers for use not only as "response biomarkers" but as predictive of ADR outcomes and prognosis.

3.3 Novel cell models
Generating physiologically relevant models is a promising approach to improving our ability to detect and predict drug-induced toxicity, as well as to unravel specific mechanisms of toxicity. Therefore, there is an increasing desire to move away from the use of cell lines that form part of screening cascades within the drug discovery process and towards primary cells with their known limitations (e.g., limited source, variability, etc.) (Eskes et al., 2017; Pamies et al., 2017, 2018; Coecke et al., 2007).

A robust, reproducible, and relatively "unlimited" source of cells with defined phenotypes and genotypes would greatly benefit the field of toxicity testing and assist in standardizing early investigational toxicological research (Pamies and Hartung, 2017). Differentiation of various types of human stem cells into the desired somatic cells might be a solution.

Moreover, introducing further complexity by culturing cells in 3D, microphysiological and organoid model systems is an approach that is gaining ground within the investigative toxicology community (Alépée et al., 2014). The idea is that such 3D and organoid models display more physiologically relevant attributes, including cell polarization, cell-cell or cell-microenvironment interactions (Anton et al., 2015; Duval et al., 2017; Retting et al., 2018) that are important drivers of tissue differentiation and function. Microfluidic and tissue printing techniques have been used to increase complexity of tissue models by add-

ing aspects such as co-culture of multiple cell types, flexibility for compartmentalization and higher-order tissue architecture, flow, gradient formation, and mechanical strain. This increase in complexity ultimately leads to improved functionality and has been demonstrated for various types of approaches in hepatocyte models including 3D spheroids (Messner et al., 2013; Bell et al., 2016; Proctor et al., 2017), 3D printed systems (Retting et al., 2018), organoids (Huch et al., 2015), and MPS systems (Huh et al., 2010; Vernetti et al., 2016). Each of these systems has added value for biological relevance, although their routine implementation in toxicological testing remains to be established.

3.3.1 Stem cell models

Cell lines and primary cells have long been the main source of cells in cell-based experiments. Cell lines provide a relatively stable and continuous source of biological material, but are highly variable with regard to the level that they maintain the features associated with their tissue of origin. Cells that are freshly isolated from primary tissue are generally considered a gold standard for their physiological relevance. The time span over which these properties are maintained, however, is typically limited. Moreover, physiological properties may disappear under certain storage conditions, logistics around these primary cells are cumbersome, and the quality of isolated cells can be highly variable.

Induced pluripotent stem cells (iPSCs) promise to be a renewable source of cells and could potentially provide large numbers of cells with well-characterized physiological properties and with genotypes that correspond to specific individuals. Today, various cell types are used for iPSC production, e.g., germ lines, liver cells, skin cells, and lymphocytes (Takahashi and Yamanaka, 2006; Yu et al., 2007; Gadue and Cotsarelis, 2008; Okita et al., 2007; Loh et al., 2009; Aoi et al., 2008). Various protocols that guide iPSC differentiation towards specific cell lineages have been published. Cell types including endothelial cells and smooth muscle cells, neuronal cells, cardiomyocytes, and hepatocyte-like cells can be differentiated with specific supplements and growth factors (Patsch et al., 2015; Hu et al., 2011; Mauritz et al., 2008). Phenotypes of various diseases, such as familial hypercholesterolemia, Wilson's disease and alpine-1-antitrypsin disease have been generated from iPSC derived hepatic cells, which could be used as cellular disease models (Cayo et al., 2012). Fundamental research on these cells could help our understanding of various disease types leading to the development of novel drugs.

Best established for toxicity testing are cardiomyocytes (Millard et al., 2018) and neurons (Wevers et al., 2016). The quality of hepatocyte differentiation is progressing. However, expression levels of xenobiotic metabolism genes in iPSCs are still not equal to those found in organs or freshly isolated primary cells. Other obstacles continue to impede progress towards using these cells for *in vitro* toxicology (variability in lines, incomplete programming within cell populations, uncharacteristic response to prototype toxicants, etc.).

Despite these limitations, iPSC derived cells are now suggested for use in toxicological screening and may provide an understanding of an individual patient's ADRs (van Hasselt and Iyengar,

2017), where the most advanced example is probably the CIPA (comprehensive *in vitro* pro-arrhythmic assay) initiative of the US FDA (Wallis et al., 2018).

Fully functional organ-specific cells derived from iPSCs will become a valuable tool for drug development or evaluation of the contribution of genetic variation to variable responses. Moreover, the technology provides a unique opportunity to distinguish between gender, ethnic background, and potentially even disease background. The field is progressing rapidly, with varying levels of limitations still remaining. However, even though the introduction of induced pluripotent or embryonic stem cells for toxicological and pharmacological studies seems inevitable, efforts towards standardization, validation, and regulation are still necessary in order to make them a widely accepted option for toxicological and pharmacological studies.
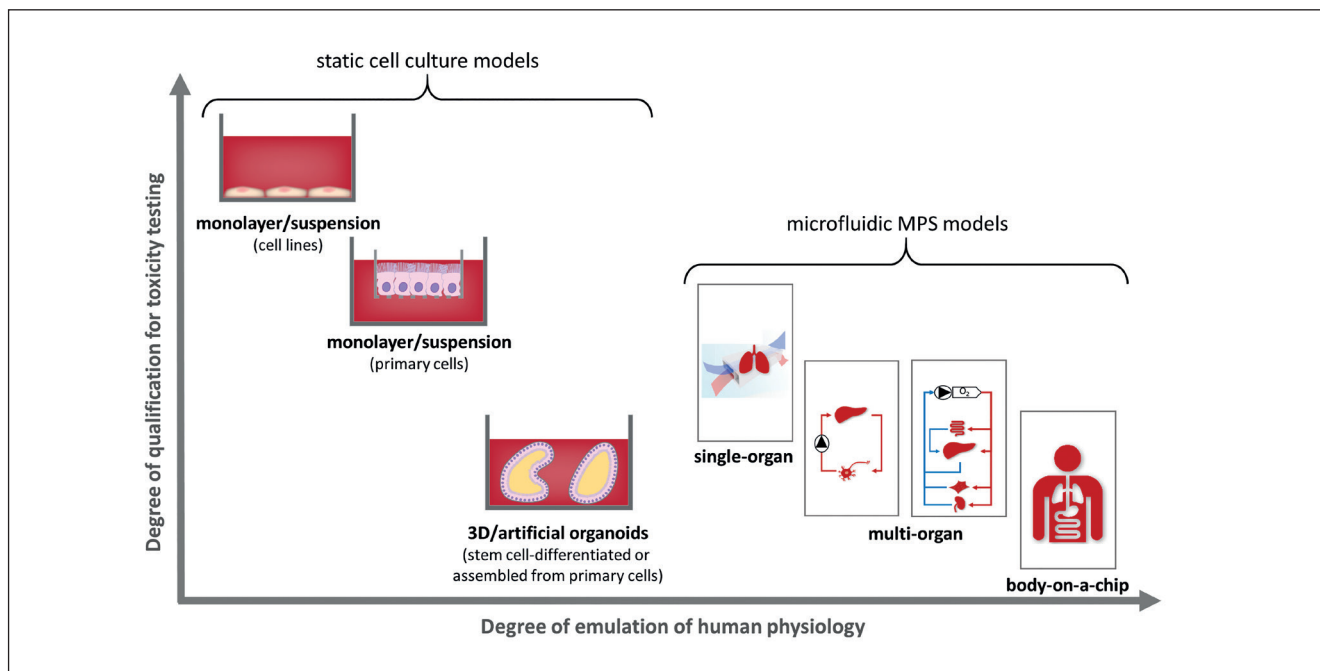
3.3.2 Organoids

Organoids are a recent paradigm in tissue culture, with culture conditions assuring preservation of the (adult) stem cell niche, while proliferation and differentiation to the essential cellular subtypes of a specific organ still occur. For example, intestinal organoids accurately predict therapy response in cystic fibrosis and were used to establish living biobanks of tumor tissue that was genetically stable over time (Artegiani and Clevers, 2018). Whereas these intestinal organoids were expanded from primary human cells or human stem cells, pluripotent stem cells have also been used to generate organoids with impressively realistic *in vivo*-like microanatomy for the brain (Lancaster and Knoblich, 2014) and kidney (Takasato et al., 2015; Freedman et al., 2015). As of today, many organoid systems have been developed, including liver (Huch et al., 2013; Takebe et al., 2013), intestine (Spence et al., 2011), thyroid (Antonica et al., 2012), pancreas (Greggio et al., 2013), lung (Lee et al., 2014), and retina function (Nakano et al., 2012).

Whereas most of these systems are currently being used mainly in the context of basic developmental and stem cell research or disease modeling (Artegiani and Clevers, 2018), it is evident that these technologies are poised to play a role in the field of toxicology. What is required is a full and thorough evaluation of physiological and pharmacological characteristics of these organotypic models alongside human tissues to establish whether such models are "fit-for-purpose", i.e., improve the prediction of target organ toxicities (Carragher et al., 2018). The utility of these *in vitro* models can be enhanced by understanding the AOPs/PoT covered by the models of interest (Hartung and McBride, 2011; Kleensang, 2014; Hartung, 2017b).

There are obvious hurdles to overcome (Lancaster and Knoblich, 2014; Carragher et al., 2018):
− artificial organoids currently mimic some, but not all, of the physiological functions of the respective human organs (Materne et al., 2013),
− they lack physiological vasculature and, consequently, whole blood perfusion, which is essential to nutrient supply, waste transport, and several other physiological processes, including creating a dynamic microenvironment,

Fig. 6: The current cell model landscape
Traditional systems for evaluation of toxicity include cell line- and primary cell-based models. Developing technologies such as 3D organoids, bio-printed tissues, and single- and multi-organ MPS will result in models with greater biological relevance, for which full validation and routine implementation remain to be established. Human body-on-chips are still at an early research stage of development.

− they lack key cell types, such as immune cells (resident or circulating), and neuronal innervation,
− primary cell-derived artificial organoids face the shortage of human cell supply,
− they lack *in vivo* relevant cellular architecture and cell-cell interactions,
− they lack mechanical forces,
− stem cell-derived artificial organoids replicate only the early stages of organ development, remaining "fetal-like" owing to lack of essential cues for final differentiation.

Therefore, organoid systems still represent a trade-off between throughput and physiological relevance, and in many cases, the effects of a drug depend on factors such as metabolic competence or tissue specific distribution and interaction that cannot be achieved within single organoids.

3.3.3 Microphysiological systems (MPS)

Within the last 5 to 10 years, advances in microfluidic and micro-engineered technology has enabled the development of organ-on-chip models or MPS (Marx et al., 2016; Smirnova et al., 2018; Esch et al., 2015). By applying engineering principles, models can now be created that accurately represent the cellular microenvironment of an organ (Bhatia and Ingber, 2014). In doing so, cells theoretically retain their physiological phenotype and respond in comparable ways to their *in vivo* counterparts. Application of these models within the investigative toxicology and safety assessment process has been reviewed recently (Ewart et al., 2018). MPS are cell source agnostic and support various tissue architectures from monolayer to 3D and artificial organoids. Because of their specialized microenvironment, they have recently been shown to be a tool that can enhance stem cell maturation (Sances et al., 2018; Ronaldson-Bouchard et al., 2018). While typically of limited throughput, higher throughput systems have been developed and applied for toxicity testing of 3D gut tubules and iPSC-derived neuronal models (Trietsch et al., 2017; Wevers et al., 2016). Multiple organs can be combined on one chip (Wagner et al., 2013; Skardal et al., 2017) to investigate the mechanisms that drive organ toxicity at organ cross talk. Finally, the impact of biological feedback loops such as the insulin-glucose regulation of liver performance can be studied using the respective organ combinations (Bauer et al., 2017).

