



Setting the stage for next-generation risk assessment with non-animal approaches: the EU-ToxRisk project experience

M. J. Moné¹ · G. Pallocca² · S. E. Escher³ · T. Exner⁴ · M. Herzler⁵ · S. Hougaard Bennekou⁶ · H. Kamp⁷ · E. D. Kroese⁸ · Marcel Leist^{2,9} · T. Steger-Hartmann¹⁰ · B. van de Water¹

Received: 21 July 2020 / Accepted: 12 August 2020 / Published online: 4 September 2020
© The Author(s) 2020

Abstract

In 2016, the European Commission launched the EU-ToxRisk research project to develop and promote animal-free approaches in toxicology. The 36 partners of this consortium used *in vitro* and *in silico* methods in the context of case studies (CSs). These CSs included both compounds with a highly defined target (e.g. mitochondrial respiratory chain inhibitors) as well as compounds with poorly defined molecular initiation events (e.g. short-chain branched carboxylic acids). The initial project focus was on developing a science-based strategy for read-across (RAx) as an animal-free approach in chemical risk assessment. Moreover, seamless incorporation of new approach method (NAM) data into this process (=NAM-enhanced RAx) was explored. Here, the EU-ToxRisk consortium has collated its scientific and regulatory learnings from this particular project objective. For all CSs, a mechanistic hypothesis (in the form of an adverse outcome pathway) guided the safety evaluation. ADME data were generated from NAMs and used for comprehensive physiological-based kinetic modelling. Quality assurance and data management were optimized in parallel. Scientific and Regulatory Advisory Boards played a vital role in assessing the practical applicability of the new approaches. In a next step, external stakeholders evaluated the usefulness of NAMs in the context of RAx CSs for regulatory acceptance. For instance, the CSs were included in the OECD CS portfolio for the Integrated Approach to Testing and Assessment project. Feedback from regulators and other stakeholders was collected at several stages. Future chemical safety science projects can draw from this experience to implement systems toxicology-guided, animal-free next-generation risk assessment.

M. J. Moné and G. Pallocca contributed equally to this work.

Disclaimer This work reflects only the authors' view and the European Commission is not responsible for any use that may be made of the information it contains.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00204-020-02866-4>) contains supplementary material, which is available to authorized users.

✉ Marcel Leist
marcel.leist@uni-konstanz.de

¹ Leiden Academic Centre for Drug Research, Leiden University, Leiden, The Netherlands

² CAAT-Europe at the University of Konstanz, Constance, Germany

³ Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM), Hannover, Germany

⁴ Edelweiss Connect GmbH, Basel, Switzerland

⁵ German Federal Institute for Risk Assessment (BfR), Berlin, Germany

⁶ Technical University of Denmark, Kongens Lyngby, Denmark

⁷ BASF SE, Ludwigshafen, Germany

⁸ TNO Innovation for Life, Utrecht, The Netherlands

⁹ In Vitro Toxicology and Biomedicine, Department Inaugurated By the Doerenkamp-Zbinden Foundation at the University of Konstanz, University of Konstanz, 78457 Constance, Germany

¹⁰ Investigational Toxicology, Bayer AG, Pharmaceuticals, Berlin, Germany

Introduction

A particular focus of EU-ToxRisk is on endpoints for repeated-dose (RDT) and developmental and reproductive toxicity (DART), and on providing guidance for practical implementation (Daneshian et al. 2016). This entails the implementation of NAM-based safety science into regulatory toxicology (Krebs et al. 2020b). The project initially focused on read-across (RAX) (Escher et al. 2019; Rovida et al. 2020a) and here we outline key outcomes and general learnings on the use of NAMs for RAX. This collection of outputs will facilitate new efforts in next-generation risk assessment (NGRA) (Dent et al. 2018; Dearfield et al. 2017; Cote et al. 2016; Krewski et al. 2014) and thus advance the field of safety assessment by more mechanistically driven and animal-free approaches (Suppl. Box 1). This report was assembled once an important project phase came to its conclusion. We will accentuate especially learnings on the regulatory implementation of new approaches. Moreover, we highlight some gaps and unresolved issues that will need future attention.