Organs-on-chips have also been developed to evaluate drug-induced toxicity (Esch et al., 2015). Organ-specific examples are the heart-on-chip (Zhang et al., 2015) and the lung-on-chip model developed by Huh and coworkers (Huh et al., 2010). A 3D bio-printed, cell-based mammalian skeletal muscle strip was successfully generated that is able to exert muscular force (Cvetkovic et al., 2014). In addition, three-dimensional bio-printed human models of liver, kidney proximal tubule, and intestinal tissue have been described for use in modeling native physiology and compound-induced toxicity (Nguyen et al., 2016; King et al., 2017; Madden et al., 2018). The progress in MPS hepatocyte culture systems (including non-parenchymal co-culture and bio-physical constraints such as oxygen tension) has led to additional improvement of tissue and organ level function (Vernetti et al., 2016, 2017; Lee-Montiel et al., 2017). 3D liver and neuronal spheroids
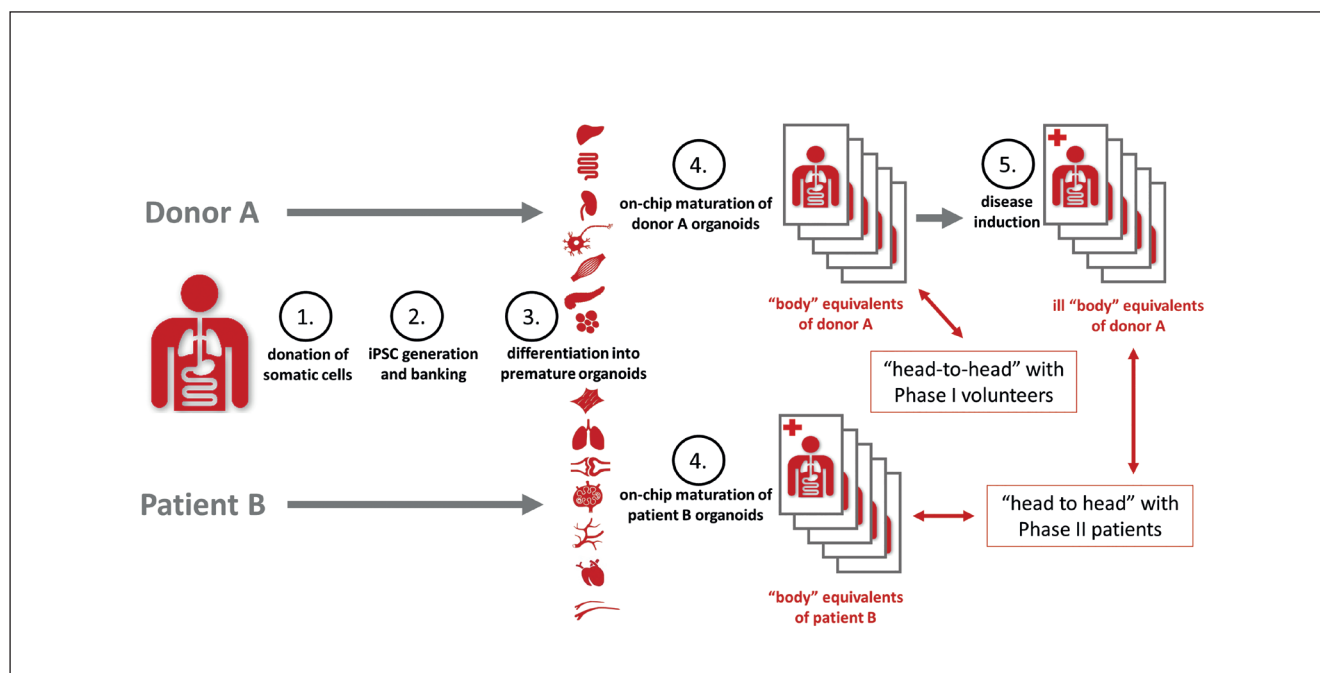
Fig. 7: Sketching a roadmap towards "clinical trials" on a chip
Assembly of minute personalized "body" equivalents on a chip derived from cells of individual healthy donors or patients under ethically acceptable conditions is still a hypothetical multistep procedure exploring different aforementioned technologies.

have successfully been co-cultured on MPS for long term toxicity testing (Materne et al., 2015). Human intestinal organoids, liver spheroids, human skin biopsies, and monolayer proximal tubular cell barriers have been combined on a four-organ MPS platform for evaluation of systemic long term toxicity (Maschmeyer et al., 2015). Hepatic and cardiac cell types have been differentiated from iPSCs using MPS (Giobbe et al., 2015). It has been hypothesized that exposure of *in vitro* assembled premature iPSC-derived organoids to the physiological cues of an MPS, such as perfusion, shear stress, electrical stimulation, and organoid cross talk in interconnected arrangements, might constitute the missing step for their final and complete *in vitro* differentiation. First progress has been made to vascularize microfluidic systems (Schimek et al., 2013; Van Duinen et al., 2017).

Figure 6 schematically illustrates the current cell model landscape and future perspectives discussed in this section.

3.3.4 Envisioned progress of *in vitro* models
The described progress in human iPSC generation at a robust large scale, their differentiation into a broad variety of premature organ-specific somatic cell based artificial organoids, and the steady increase in the number of organ equivalents on MPS platforms has created an historically unique opportunity for the introduction of humanized models in safety assessment (Miller and Shuler, 2016; Xiao et al., 2017; Edington et al., 2018). The combination of these three approaches may well lead to the establishment of personalized minute equivalents of a healthy donor- or a patient-on-a-chip. Figure 7 summarizes the long-term vision of

using such MPS-based personalized patient equivalents for studies to mimic Phase 1 and Phase 2 clinical trials.

Eventual success along this path may enable us to perform individualized studies mimicking clinical trials of a particular donor using statistically relevant numbers of almost identical replicates of donor or patient "bodies" on chips. This is in line with the current use of inbred, genetically identical laboratory animals in preclinical evaluation, with the difference that such miniaturized "bodies"-on-a-chip are of personalized human origin. Furthermore, it might allow head-to-head analysis of the outcome of the real donor or patient study with its body-on-a-chip counterparts. Finally, the use of "body" equivalents from donors and patients of different gender, ethnic groups, and genetic backgrounds will allow an evaluation of the impact of such parameters on the safety and efficacy of an NCE/NBE in the preclinical setting, which illustrates the high potential of such tools for the drug development cycle (Marx et al., 2016).

3.4 Imaging technologies
The past decades have seen enormous development and integration of high-content imaging in investigative toxicology departments (van Vliet et al., 2014; Uteng et al., 2014). The development of a variety of small molecule fluorescent probes allows the detection of numerous biochemical perturbations and live/dead endpoint measurements. For example, probes have been used to follow the accumulation of fatty acids in cells leading to steatosis (Germano et al., 2015), one of the critical endpoints of DILI. Fluorescent bile acids have been applied to visualize the accumu-

lation of bile acids as a consequence of bile acid transport inhibition and may contribute to identifying compounds with a liability for drug-induced cholestasis (Germano et al., 2015). Likewise, fluorescent probes have been used for assessment of phospholipidosis (Morelli et al., 2006), oxidative stress, and mitochondrial membrane potential (Billis et al., 2014). For the pharmaceutical industry, this high-content imaging approach has become an essential tool within the field of predictive toxicology with the aim to design and prioritize drug candidates with a superior safety profile (Persson and Hornberg, 2016). While the technologies have primarily used 2D cell systems (either cell lines or primary cells; Pampaloni et al., 2007), the challenge for the future is to capture this in advanced 3D cell models and allow sufficient resolution for single-cell-based quantification of probe activity. Novel high-content imaging machines still have the limitation that they cannot capture the fluorescence of cells in the center of multicellular 3D spheroids. Challenges for the future are to bring light-sheet microscopy to the level of high-content screening and integrate this in screening labs (Joshi and Lee, 2015) to allow the detailed analysis of biochemical changes in complex MPS. Novel approaches involve phenotypic screening of cell morphologies, allowing the quantification of hundreds of (related) parameters in parallel (Joshi and Lee, 2015; Leary et al., 2018). Further challenges include the integration of other mechanistic biomarkers that would represent key events of AOPs into high-content imaging strategies.

Further advances in molecular imaging capability and deployment are also continuing through the development of label-free bio-imaging of tissues and cells (and potentially single cells and organelles) using mass spectrometry (MS) based approaches (Passarelli and Ewing, 2013).

3.5 Omics profiling

Omics technologies, which have gained prominence and relevance over the last few years, can be divided into four parts that focus on different steps in the generation of bioactive molecules: genomics, transcriptomics, proteomics, and metabolomics. Systematic studies on transcriptome analysis within large consortia and industry settings have helped establish extensive datasets of drug-induced transcriptome profiles in different target organs as well as *in vitro* in cells. This is exemplified by the TG-GATEs[10] and DrugMatrix[11] datasets that are available in the public domain (Igarashi et al., 2015; Ganter et al., 2005). The initial hope for toxicogenomics as the solution for the ultimate prediction of target organ toxicity has not been completely fulfilled as the technology and the diversity of transcriptional profiles has turned out to be more complex than anticipated (Pognan, 2007). While toxicogenomics-based predicted gene profiles have been established for several adverse outcomes, including genotoxicity testing (Ellinger-Ziegelbauer et al., 2009), a widespread generic application in drug safety testing has not been implemented. Generally, toxicogenomics is applied to support the mechanistic understanding of identified target organ toxicities. With RNA sequencing being the current major tool in

transcriptome analysis, and sequencing becoming much cheaper, new applications of toxicogenomics are emerging and may require new attention and funding. In addition, improved bioinformatics tools that integrate large omics datasets into co-regulated gene networks, allow quantitative analysis of the association between such gene networks and adverse outcomes (Stiehl et al., 2017; Sutherland et al., 2018). There is a rapid development of sequencing strategies, where chromatin-immuno-precipitation (ChIP) sequencing will contribute to a further refinement of the transcription factors that drive these transcriptional networks in different target tissues. These complementary sequencing approaches should ultimately define the quantitative relationships between both safe and adverse ranges of pathway activation that will determine the safety margins. The IMI TransQST project will contribute to these quantitative systems toxicology evaluations[12].

In concert with transcriptome analysis, sensitive proteomics platforms have also evolved that have allowed the analysis of cell and tissue proteomes under healthy and disease settings and after drug exposure (Cox and Mann, 2011). In particular, phosphoproteomics has allowed the assessment of early signals of cell signaling activation in relation to drug exposure (Pines et al., 2011). To date (phospho)-proteomics is not yet a common tool in drug safety assessment and investigative toxicology. However, integration of proteomics with transcriptomics has helped to gain a more precise understanding of drug action (Puigvert et al., 2013). Recent integration of biology information with proteomics has allowed the identification of drug targets of a large panel of kinase inhibitors (Klaeger et al., 2017). The integration of activity-based target profiling in investigative toxicology with the help of proteomics will further clarify the spectrum of off-targets of candidate drugs and contribute to an improved drug safety prediction.

Metabolomics is defined as analysis (identification and quantification) of active metabolites, including carbohydrates, lipids, and more complex bioactive molecules, such as hormones. Its role in toxicology is increasing (Bouhifd et al., 2013; Ramirez et al., 2013), fueled also by increasing quality assurance demands (Bouhifd et al., 2015). The metabolome can be determined in human and animal matrices (e.g., blood, plasma, urine, or sweat) with the focus on the entire body but also on organ-specific toxicity. Moreover, organ specific metabolomes for *in vitro* systems have been reported (Ramirez et al., 2013).

In parallel, targeted approaches for metabolomics have been developed, with increased sample throughput, enhanced analytical robustness, and facilitated data analyses. Targeted metabolomics carries the promise of a high translational potential for clinical studies. An example for targeted metabolomics is the application of multiplexed LC (liquid chromatography) MS/MS methods for bile acid analysis (both unconjugated and conjugated) for the assessment of the cholestatic or steatotic potential of drug candidates (Schadt et al., 2016). While the metabolome analysis of plasma and urine requires animal testing, it is recommended to consider the 3Rs strategy (focus on reduction) and therefore include omics technology in animal studies. Hence

---

[10] http://toxico.nibio.go.jp/english/index.html

[11] https://ntp.niehs.nih.gov/drugmatrix/index.html

[12] http://transqst.org/

in vitro metabolomics is an important step towards avoiding animal studies and therefore supporting the 3Rs strategy (focusing on replacement).

3.6 Intracellular sensors

Perturbations of normal cell physiology leading to perturbations of biology on an organ or body level culminate in adverse outcome. There is a limited set of cellular perturbations that will drive this adversity. The establishment of AOPs and AOP networks, with the help of omics and cell biology, has already defined some critical pathways related to toxicity. This involves both biochemical perturbations as well as cellular disturbances that drive cell signaling and onset of adaptive rescue programs, or alternatively the onset of pathways that drive cellular demise. Cell biologists and toxicologists have taken advantage of this information and integrated biomarker genes tagged with fluorescent proteins that represent various adaptive cellular stress response pathways. Bacterial artificial chromosome genome editing technologies have allowed the green fluorescent protein (GFP) tagged expression of sensors, transcription factors, and downstream target genes of different cellular adaptive stress response pathways, including the Nrf2-mediated oxidative stress pathway, the p53-mediated DNA damage response pathway, NFκB-mediated inflammatory signaling, and the ATF4 and XBP1-based unfolded protein response (Wink et al., 2014). These fluorescent reporter cell systems can be integrated with high-content imaging and allow the dynamic analysis of stress response pathway activation (Wink et al., 2017). Wink and collaborators recently demonstrated the application of a panel of these reporters in the prediction of DILI using a panel of > 120 DILI compounds (Wink et al., 2018). A full coverage of cellular components that drive the adverse responses of drugs would contribute to the toolbox to evaluate drug safety.

Developments in cell and molecular biology have allowed the refinement of fluorescent protein probes based on fluorescent resonance energy transfer (FRET) (Ni et al., 2017). This allows the dynamic imaging of cell signaling activity. The integration of such tools in high-throughput microscopy setups would allow a refined understanding of the balance between cell adaptation and adversity and has been used in the safety assessment of drugs (Shuhendler et al., 2014).

4 Challenges

4.1 Definitions, terminology and the need for ontologies

Toxicology as a subject is important in many industry sectors next to the pharmaceutical area, such as the chemical industry, cosmetics, consumer products, and food industry. Moreover, it plays an important role in environmental health. Since every sector discusses specific needs and issues, the terminology around toxicology varies widely. Therefore, it is not surprising how differently the term "investigative toxicology", together with its related key-

words, is used in each respective context. Even when discussed by toxicologists within the pharmaceutical industry, "investigative toxicology" is used with different meanings. Hence the challenge that emerged at the beginning of the think tank was to find a common understanding of the keywords related to investigative toxicology. The need to harmonize descriptions of the keywords to enable easier discussion among all stakeholders became obvious. This is particularly important in the effort to implement investigative toxicology also into the regulatory process when discussing safety aspects with regulatory authorities.

As aforementioned and much discussed during the workshop, a holistic definition of investigative toxicology in drug development could be "*the complementary discipline to 'regulatory toxicology' that includes all aspects of scientific investigations into drug candidates starting with screening and selection of tolerable compounds (prospective approach) and leading to the mechanistic elucidation of adverse effects observed in preclinical or clinical phases (retrospective approach)*". Whether (or how) this definition can be translated to academia and other industry sectors will be seen in the near future. Such translation is of high importance since both approaches (prospective and retrospective) are currently in a transition phase, triggered by exponentially evolving testing options, e.g., complex human *in vitro* models, gene targeting, or enhanced screening possibilities. For all stakeholders, a common, more precise understanding of needs and objectives of these approaches is necessary to successfully align development efforts in academia and industry. Cooperative advancement is necessary to deliver comprehensive tools that can support all stakeholders according to their objectives.

Furthermore, the terminologies used in the context of investigative toxicology need to be harmonized since different (sub)disciplines may have a different understanding of key terms, including biomarkers, safety assessment, mode of action, off-target/on-target toxicity, hampering the focused integration of new technologies. But, also in the field of regulatory toxicology, ontologies of toxicological observations are not fully harmonized (Hardy et al., 2012a,b). How this can be improved was demonstrated in the ontology discussion within the IMI eTOX[5] project. Histopathology ontology was developed with the intent to standardize histopathology findings, and it has now been made openly available (Ravagli et al., 2017). This is a key enabler, making searches on one of the most important data types in toxicology studies possible and providing a database for new *in silico* and *in vitro* models[13].

Further aspects that were found to be critical and challenging for an effective development of new tools (e.g., models, methods) include the definition of performance standards that lead to successful proof of concept studies. These need to be clearly defined to demonstrate the value of an evolving complex cell model or entire testing method for the respective objective. In turn, proof of concept studies may form the basis of the critical aspect of validation of a test method. The best case scenario is a formal validation that would open the door for regulatory use needs to be clearly defined to not only deliver reproducible results but to ensure the

---

[13] http://etransafe.eu/histopathology-ontology/

results are relevant and informative for hazard identification and risk assessment (Amur et al., 2015).

Many keywords that have emerged in the field in recent years are barely defined across stakeholders/sectors. For example, translational biomarkers and AOPs are frequently used in innovative publications (Antoine et al., 2013). However, at least in the case of safety biomarkers, the underlying understanding is still diverse, ranging from very specific (e.g., FDA approved kidney toxicity marker (Brott et al., 2014)) to broad use (e.g., gene expression pattern to define toxicity pathway (Ferrario et al., 2014)). Also, the OECD-introduced AOP concept describes the dissection of molecular toxicity pathways and defines key molecular events (OECD, 2017). But to make it a useful concept for risk assessment, hence allowing differentiation between adaptation and adversity across the different sectors, a common understanding is needed that finally leads to quantitative read-outs or thresholds.

Taken together, the think tank discussion, although initiated by one industry sector only, clearly showed the need to harmonize the use of key terms in the field of investigative toxicology. A common understanding of the relevant vocabulary was stated to be the basis to drive this discipline.

4.2 Study design
Before starting an investigative study or incorporating the investigative part into a regulatory GLP toxicology study, the expectation and the conclusions that might be drawn from the results should be anticipated, clearly described in a study plan, and scientifically justified. This places more weight on the scientific justification of the experimental design and may obviate the need for a full validation or description of the methods in regulatory guidelines. In case a model is more advanced, performance of multi-center studies to prove the robustness of the method is a key requirement. Thus, the community of investigative toxicologists together with the developers of the new models shall enable the pre-validation of these models.

For stand-alone *in vitro* studies, a fit-for-purpose validation (see Section 5) including scientifically justified negative and positive controls is key, whereas for experimental biomarkers, histopathological correlates may be sufficient. In the field of investigative toxicology, very often an iterative investigation cycle with multiple steps leads to the generation of a hypothesis and then to the generation of mechanistic data that support the mechanism of toxicity.

Another important aspect regarding the design of a preclinical safety study is the question whether to use disease models or not. Although disease models are not often used in animal toxicity studies, the underlying disease may be important to identify the most critical key events, as well as thresholds, leading to an ADR (Morgan et al., 2013). The technologies for human *in vitro* disease models as described in Section 2.1.2 as well as animal disease models humanized by gene editing technologies (see Section 2.2.3) will provide new translational approaches for the identification of the mentioned critical key events. This idea, combined with the aspect of human diversity, may eventually lead to the development of individual patient-derived cell models that may be included in a clinical trial of the respective patient. Whether these cell models still need to serve as organ toxicity models is debatable, since the concept of AOPs is not restricted to individual organs and in the best case an appropriate cell model may depict the fundamental toxicity mechanism in the body across different organs. However, the most important challenge for the AOP concept is the quantitative aspect (Hartung, 2017b). What is the cut-off value at which read-out or key event indicates a safety risk? In other words, for example, the pharmaceutical industry asks: "*How can one define a safe human dose from an AOP investigation?*" This question is still unanswered. Considering the importance of this aspect, investigative toxicology should strive to make this a powerful tool to answer this question.

4.3 Communication
The think tank discussed the need for "disruptive technologies" intensively. The meaning of "disruptive" in this context is a technology that has the potential to change the current toxicity strategies, which mainly rely on animal studies for risk assessment, entirely. "Big data", which has become a buzz word across very different industry sectors, sparks hopes of contributing to disruptive technologies, such as *in silico* modelling for precise prediction of specific organ toxicities (Mulliner et al., 2016b). However, valuable expertise of handling and connecting big data with relevant information lies outside the life science community (Haslehurst and Johnson, 2018). For example, technology and know-how of social network and internet companies may be useful to develop new strategies to translate toxicity data into meaningful context. Therefore, a suggestion is to bring together different industry sectors for new discussions about needs and opportunities that the internet and big data offer to toxicology. Examples for such new fields of collaboration are smart phone apps, which have found their use in the storage, tracking, and sharing of diagnostic data of diabetes patients, resulting in an improved self-care (Osborn et al., 2017). IT technology companies like Apple, Amazon, and Google have already made initial steps towards the life science and the healthcare sectors, e.g., by successfully applying their automated deep learning algorithms to diagnostic image analysis for the detection of retinopathy (Gulshan et al., 2016).

5 Qualification and validation

Over the last few years, many novel cell systems and technologies have emerged on to the market, most with little validation or supportive data. As a discipline, investigative toxicology needs to become proactive in finding the right balance between project-oriented issue resolution and ready to use/fit-for-purpose technology evaluation. Problem orientation also requires moving away from the formal requirement of validation, although important, as this often slows down the implementation of 3Rs methods, towards a "fit-for-purpose"[14] evaluation of new approaches. The ICH "Note for guidance on Non-clinical Safety Studies for the

---

[14] "Fit-for-purpose" is defined as a level of validation which proves that the assay or biomarker is sufficient for use in a particular defined context. It does not require a regulatory submission (for reference see footnote 13).

Conduct of Human Clinical trials for Pharmaceuticals"[1] (ICH M3 (R2) states that "(…), *consideration should be given to use of new* in vitro *alternative methods for safety evaluation. These methods, if validated and accepted by all ICH regulatory authorities, can be used to replace current standard methods*." But what is "validated"? ICH makes almost no reference to validation, except for analytical methods (ICH Q2A, Q2B, M10). Validation is defined by OECD (2005) as "*the process by which the reliability and relevance of a particular approach, method, process or assessment is established for a defined purpose*".

The European Medicines Agency (EMA) (EMA, 2016) defines "reliability" as "*a measure of the extent that a test method can be performed reproducibly over time when using the same protocol*". The reliability of current pivotal *in vivo* toxicology studies should be assured by compliance with GLP. Even though GLP principles are general and applicable in many areas, it has become evident that GLP, which was mainly tailored for *in vivo* methods available at the time of its development, needs to be adapted to the requirements of *in vitro* assays (OECD, 2004). More recently, the EU Joint Research Center (JRC), at the request of OECD, developed a guidance on Good In Vitro Method Practices - GIVIMP (OECD, 2018). The GIVIMP document describes the factors relevant to reliability and relevance of *in vitro* data generated for human safety assessment purposes and has been written with different users in mind, including GLP test facilities and research laboratories developing new *in vitro* methods.

"Relevance" is defined by EMA (2016) as "*the extent to which the test correctly measures or predicts the biological effect of interest*". So, how should relevance be established? According to EMA, "*Relevance incorporates consideration of the accuracy (e.g. concordance with comparable validated test method with established performance standards) of a test method*". However, this assumes that the existing "validated test method" is adequately relevant, which is not always the case for some *in vivo* animal test systems. A direct example of this is AOP-based *in vitro* skin sensitization testing, where concordance of the *in vitro* results with the standard *in vivo* mouse local lymph node assay (LLNA) is poor (Dumont et al., 2016) due to both limited relevance and high variability of the reference data. In this case, concordance of the *in vitro* data with human data is much better (concordance of the *in vivo* LLNA with human data is relatively poor) (Natsch and Emter, 2015) as was already shown for the guinea pig assay preceding the LLNA (Luechtefeld et al., 2016; Adriaens et al., 2014; Hoffmann, 2015). In the meantime, several testing and assessment strategies have been published (Urbisch et al., 2015; Roberts and Patlewicz, 2018) that have contributed to the development of a consolidated integrative approach to testing and assessment (IATA) for skin sensitization by OECD (2014).

For the pharmaceutical sector, EMA and FDA have put in place a qualification process that addresses innovative drug development methods and tools developed for a specific intended use in a pharmaceuticals research and development context (non-clinical or clinical studies) (EMA, 2014). These are voluntary, scientific pathways leading to a regulatory conclusion that an innova-

tive method or tool has data to support its specific context of use in drug development (Ohno, 2002). The qualification concept also applies to the submission of biomarkers; however, in this case it is independent of the specific test or assay performing the measurement[15].

Overall, while the terminology (i.e., qualification versus validation) differs and the processes evolve to keep pace with scientific progress and to benefit from it, the underlying purpose and principles of qualification and validation remain relatively constant as laid down in the modular approach to validation and in the OECD Guidance Document (Hartung et al., 2004; OECD, 2005). As described by Hartung et al. (2013), the validation procedure should follow fit-for-purpose approaches. A procedure that strictly follows the modular approach system, comparing *in vitro* data with animal data, still relies on the concept of one-to-one replacement. New concepts need to be developed, taking into account combinations of assays (IATA) and human data. A combination of *in silico* and *in vitro* might be the key to solve this problem.

6 Recommendation: the way forward

Investigative toxicology is a recent discipline in drug safety assessment that strives for a holistic view on properties of candidate molecules (NCEs/NBEs) and their predicted effect in humans by combining *in silico*, *in vitro*, *in vivo*, and clinical data and making use of innovative technologies and novel approaches (Fig. 8). Thus, it should be seen as complementary to regulatory toxicology, which is often confined by international guidelines. Investigative toxicology can embrace novel technologies more readily. Thereby, it is able to evaluate these technologies and produces evidence for their further use also in the regulatory context.
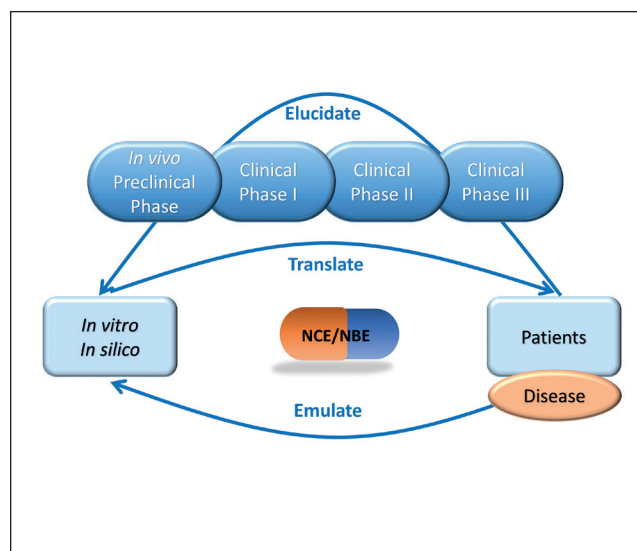
Fig. 8: General view on how investigative toxicology will contribute to the development of new drugs

---

[15] https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/Biomarker QualificationProgram/default.htm (accessed 24.07.2018).

In a prospective manner, efforts in investigative toxicology focus on early screening tools allowing candidate selection as well as de-risking activities, where tailored work packages aim to unravel mechanisms of adverse events observed during pre-clinical and clinical stages and assessment of potential ways forward. The key to success of these activities is close interaction with experts from academia and regulatory bodies. The objective is to elaborate robust, reliable, and accepted investigative toxicology concepts for decision making by virtue of multiple and diverse tools, technologies, and readouts.

In the rapidly evolving field of drug development, with new drug modalities arising, new chemical spaces being explored, complex pharmacological strategies being pursued, and diseases and pathways involving the immune system becoming more prominent, it is instrumental that safety assessment is continuously innovated to be able to address new challenges arising from these trends. Together with high regulatory burden, cost and time pressures, and the need to reduce animal testing demands for novel concepts, strategies, and tools, investigative toxicology plays a key role. Collaborative efforts in the pre-competitive space are considered necessary to explore and establish these innovative new tools and concepts.

An important role is played by education to create a new generation of scientists (Daneshian et al., 2011). Novel technologies, innovative scientific approaches, and new terminologies demand modern and up-dated ways of teaching-learning (Flecha, 2018). Professional knowledge and expertise in this discipline is mandatory, considering that the possible target audience is wide and includes both undergraduate students and graduates (M.Sc., Ph.D., post-docs), as wells as academic and industry scientists, and regulators. Teaching must go hand in hand with research, and common strategies and novel approaches are urgently required. Currently, there are few university graduate and post-graduate level courses in life science that provide adequate training although their implementation is fundamental for a competency-based education and, consequently, social impact.

Among the key themes are:
− Making use of clinical and real-world data to inform and improve testing algorithms early on, establishing modeling approaches to predict the human response from *in vitro* data,
− Use of new human cell models, such as microphysiological systems and human organs-on-chips, for *in vitro* safety and efficacy assessment and generation of a cellular therapeutic index, modeling disease aspects by use of patient-derived tissues and iPSCs for patient stratification, and performing phenotypic screens and identification of translational biomarkers,
− Use of image-based technologies and omics readouts to support *in vitro* to *in vivo* assessment as well as back translation from clinical specimens to cell models to understand safety and disease as well as patient-specific, personalized aspects; monitoring AOPs and identifying key events from early *in vitro* discovery stages up to clinical phases,
− Focusing on quantitative systems biology level-based risk assessments that integrate diverse pharmacological and toxicological data sets, including data from advanced humanized models and human tissue, to underpin mechanistic under-

standing of adverse drug effects and permit translational, exposure-based modeling of toxicological impact,
− Education based on a new integrated teaching strategy.

The participants of the workshop envisage an increasing importance of investigative toxicology in supporting the entire drug development process. Cross-industry collaboration in the pre-competitive part of this work is helping to accumulate experiences on novel approaches faster and more reliably. This process can impact on the complementation and replacement of traditional methods in regulatory toxicology. This also means that the promotion of new technologies supports a lesser reliance on animal studies, moving the industry to more human-relevant assessment approaches.

References

Adriaens, E., Barroso, J., Eskes, C. et al. (2014). Retrospective analysis of the Draize test for serious eye damage/eye irritation: Importance of understanding the in vivo endpoints under UN GHS/EU CLP for the development and evaluation of in vitro test methods. *Arch Toxicol 88*, 701-723. doi:10.1007/s00204-013-1156-8

Alépée, N., Bahinski, A., Daneshian, M. et al. (2014). State-of-the-art of 3D cultures (organs-on-a-chip) in safety testing and pathophysiology. *ALTEX 31*, 441-477. doi:10.14573/altex1406111

Amberg, A., Czich, A. and Thybaud, V. (2014). In silico/computational assessment for the evaluation of genotoxic impurities. In H. Lee (ed.), *Pharmaceutical Industry Practices on Genotoxic Impurities* (528-553). 1st edition. Boca Raton, FL, USA: CRC Press. doi:10.1201/b17350

Amur, S., Lavange, L., Zineh, I. et al. (2015). Biomarker qualification: Toward a multiple stakeholder framework for biomarker development, regulatory acceptance, and utilization. *Clin Pharmacol Ther 98*, 34-46. doi:10.1002/cpt.136

Ankley G. T., Bennett, R. S., Erickson, R. J. et al. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem 29*, 730-741. doi:10.1002/etc.34

Antoine, D. J., James, W., Dear, P. S. L. et al. (2013). Mechanistic biomarkers provide early and sensitive detection of acetaminophen-induced acute liver injury at first presentation to hospital. *Hepatology 58*, 777-787. doi:10.1002/hep.26294

Anton, D., Burckel, H., Josset, E. and Noel, G. (2015). Three-dimensional cell culture: A breakthrough in vivo. *Int J Mol Sci 16*, 5517-5527. doi:10.3390/ijms16035517

Antonica, F., Kasprzyk, D. F., Opitz, R. et al. (2012). Generation of functional thyroid from embryonic stem cells. *Nature 491*, 66-71. doi:10.1038/nature11525

Aoi, T., Yae, K., Nakagawa, M. et al. (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science 321*, 699-702. doi:10.1126/science.1154884

Artegiani, B. and Clevers, H. (2018). Use and application of 3D-organoid technology. *Hum Mol Genet 27*, R99-107. doi:10.1093/hmg/ddy187

Atienzar, F. A., Blomme, E. A., Chen, M. et al. (2016). Key

challenges and opportunities associated with the use of in vitro models to detect human DILI: Integrated risk assessment and mitigation plans. *BioMed Res Int 2016*, 9737920. doi:10.1155/2016/9737920

Basketter, D. A., Clewell, H., Kimber, I. et al. (2012). A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing – t4 report. *ALTEX 29*, 3-91. doi:10.14573/altex.2012.1.003

Bauer, S., Wennberg Huldt, C., Kanebratt, K. P. et al. (2017). Functional coupling of human pancreatic islets and liver spheroids on-a-chip: Towards a novel human ex vivo type 2 diabetes model. *Sci Rep 7*, 14620. doi:10.1038/s41598-017-14815-w

Bell, C. C., Hendriks, D. F., Moro, S. M. et al. (2016). Characterization of primary human hepatocyte spheroids as a model system for drug-induced liver injury, liver function and disease. *Sci Rep 6*, 25187. doi:10.1038/srep25187

Benam, K. H., Dauth, S., Hassell, B. et al. (2015). Engineered in vitro disease models. *Ann Rev Pathol 10*, 195-262. doi:10.1146/annurev-pathol-012414-040418

Bhatia, S. N. and Ingber, D. E. (2014). Microfluidic organs-on-chips. *Nat Biotechnol 32*, 760-772. doi:10.1038/nbt.2989

Billis, P., Will, Y. and Nadanaciva, S. (2014). High-content imaging assays for identifying compounds that generate superoxide and impair mitochondrial membrane potential in adherent eukaryotic cells. *Curr Protoc Toxicol 59*, 25.1.1-25.1.14. doi:10.1002/0471140856.tx2501s59

Blaauboer, B. J., Boekelheide, K., Clewell, H. J. et al. (2012). The use of biomarkers of toxicity for integrating in vitro hazard estimates into risk assessment for humans. *ALTEX 29*, 411-425. doi:10.14573/altex.2012.4.411

Bouhifd, M., Hartung, T., Hogberg, H. T. et al. (2013). Review: Toxicometabolomics. *J Appl Toxicol 33*, 1365-1383. doi:10.1002/jat.2874

Bouhifd, M., Beger, R., Flynn, T. et al. (2015). Quality assurance of metabolomics. *ALTEX 32*, 319-326. doi:10.14573/altex.1509161

Brennan, F. R., Baumann, A., Blaich, G. et al. (2015). Nonclinical safety testing of biopharmaceuticals – Addressing current challenges of these novel and emerging therapies. *Regul Toxicol Pharmacol 73*, 265-275. doi:10.1016/j.yrtph.2015.07.019

Brooks, J., Watson, A. and Korcsmaros, T. (2017). Omics approaches to identify potential biomarkers of inflammatory diseases in the focal adhesion complex. *Genomics Proteomics Bioinformatics 15*, 101-109. doi:10.1016/j.gpb.2016.12.003

Brott, D. A., Adler, S. H., Arani, R. et al. (2014). Characterization of renal biomarkers for use in clinical trials: Biomarker evaluation in healthy volunteers. *Drug Des Devel Ther 8*, 227-237. doi:10.2147/DDDT.S54956

Burden, N., Sewell, F., Andersen, M. E. et al. (2015). Adverse outcome pathways can drive non-animal approaches for safety assessment. *J Appl Toxicol 35*, 971-975. doi:10.1002/jat.3165

Bussiere, J. L., Martin, P., Horner, M. et al. (2009). Alternative strategies for toxicity testing of species-specific biopharmaceuticals. *Int J Toxicol 28*, 230-253. doi:10.1177/1091581809337262

Carragher, N., Piccinini, F., Tesei, A. et al. (2018). Concerns, challenges and promises of high-content analysis of 3D cellular models. *Nat Rev Drug Discov 17*, 606-606. doi:10.1038/nrd.2018.99

Cayo, M. A., Cai, J., Delaforest, A. et al. (2012). JD induced pluripotent stem cell-derived hepatocytes faithfully recapitulate the pathophysiology of familial hypercholesterolemia. *Hepatology 56*, 2163-2171. doi:10.1002/hep.25871

Chau, C. H., Rixe, O., McLeod, H. and Figg, W. D. (2008). Validation of analytic methods for biomarkers used in drug development. *Clin Cancer Res 14*, 5967-5976. doi:10.1158/1078-0432.CCR-07-4535

Chen, M., Zhang, M., Borlak, J. and Tong, W. (2012). A decade of toxicogenomic research and its contribution to toxicological science. *Toxicol Sci 130*, 217-228. doi:10.1093/toxsci/kfs223

Church, R. J., Kullak-Ublick, G. A., Aubrecht, J. et al. (2018). Candidate biomarkers for the diagnosis and prognosis of drug-induced liver injury: An international collaborative effort. *Hepatology 69*, 760-773. doi:10.1002/hep.29802

Clarke, J. I., Dear, J. W. and Antoine, D. J. (2016). Recent advances in biomarkers and therapeutic interventions for hepatic drug safety – False dawn or new horizon? *Expert Opin Drug Saf 15*, 625-634. doi:10.1517/14740338.2016.1160057

Clark, M. and Steger-Hartmann, T. (2018). A big data approach to the concordance of the toxicity of pharmaceuticals in animals and humans. *Regul Toxicol Pharmacol 96*, 94-105. doi:10.1016/J.YRTPH.2018.04.018

Coecke, S., Balls, M., Bowe, G. et al. (2007). Guidance on good cell culture practice. In R. Smith (ed), *Cell Technology for Cell Products* (313-315). Dordrecht, The Netherlands: Springer. doi:10.1007/978-1-4020-5476-1_49

Cook, D., Brown, D., Alexander, R. et al. (2014). Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework. *Nat Rev Drug Discov 13*, 419-431. doi:10.1038/nrd4309

Corsini, A., Ganey, P., Ju, C. et al. (2012). Current challenges and controversies in drug-induced liver injury. *Drug Saf 35*, 1099-1117. doi:10.2165/11632970-000000000-00000

Cox, J. and Mann, M. (2011). Quantitative, high-resolution proteomics for data-driven systems biology. *Ann Rev Biochem 80*, 273-299. doi:10.1146/annurev-biochem-061308-093216

Crawford, S. E., Hartung, T., Hollert, H. et al. (2017). Green toxicology: A strategy for sustainable chemical and material development. *Environ Sci Eur 29*, 16. doi:10.1186/s12302-017-0115-z

Cvetkovic, C., Raman, R., Chan, V. et al. (2014). Three-dimensionally printed biological machines powered by skeletal muscle. *Proc Natl Acad Sci U S A 111*, 10125-10130. doi:10.1073/pnas.1401577111

Daneshian, M., Akbarsha, M. A., Blaauboer, B. et al. (2011). A framework program for the teaching of alternative methods (replacement, reduction, refinement) to animal experimentation. *ALTEX 28*, 341-352. doi:10.14573/altex.2011.4.341

Dixit, R. and Boelsterli, U. A. (2007). Healthy animals and animal models of human disease(s) in safety assessment of human pharmaceuticals, including therapeutic antibodies. *Drug Discov Today 12*, 337-342. doi:10.1016/j.drudis.2007.02.018

Dorato, M. A. and Engelhardt, J. A. (2005). The no-observed-adverse-effect-level in drug safety evaluations: Use, issues, and definition(s). *Regul Toxicol Pharmacol 42*, 265-274. doi:10.1016/j.yrtph.2005.05.004

Dumont, C., Barroso, J., Matys, I. et al. (2016). Analysis of the local lymph node assay (LLNA) variability for assessing the prediction of skin sensitisation potential and potency of chemicals with non-animal approaches. *Toxicol In Vitro 34*, 220-228. doi:10.1016/j.tiv.2016.04.008

Duval, K., Grover, H., Han, L.-H. et al. (2017). Modeling physiological events in 2D vs. 3D cell culture. *Physiology 32*, 266-277. doi:10.1152/physiol.00036.2016

Edington, C. D., Chen, W. L. K., Geishecker, E. et al. (2018). Interconnected microphysiological systems for quantitative biology and pharmacology studies. *Sci Rep 8*, 4530. doi:10.1038/s41598-018-22749-0

Edwards, S. W., Tan, Y.-M., Villeneuve, D. L. et al. (2016). Adverse outcome pathways – Organizing toxicological information to improve decision making. *J Pharmacol Exp Ther 356*, 170-181. doi:10.1124/jpet.115.228239

Ellinger-Ziegelbauer, H., Aubrecht, J., Kleinjans, J. C. and Ahr., H. J. (2009). Application of toxicogenomics to study mechanisms of genotoxicity and carcinogenicity. *Toxicol Lett 186*, 36-44. doi:10.1016/j.toxlet.2008.08.017

EMA (2014). Qualification of novel methodologies for drug development: Guidance to applicants. EMA/CHMP/SAWP/72894/2008. https://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004201.pdf

EMA (2016). Guideline on the principles of regulatory acceptance of 3Rs (replacement, reduction, refinement) testing approaches. EMA/CHMP/CVMP/JEG-3Rs/450091/2012. https://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2017/03/WC500224306.pdf

Esch, E. W., Bahinski, A. and Huh, D. (2015). Organs-on-chips at the frontiers of drug discovery. *Nat Rev Drug Discov 14*, 248-260. doi:10.1038/nrd4539

Eskes, C., Boström, A. C., Bowe, G. et al. (2017). Good cell culture practices & in vitro toxicology. *Toxicol In Vitro 45*, 272-277. doi:10.1016/j.tiv.2017.04.022

Ewart, L., Dehne, E. M., Fabre, K. et al. (2018). Application of microphysiological systems to enhance safety assessment in drug discovery. *Ann Rev Pharmacol Toxicol 58*, 65-82. doi:10.1146/annurev-pharmtox-010617-052722

Ferrario, D., Brustio, R. and Hartung, T. (2014). Glossary of reference terms for alternative test methods and their validation. *ALTEX 31*, 319-335. doi:10.14573/altex.1403311

Flecha, R. (2018). Social impact of community-based educational programs in Europe. In G. W. Nobilt, *Oxford Research Encyclopedia of Education* (1-21). doi:10.1093/acrefore/9780190264093.013.184

Freedman, B. S., Brooks, C. R., Lam, A. Q. et al. (2015). Modelling kidney disease with CRISPR-mutant kidney organoids derived from human pluripotent epiblast spheroids. *Nat Comm 6*, 8715. doi:10.1038/ncomms9715

Gadue, P. and Cotsarelis, G. (2008). Epidermal cells rev up reprogramming. *Nat Biotechnol 26*, 1243-1244. doi:10.1038/nbt1108-1243

Ganter, B., Tugendreich, S., Pearson, C. I. et al. (2005). Development of a large-scale chemogenomics database to improve drug candidate selection and to understand mechanisms of chemical toxicity and action. *J Biotechnol 119*, 219-244. doi:10.1016/j.jbiotec.2005.03.022

Garcia-Serna, R., Vidal, D., Remez, N. and Mestres, J. (2015). Large-scale predictive drug safety: From structural alerts to biological mechanisms. *Chem Res Toxicol 28*, 1875-1887. doi:10.1021/acs.chemrestox.5b00260

Germano, D., Uteng, M., Pognan, F. et al. (2015). Determination of liver specific toxicities in rat hepatocytes by high content imaging during 2-week multiple treatment. *Toxicol In Vitro 30*, 79-94. doi:10.1016/j.tiv.2014.05.009

Giobbe, G. G., Michielin, F., Luni, C. et al. (2015). Functional differentiation of human pluripotent stem cells on a chip. *Nat Methods 12*, 637-640. doi:10.1038/nmeth.3411

Goh, J. Y., Weaver, R. J., Dixon, L. et al. (2015). Development and use of in vitro alternatives to animal testing by the pharmaceutical industry 1980-2013. *Toxicol Res 4*, 1297-1307. doi:10.1039/c5tx00123d

Greggio, C., De Franceschi, F., Figueiredo-Larsen, M. et al. (2013). Artificial three-dimensional niches deconstruct pancreas development in vitro. *Development 140*, 4452-4462. doi:10.1242/dev.096628

Gulshan, V., Peng, L., Coram, M. et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA 316*, 2402-2410. doi:10.1001/jama.2016.17216

Hardwick, R. N., Viergever, C., Chen, A. E. and Nguyen, D. G. (2017). 3D bioengineered tissues: From advancements in in vitro safety to new horizons in disease modeling. *Clin Pharmacol Ther 101*, 453-457. doi:10.1002/cpt.569

Hardy, B., Apic, G., Carthew, P. et al. (2012a). Food for thought ... a toxicology ontology roadmap. *ALTEX 29*, 129-137. doi:10.14573/altex.2012.2.129

Hardy, B., Apic, G., Carthew, P. et al. (2012b). Toxicology ontology perspectives. *ALTEX 29*, 139-156. doi:10.14573/altex.2012.2.139

Hartung, T., Bremer, S., Casati, S. et al. (2004). A modular approach to the ECVAM principles on test validity. *Altern Lab Anim 32*, 467-472.

Hartung, T. and McBride, M. (2011). Food for thought ... on mapping the human toxome. *ALTEX 28*, 83-93. doi:10.14573/altex.2011.2.083

Hartung, T., van Vliet, E., Jaworska, J. et al. (2012). Systems toxicology. *ALTEX 29*, 119-128. doi:10.14573/altex.2012.2.119

Hartung, T., Hoffmann, S. and Stephens, M. (2013). Mechanistic validation. *ALTEX 30*, 119-130. doi:10.14573/altex.2013.2.119

Hartung, T. (2016). Making big sense from big data in toxicology by read-across. *ALTEX 33*, 83-93. doi:10.14573/altex.1603091

Hartung, T. (2017a). Perspectives on in vitro to in vivo extrapolations. *Appl In Vitro Toxicol 4*, aivt.2016.0026. doi:10.1089/aivt.2016.0026

Hartung, T. (2017b). Utility of the adverse outcome pathway concept in drug Development. *Expert Opin Drug Metab Toxicol 13*, 1-3. doi:10.1080/17425255.2017.1246535

Hartung, T., FitzGerald, R. E., Jennings, P. et al. (2017). Systems toxicology: Real world applications and opportunities. *Chem Res Toxicol 30*, 870-882. doi:10.1021/acs.chemrestox.7b00003

Haslehurst, R. and Johnson, J. (2018). As Amazon turns its gaze to healthcare, the industry may be in for a wild ride. L.E.K. Consulting. https://www.lek.com/insights/amazon-turns-its-gaze-healthcare-industry-may-be-wild-ride

Hoffmann, S. (2015). LLNA variability: An essential ingredient for a comprehensive assessment of non-animal skin sensitization test methods and strategies. *ALTEX 32*, 379-383. doi:10.14573/altex.1505051

Hornberg, J. J., Laursen, M., Brenden, N. et al. (2014a). Exploratory toxicology as an integrated part of drug discovery. Part I: Why and how. *Drug Discov Today 19*, 1131-1136. doi:10.1016/j.drudis.2013.12.008

Hornberg, J. J., Laursen, M., Brenden, N. et al. (2014b). Exploratory toxicology as an integrated part of drug discovery. Part II: Screening strategies. *Drug Discov Today 19*, 1137-1144. doi:10.1016/j.drudis.2013.12.009

Hsin, K. Y., Ghosh, S. and Kitano, H. (2013). Combining machine learning systems and multiple docking simulation packages to improve docking prediction reliability for network pharmacology. *PLoS One 8*, e83922. doi:10.1371/journal.pone.0083922

Hu, Z., Qian, M. and Zhang, M. Q. (2011). Novel Markov model of induced pluripotency predicts gene expression changes in reprogramming. *BMC Syst Biol 5, Suppl 2*, S8. doi:10.1186/1752-0509-5-S2-S8.

Huch, M., Dorrell, C., Boj, S. F. et al. (2013). In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. *Nature 494*, 247-250. doi:10.1038/nature11826

Huch, M., Gehart, H., van Boxtel, R. et al. (2015). Long-term culture of genome-stable bipotent stem cells from adult human liver. *Cell 160*, 299-312. doi:10.1016/j.cell.2014.11.050

Hübner, J., Raschke, M., Rütschle, I. et al. (2018). Simultaneous evaluation of anti-EGFR-induced tumour and adverse skin effects in a microfluidic human 3D co-culture model. *Sci Rep 8*, 15010. doi:10.1038/s41598-018-33462-3

Huh, D., Matthews, B. D., Mammoto, A. et al. (2010). Reconstituting organ-level lung functions on a chip. *Science 328*, 1662-1668. doi:10.1126/science.1188302

Igarashi, Y., Nakatsu, N., Yamashita, T. et al. (2015). Open TG-GATEs: A large-scale toxicogenomics database. *Nucleic Acids Res 43*, D921-927. doi:10.1093/nar/gku955

Joshi, P. and Lee, M. Y. (2015). High content imaging (HCI) on miniaturized three-dimensional (3D) cell cultures. *Biosensors 5*, 768-790. doi:10.3390/bios5040768

Kangas, J. D., Naik, A. W. and Murphy, R. F. (2014). Efficient discovery of responses of proteins to compounds using active learning. *BMC Bioinformatics 15*, 143. doi:10.1186/1471-2105-15-143

Kausar, S. and Falcao, A. O. (2018). An automated framework for QSAR model building. *J Cheminform 10*, 1. doi:10.1186/s13321-017-0256-5

Khan, S. R., Baghdasarian, A., Fahlman, R. P. et al. (2014). Current status and future prospects of toxicogenomics in drug discovery. *Drug Discov Today 19*, 562-578. doi:10.1016/j.drudis.2013.11.001

King, S. M., Higgins, J. W., Nino, C. R. et al. (2017). 3D proximal tubule tissues recapitulate key aspects of renal physiology to enable nephrotoxicity testing. *Front Physiol 8*, 123. doi:10.3389/fphys.2017.00123

Klaeger, S., Heinzlmeir, S., Wilhelm, M. et al. (2017). The target landscape of clinical kinase drugs. *Science 358*, eaan4368. doi:10.1126/science.aan4368

Kleensang, A., Maertens, A., Rosenberg, M. et al. (2014). Pathways of toxicity. *ALTEX 31*, 53-61. doi:10.14573/altex.1309261

Klymiuk, N., Seeliger, F., Bohlooly-Y, M. et al. (2016). Tailored pig models for preclinical efficacy and safety testing of targeted therapies. *Toxicol Pathol 44*, 346-357. doi:10.1177/0192623315609688

Knapen, D., Angrish, M. M., Fortin, M. C. et al. (2018). Adverse outcome pathway networks I: Development and applications. *Environ Toxicol Chem 37*, 1723-1733. doi:10.1002/etc.4125

Kotsampasakou, E., Montanari, F. and Ecker, G. F. (2017). Predicting drug-induced liver injury: The importance of data curation. *Toxicology 389*, 139-145. doi:10.1016/j.tox.2017.06.003

Kretschmann, A., Ashauer, R., Hollender, J. and Escher, B. I. (2012). Toxicokinetic and toxicodynamic model for Diazinon toxicity – Mechanistic explanation of differences in the sensitivity of Daphnia magna and Gammarus pulex. *Environ Toxicol Chem 31*, 2014-2022. doi:10.1002/etc.1905

Lancaster, M. A. and Knoblich, J. A. (2014). Organogenesis in a dish: Modeling development and disease using organoid technologies. *Science 345*, 1247125. doi:10.1126/science.1247125

Laverty, H. G., Benson, C., Cartwright, E. J. et al. (2011). How can we improve our understanding of cardiovascular safety liabilities to develop safer medicines? *Br J Pharmacol 163*, 675-693. doi:10.1111/j.1476-5381.2011.01255.x

Leary, E., Rhee, C., Wilks, B. T. and Morgan, J. R. (2018). Quantitative live-cell confocal imaging of 3D spheroids in a high-throughput format. *SLAS Technol 23*, 231-242. doi:10.1177/2472630318756058

Lee, J. H., Bhang, D. H., Beede, A. et al. (2014). Lung stem cell differentiation in mice directed by endothelial cells via a BMP4-NFATc1-thrombospondin-1 axis. *Cell 156*, 440-455. doi:10.1016/j.cell.2013.12.039

Lee-Montiel, F. T., George, S. M., Gough, A. H. et al. (2017). Control of oxygen tension recapitulates zone-specific functions in human liver microphysiology systems. *Exp Biol Med 242*, 1617-1632. doi:10.1177/1535370217703978

Leist, M., Hasiwa, N., Rovida, C. et al. (2014). Consensus report on the future of animal-free systemic toxicity testing. *ALTEX 31*, 341-356. doi:10.14573/altex.1406091

Lin, C., Ballinger, K. R. and Khetani, S. R. (2015). The application of engineered liver tissues for novel drug discovery. *Expert Opin Drug Discov 10*, 519-540. doi:10.1517/17460441.2015.1032241

Loh, Y.-H., Agarwal, S., Park, I.-H. et al. (2009). Generation of induced pluripotent stem cells from human blood. *Blood 113*, 5476-5479. doi:10.1182/blood-2009-02-204800

Louisse, J., Beekmann, K. and Rietjens, I. M. (2017). Use of physiologically based kinetic modeling-based reverse dosimetry to predict in vivo toxicity from in vitro data. *Chem Res Toxicol 30*, 114-125. doi:10.1021/acs.chemrestox.6b00302

Lu, B., Nakamura, T., Inouye, K. et al. (2012). Novel role of PKR in inflammasome activation and HMGB1 release. *Nature 488*, 670-674. doi:10.1038/nature11290

Luechtefeld, T., Maertens, A., Russo, D. P. et al. (2016). Analysis of publically available skin sensitization data from REACH registrations 2008-2014. *ALTEX 33*, 135-148. doi:10.14573/altex.1510055

Madden, L. R., Nguyen, T. V., Garcia-Mojica, S. et al. (2018). Bioprinted 3D primary human intestinal tissues model aspects of native physiology and ADME/Tox functions. *iScience 2*, 156-167. doi:10.1016/j.isci.2018.03.015

Maertens, A., Anastas, N., Spencer, P. J. et al. (2014). Green toxicology. *ALTEX 31*, 243-249. doi:10.14573/altex.1406181

Maertens, A. and Hartung, T. (2018). Green toxicology – Know early about and avoid toxic product liabilities. *Toxicol Sci 161*, 285-289. doi:10.1093/toxsci/kfx243.

Marrer, E. and Dieterle, F. (2010). Impact of biomarker development on drug safety assessment. *Toxicol Appl Pharmacol 243*, 167-179. doi:10.1016/j.taap.2009.12.015

Marx, U., Andersson, T. B., Bahinski, A. et al. (2016). Biology-inspired microphysiological system approaches to solve the prediction dilemma of substance testing. *ALTEX 33*, 272-321. doi:10.14573/altex.1603161

Maschmeyer, I., Lorenz, A. K., Schimek, K. et al. (2015). A four-organ-chip for interconnected long-term co-culture of human intestine, liver, skin and kidney equivalents. *Lab Chip 15*, 2688-2699. doi:10.1039/C5LC00392J

Materne, E.-M., Tonevitsky, A. G. and Marx, U. (2013). Chip-based liver equivalents for toxicity testing-organotypicalness versus cost-efficient high throughput. *Lab Chip 13*, 3481-3495. doi:10.1039/c3lc50240f

Materne, E.-M., Ramme, A. P., Terrasso, A. P. et al. (2015). A multi-organ chip co-culture of neurospheres and liver equivalents for long-term substance testing. *J Biotechnol 205*, 36-46. doi:10.1016/j.jbiotec.2015.02.002

Matheis, K., Laurie, D., Andriamandroso, C. et al. (2011). A generic operational strategy to qualify translational safety biomarkers. *Drug Discov Today 16*, 600-608. doi:10.1016/j.drudis.2011.04.011

Mauritz, C., Schwanke, K., Reppel, M. et al. (2008). Generation of functional murine cardiac myocytes from induced pluripotent stem cells. *Circulation 118*, 507-517. doi:10.1161/circulationaha.108.778795

Mayr, A., Klambauer, G., Unterthiner, T. and Hochreiter, S. (2016). DeepTox: Toxicity prediction using deep learning. *Front Environ Sci 3*, 80. doi:10.3389/fenvs.2015.00080

Messner, S., Agarkova, I., Moritz, W. and Kelm, J. M. (2013). Multi-cell type human liver microtissues for hepatotoxicity testing. *Arch Toxicol 87*, 209-213. doi:10.1007/s00204-012-0968-2

Millard, D., Dang, Q., Shi, H. et al. (2018). Cross-site reliability of human induced pluripotent stem cell-derived cardiomyocyte based safety assays using microelectrode arrays: Results from a blinded CiPA pilot study. *Toxicol Sci 164*, 550-562. doi:10.1093/toxsci/kfy110

Miller, P. G. and Shuler, M. L. (2016). Design and demonstration of a pumpless 14 compartment, microphysiological system. *Biotechnol Bioeng 113*, 2213-2227. doi:10.1002/bit.25989

Moggs, J., Moulin, P., Pognan, F. et al. (2012). Investigative safety science as a competitive advantage for pharma. *Expert Opin Drug Metabol Toxicol 8*, 1071-1082. doi:10.1517/17425255.2012.693914

Monticello, T. M. (2015). Drug development and nonclinical to clinical translational databases: Past and current efforts. *Toxicol Pathol 43*, 57-61. doi:10.1177/0192623314557189

Morelli, J. K., Buehrle, M., Pognan, F. et al. (2006). Validation of an in vitro screen for phospholipidosis using a high-content biology platform. *Cell Biol Toxicol 22*, 15-27. doi:10.1007/s10565-006-0176-z

Morgan, S. J., Elangbam, C. S., Berens, S. et al. (2013). Use of animal models of human disease for nonclinical safety assessment of novel pharmaceuticals. *Toxicol Pathol 41*, 508-518. doi:10.1177/0192623312457273

Mueller, D., Krämer, L., Hoffmann, E. et al. (2014). 3D organotypic HepaRG cultures as in vitro model for acute and repeated dose toxicity studies. *Toxicol In Vitro 28*, 104-112. doi:10.1016/j.tiv.2013.06.024

Mulliner, D., Schmidt, F., Stolte, M. et al. (2016a). Computational models for human and animal hepatotoxicity with a global application scope. *Chem Res Toxicol 29*, 757-767. doi:10.1021/acs.chemrestox.5b00465

Mulliner, D., Schmidt, F., Stolte, M. et al. (2016b). Computational models for human and animal hepatotoxicity with a global application scope. *Chem Res Toxicol 29*, 757-767. doi:10.1021/acs.chemrestox.5b00465

Murphy, R. F. (2011). An active role for machine learning in drug development. *Nat Chem Biol 7*, 327-330. doi:10.1038/nchembio.576

Nakano, T., Ando, S., Takata, N. et al. (2012). Self-formation of optic cups and storable stratified neural retina from human ESCs. *Cell Stem Cell 10*, 771-785. doi:10.1016/j.stem.2012.05.009

Natsch, A. and Emter, R. (2015). Reporter cell lines for skin sensitization testing. *Arch Toxicol 89*, 1645-1668. doi:10.1007/s00204-015-1555-0

Nguyen, D. G., Funk, J., Robbins, J. B. et al. (2016). Bioprinted 3D primary liver tissues allow assessment of organ-level response to clinical drug induced toxicity in vitro. *PLoS One 11*, e0158674. doi:10.1371/journal.pone.0158674

Ni, Q., Mehta, S. and Zhang, J. (2017). Live-cell imaging of cell signaling using genetically encoded fluorescent reporters. *FEBS J 285*, 203-219. doi:10.1111/febs.14134

OECD (2004). The Application of the Principles of GLP to in vitro Studies. *OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring No. 14*. OECD Publishing, Paris. doi:10.1787/9789264084971-en

OECD (2005). Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. *OECD Series on Testing and Assessment No. 34*. OECD Publishing, Paris. https://bit.ly/2EffJp7

OECD (2014). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins., *OECD Series on Testing and Assessment No.168*. OECD Publishing, Paris. doi:10.1787/9789264221444-en

OECD (2017). Revised Guidance Document on Developing and Assessing Adverse Outcome Pathways. *OECD Series on Testing and Assessment No. 184*, second edition. https://bit.ly/1R-JXOEx

OECD (2018). Guidance Document on Good In Vitro Method Practices (GIVIMP). *OECD Series on Testing and Assessment No. 286*. OECD Publishing, Paris, doi:10.1787/9789264304796-en

Ohno, Y. (2002). ICH guidelines – Implementation of the 3Rs (refinement, reduction, and replacement): Incorporating best scientific practices into the regulatory process. *ILAR J 43, Suppl 1*, S95-98. doi:10.1093/ilar.43.Suppl_1.S95

Okita, K., Ichisaka, T. and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature 448*, 313-317. doi:10.1038/nature05934

Olson, H., Betton, G., Robinson, D. et al. (2000). Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regul Toxicol Pharmacol 32*, 56-67. doi:10.1006/rtph.2000.1399

Osborn, C. Y., van Ginkel, J. R., Rodbard, D. et al. (2017). One Drop | Mobile: An evaluation of hemoglobin A1c improvement linked to app engagement. *JMIR Diabetes 2*, e21. doi:10.2196/diabetes.8039

Ozer, J., Ratner, M., Shaw, M. et al. (2008). The current state of serum biomarkers of hepatotoxicity. *Toxicology 245*, 194-205. doi:10.1016/j.tox.2007.11.021

Pallardy, M. and Bechara, R. (2017). Chemical or drug hypersensitivity: Is the immune system clearing the danger? *Toxicol Sci 158*, 14-22. doi:10.1093/toxsci/kfx084

Pamies, D. and Hartung, T. (2017). 21st century cell culture for 21st century toxicology. *Chem Res Toxicol 30*, 43-52. doi:10.1021/acs.chemrestox.6b00269

Pamies, D., Bal-Price, A., Simeonov, A. et al. (2017). Good cell culture practice for stem cells & stem-cell-derived models. *ALTEX 34*, 95-132. doi:10.14573/altex.1607121

Pamies, D., Bal-Price, A., Chesné, C. et al. (2018). Advanced good cell culture practice for human primary, stem cell-derived and organoid models as well as microphysiological systems. *ALTEX 35*, 353-378. doi:10.14573/altex.1710081

Pampaloni, F., Reynaud, E. G. and Stelzer, E. H. (2007). The third dimension bridges the gap between cell culture and live tissue. *Nat Rev Mol Cell Biol 8*, 839-845. doi:10.1038/nrm2236

Park, B. K., Kitteringham, N. R., Powell, H. and Pirmohamed, M. (2000). Advances in molecular toxicology – Towards understanding idiosyncratic drug toxicity. *Toxicology 153*, 39-60. doi:10.1016/S0300-483X(00)00303-6.

Passarelli, M. K. and Ewing, A. G. (2013). Single-cell imaging mass spectrometry. *Curr Opin Chem Biol 17*, 854-859. doi:10.1016/j.cbpa.2013.07.017

Patsch, C., Challet-Meylan, L., Thoma, E. C. et al. (2015). Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. *Nat Cell Biol 17*, 994-1003. doi:10.1038/ncb3205

Persson, M. and Hornberg, J. J. (2016). Advances in predictive toxicology for discovery safety through high content screening. *Chem Res Toxicol 29*, 1998-2007. doi:10.1021/acs.chemrestox.6b00248

Pines, A., Kelstrup, C. D., Vrouwe, M. G. et al. (2011). Global phosphoproteome profiling reveals unanticipated networks responsive to cisplatin treatment of embryonic stem cells. *Mol Cell Biol 31*, 4964-4977. doi:10.1128/MCB.05258-11

Plewczynski, D. and Klingström, T. (2011). GIDMP: Good protein-protein interaction data metamining practice. *Cell Mol Biol Lett 16*, 258-263. doi:10.2478/s11658-011-0004-1

Pognan, F. (2007). Toxicogenomics applied to predictive and exploratory toxicology for the safety assessment of new chemical entities: A long road with deep potholes. *Prog Drug Res 64*, 217-238. doi:10.1007/978-3-7643-7567-6_9

Pognan, F. (2018). Detection, elimination, mitigation, and prediction of drug-induced liver injury in drug discovery. In M. Chen and Y. Will (eds.). *Drug-Induced Liver Toxicity*, 21-43. New York, NY, USA: Humana Press. doi:10.1007/978-1-4939-7677-5_2

Proctor, W. R., Foster, A. J., Vogt, J. et al. (2017). Utility of spherical human liver microtissues for prediction of clinical drug-induced liver Injury. *Arch Toxicol 91*, 2849-2863. doi:10.1007/s00204-017-2002-1

Puigvert, J. C., von Stechow, L., Siddappa, R. et al. (2013). Systems biology approach identifies the kinase Csnk1a1 as a regulator of the DNA damage response in embryonic stem cells. *Sci Signal 6*, ra5. doi:10.1126/scisignal.2003208

Ramirez, T., Daneshian, M., Kamp, H. et al. (2013). t4 report metabolomics in toxicology and preclinical research. *ALTEX 30*, 209-225. doi:10.14573/altex.2013.2.209

Ravagli, C., Pognan, F. and Marc, P. (2017). OntoBrowser: A collaborative tool for curation of ontologies by subject matter experts. *Bioinformatics 33*, 148-149. doi:10.1093/bioinformatics/btw579

Retting, K., Carter, D. E., Crogan-Grundy, C. et al. (2018). Modeling liver biology and the tissue response to injury in bioprinted human liver tissues. *Appl In Vitro Toxicol 4*. doi:10.1089/aivt.2018.0015

Roberts, D. W. and Patlewicz, G. (2018). Non-animal assessment of skin sensitization hazard: Is an integrated testing strategy needed, and if so what should be integrated? *J Appl Toxicol 38*, 41-50. doi:10.1002/jat.3479

Ronaldson-Bouchard, K., Ma, S. P., Yeager, K. et al. (2018). Advanced maturation of human cardiac tissue grown from pluripotent stem cells. *Nature 556*, 239-243. doi:10.1038/s41586-018-0016-3

Roth, A. B. (2017). Setting the scene – The investigative toxicology landscape in the European pharmaceutical industry. In M. Iscan and H. Kandarova (eds.), Abstracts of the 53rd Congress of the European Societies of Toxicology (EUROTOX) Bratislava, Slovakia, 10th-13th September,

2017. *Toxicol Lett 280, Suppl 1*, S76. doi:10.1016/j.toxlet.2017.07.199

Rovida, C., Alépée, N., Api, A. M. et al. (2015). Integrated testing strategies (ITS) for safety assessment. *ALTEX 32*, 25-40. doi:10.14573/altex.1411011

Sacks, L. V., Shamsuddin, H. H., Yasinskaya, Y. I. et al. (2014). Scientific and regulatory reasons for delay and denial of FDA approval of initial applications for new drugs, 2000-2012. *JAMA 311*, 378-384. doi:10.1001/jama.2013.282542

Sager, P. T., Gintant, G., Turner, J. R. et al. (2014). Rechanneling the cardiac proarrhythmia safety paradigm: A meeting report from the cardiac safety research consortium. *Am Heart J 167*, 292-300. doi:10.1016/j.ahj.2013.11.004

Sances, S., Ho, R., Vatine, G. et al. (2018). Human iPSC-derived endothelial cells and microengineered organ-chip enhance neuronal development. *Stem Cell Reports 10*, 1222-1236. doi:10.1016/j.stemcr.2018.02.012

Scaffidi, P., Misteli, T. and Bianchi, M. E. (2002). Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. *Nature 418*, 191-195. doi:10.1038/nature00858

Schadt, H. S., Wolf, A., Pognan, F. et al. (2016). Bile acids in drug induced liver injury: Key players and surrogate markers. *Clin Res Hepatol Gastroenterol 40*, 257-266. doi:10.1016/j.clinre.2015.12.017

Schimek, K., Busek, M., Brincker, S. et al. (2013). Integrating biological vasculature into a multi-organ-chip microsystem. *Lab Chip 13*, 3588. doi:10.1039/c3lc50217a

Schmidt, B. Z., Lehmann, M., Gutbier, S. et al. (2017). In vitro acute and developmental neurotoxicity screening: An overview of cellular platforms and high-throughput technical possibilities. *Arch Toxicol 91*, 1-33. doi:10.1007/s00204-016-1805-9

Sewell, F., Doe, J., Gellatly et al. (2017). Steps towards the International regulatory acceptance of non-animal methodology in safety assessment. *Regul Toxicol Pharmacol 89*, 50-56. doi:10.1016/j.yrtph.2017.07.001

Shuhendler, A. J., Pu, K., Cui, L. et al. (2014). Real-time imaging of oxidative and nitrosative stress in the liver of live animals for drug-toxicity testing. *Nat Biotechnol 32*, 373-380. doi:10.1038/nbt.2838

Skardal, A., Murphy, S. V., Devarasetty, M. et al. (2017). Multi-tissue interactions in an integrated three-tissue organ-on-a-ohip platform. *Sci Rep 7*, 8837. doi:10.1038/s41598-017-08879-x

Smirnova, L., Kleinstreuer, N., Corvi, R. et al. (2018). 3S – Systematic, systemic, and systems biology and toxicology. *ALTEX 35*, 139-162. doi:10.14573/altex.1804051

Soldatow, V. V. Y., Lecluyse, E. L., Griffith, L. G. and Rusyn, I. (2013). In vitro models for liver toxicity testing. *Toxicol Res 2*, 23-39. doi:10.1039/c2tx20051a

Spence, J. R., Mayhew, C. N., Rankin, S. A. et al. (2011). Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature 470*, 105-110. doi:10.1038/nature09691

Steger-Hartmann, T. and Pognan, F. (2018). Improving the safety assessment of chemicals and drug candidates by the integration of bioinformatics and chemoinformatics data. *Basic Clin Pharmacol Toxicol 123, Suppl 5*, 29-36. doi:10.1111/bcpt.12956

Stiehl, D. P., Tritto, E., Dine Chibout, S. et al. (2017). The utility of gene expression profiling from tissue samples to support drug safety assessments. *ILAR J 58*, 69-79. doi:10.1093/ilar/ilx016

Sutherland, J. J., Webster, Y. W., Willy, J. A. et al. (2018). Toxicogenomic module associations with pathogenesis: A network-based approach to understanding drug toxicity. *Pharmacogenom J 18*, 377-390. doi:10.1038/tpj.2017.17

Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell 126*, 663-676. doi:10.1016/j.cell.2006.07.024

Takasato, M., Er, P. X., Chiu, H. S. et al. (2015). Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature 526*, 564-568. doi:10.1038/nature15695

Takebe, T., Sekine, K., Enomura, M. et al. (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature 49*, 481-484. doi:10.1038/nature12271

Tollefsen, K. E., Scholz, S., Cronin, M. T., (2014). Applying adverse outcome pathways (AOPs) to support integrated approaches to testing and assessment (IATA). *Regul Toxicol Pharmacol 70*, 629-640. doi:10.1016/j.yrtph.2014.09.009

Trietsch, S. J., Naumovska, E., Kurek, D. et al. (2017). Membrane-free culture and real-time barrier integrity assessment of perfused intestinal epithelium tubes. *Nat Commun 8*, 262. doi:10.1038/s41467-017-00259-3

Tsaioun, K., Blaauboer, B. J. and Hartung, T. (2016). Evidence-based absorption, distribution, metabolism, excretion (ADME) and its interplay with alternative toxicity methods. *ALTEX 33*, 343-358. doi:10.14573/altex.1610101

Uetrecht, J. (2013). Role of the adaptive immune system in idiosyncratic drug-induced liver injury. In N. Kaplowitz and L. D. DeLeve (eds.), *Drug-Induced Liver Disease* (Chapter 11, 175-193). 3rd edition. New York, USA: Elsevier. doi:10.1016/B978-0-12-387817-5.00011-X

Urbisch, D., Mehling, A., Guth, K. et al. (2015). Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul Toxicol Pharmacol 71*, 337-351. doi:10.1016/j.yrtph.2014.12.008

Uteng, M., Germano, D., Balavenkatraman, K. K. et al. (2014). High content imaging approaches for in vitro toxicology. In A. Bal-Price and P. Jennings (eds.), *In Vitro Toxicology Systems. Methods in Pharmacology and Toxicology* (377-397). New York, NY, USA: Humana Press. doi:10.1007/978-1-4939-0521-8_17

Valentin, J. P., Pollard, C., Lainée, P. and Hammond, T. (2010). Value of non-clinical cardiac repolarization assays in supporting the discovery and development of safer medicines. *Br J Pharmacol 159*, 25-33. doi:10.1111/j.1476-5381.2009.00530.x

van Duinen, V., van den Heuvel, A., Trietsch, S. J. et al. (2017). 96 perfusable blood vessels to study vascular permeability in vitro. *Sci Rep 7*, 18071. doi:10.1038/s41598-017-14716-y

van Hasselt, J. G. C. and Iyengar, R. (2017). Systems pharmacology-based identification of pharmacogenomic determinants of adverse drug reactions using human iPSC-derived cell lines. *Curr Opin Systems Biol 4*, 9-15. doi:10.1016/j.coisb.2017.05.006

van Vliet, E., Daneshian, M., Beilmann, M. et al. (2014). Current approaches and future role of high content imaging in safety sciences and drug discovery. *ALTEX 31*, 479-493. doi:10.14573/altex.1405271

Vernetti, L. A, Senutovitch, N., Boltz, R. et al. (2016). A human liver microphysiology platform for investigating physiology, drug safety, and disease models. *Exp Biol Med 241*, 101-114. doi:10.1177/1535370215592121

Vernetti, L. A., Vogt, A., Gough, A. and Taylor, D. L. (2017). Evolution of experimental models of the liver to predict human drug hepatotoxicity and efficacy. *Clin Liver Dis 21*, 197-214. doi:10.1016/j.cld.2016.08.013

Villeneuve, D. L., Crump, D., Garcia-Reyero, N. et al. (2014). Adverse outcome pathway (AOP) development I: Strategies and principles. *Toxicol Sci 142*, 312-320. doi:10.1093/toxsci/kfu199

Wagner, I., Materne, E.-M., Brincker, S. et al. (2013). A dynamic multi-organ-chip for long-term cultivation and substance testing proven by 3D human liver and skin tissue co-culture. *Lab Chip 13*, 3538. doi:10.1039/c3lc50234a

Wallace, T. M., Levy, J. C. and Matthews, D. R. (2004). Use and abuse of HOMA modeling. *Diabetes Care 27*, 1487-1495. doi:10.2337/diacare.27.6.1487

Wallis, R., Benson, C., Darpo, B. et al. (2018). CiPA challenges and opportunities from a non-clinical, clinical and regulatory perspectives. An overview of the safety pharmacology scientific discussion. *J Pharmacol Toxicol Methods 93*, 15-25. doi:10.1016/j.vascn.2018.06.005

Ward, J., Kanchagar, C., Veksler-Lublinsky, I. et al. (2014). Circulating microRNA profiles in human patients with acetaminophen hepatotoxicity or ischemic hepatitis. *Proc Natl Acad Sci U S A 111*, 12169-12174. doi:10.1073/pnas.1412608111

Waring, M. J., Arrowsmith, J., Leach, A. R. et al. (2015). An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat Rev Drug Discov 14*, 475-486. doi:10.1038/nrd4609

Weaver, R. J. (2001). Assessment of drug-drug interactions: Concepts and approaches. *Xenobiotica 31*, 499-538. doi:10.1080/00498250110060950

Weaver, R. J., Betts, C., Blomme, E. A. G. et al. (2017). Test systems in drug discovery for hazard identification and risk assessment of human drug-induced liver injury. *Expert Opin Drug Metab Toxicol 13*, 767-782. doi:10.1080/17425255.2017.1341489

Webster, R., Leishman, D. and Walker, D. (2002). Towards a drug concentration effect relationship for QT prolongation and torsades de pointes. *Curr Opin Drug Discov Devel 5*, 116-126.

Wevers, N. R., van Vught, R., Wilschut, K. J. et al. (2016). High-throughput compound evaluation on 3D networks of neurons and glia in a microfluidic platform. *Sci Rep 6*, 38856. doi:10.1038/srep38856

Wink, S., Hiemstra, S., Huppelschoten, S. et al. (2014). Quantitative high content imaging of cellular adaptive stress response pathways in toxicity for chemical safety assessment. *Chem Res Toxicol 27*, 338-55. doi:10.1021/tx4004038

Wink, S., Hiemstra, S., Herpers, B. and van de Water, B. (2017). High-content imaging-based BAC-GFP toxicity pathway reporters to assess chemical adversity liabilities. *Arch Toxicol 91*, 1367-1383. doi:10.1007/s00204-016-1781-0

Wink, S., Hiemstra, S. W., Huppelschoten, S. et al. (2018). Dynamic imaging of adaptive stress response pathway activation for prediction of drug induced liver injury. *Arch Toxicol 92*, 1797-1814. doi:10.1007/s00204-018-2178-z

Xiao, S., Coppeta, J. R., Rogers, H. B. et al. (2017). A microfluidic culture model of the human reproductive tract and 28-day menstrual cycle. *Nat Commun 8*, 4584. doi:10.1038/ncomms14584

Yu, J., Vodyanik, M. A., Smuga-Otto, K. et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science 318*, 1917-1920. doi:10.1126/science.1151526

Zhang, M., Chen, M. and Tong, W. (2012). Is toxicogenomics a more reliable and sensitive biomarker than conventional indicators from rats to predict drug-induced liver injury in humans? *Chem Res Toxicol 25*, 122-129. doi:10.1021/tx200320e

Zhang, Y. S., Aleman, J., Arneri, A. et al. (2015). From cardiac tissue engineering to heart-on-a-chip: Beating challenges. *Biomed Mater 10*, 034006. doi:10.1088/1748-6041/10/3/034006

Conflict of interest

Acknowledgements