Incorporation of NAMs in read-across approaches

The main activity of the project addressed the application of NAMs to support the most widely applied method of data-gap filling: RAX. RAX may support safety assessment for the complex toxicological endpoints RDT and DART, yet it often fails to be accepted by regulatory agencies. Reasons for rejection often include the notion that associated uncertainties are perceived as being too large. Classical RAX is based on structural similarity as the major argument, and

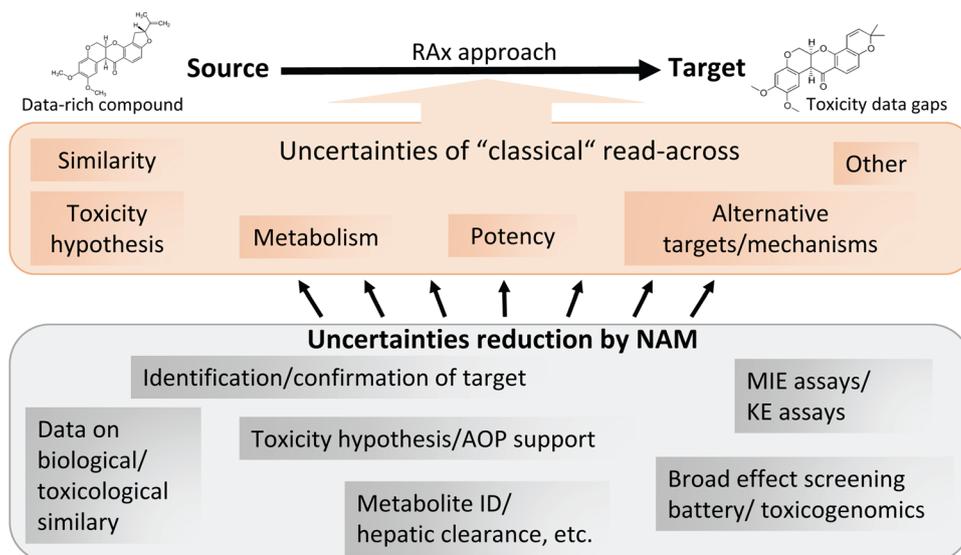
there are many examples that chemical formula alone does not sufficiently predict toxicity. This uncertainty can be substantially reduced when NAMs demonstrate that source and target chemicals share similar biological and toxicokinetic profiles, providing justification for reading across source chemicals' in vivo data.

The project hypothesis was that NAMs can reduce uncertainty in traditional RAX (Fig. 1). For instance, they help characterize the biological properties of source and target compounds. NAMs can point out a specific molecular initiation event [e.g. receptor (ant)agonism], provide data on test compound hazard (i.e. expected types of adverse outcomes), elucidate a mode of action (i.e. pathways and targets affected), or assess relative potencies of the observed effects. Also, the absence of a certain mechanism or effect can be demonstrated, as well as the relatively low potency for a certain testing endpoint. The EU-ToxRisk RAX testing framework demonstrated how NAMs can be integrated into practical RAX procedures (Escher et al. 2019).

If the AOP for a set of chemicals is known, NAM testing can be designed along this mechanistic framework to explore molecular initiating events (MIEs) or key event (KE) responses. Targeted testing aims to generate mechanism-related data for all grouped compounds, to either confirm (dis)similarity or observe a consistent trend. If no AOP is known, the battery of NAMs can be chosen in a way to capture as many as possible potential underlying mechanisms. NAMs can alternatively be used for broad untargeted testing to either (1) generate a RAX hypothesis based on shared in vitro effects, or (2) to prove the absence of effects (up to concentrations corresponding or exceeding those obtained in humans under realistic exposure situations).

Besides resolving many conceptual uncertainties, NAM data, like any other type of experimental result, may

Fig. 1 Reduction of RAX uncertainties by NAMs. The top box lists several uncertainties that may weaken a RAX approach (e.g. uncertainty on the similarity of source and target, uncertainty on metabolism, or uncertainty concerning the potency ratio of source and target). The bottom box indicates information that can be provided by NAMs to reduce uncertainty (e.g. data on potency ratios in key event (KE) or molecular initiation event (MIE) assays, or the identification (ID) of relevant metabolites formed)



introduce their own uncertainty (assay variations, random noise, and congruence of different tests). This issue can be addressed by applying a structured and weighted description of uncertainties. This may use generalized Bayesian approaches or Dempster–Shafer decision theory (Dempster 1967; Rathman et al. 2018; Shafer 1976), or it may use semi-quantitative approaches based on expert judgment (Escher et al. 2019; Schultz et al. 2018). For instance, the Dempster–Shafer approach allows for a fully quantitative combination of various types of test data, taking into account the individual test performances/uncertainties, and to derive likelihoods of test data being correct. EU-ToxRisk puts significant emphasis on diligent uncertainty descriptions and in taking initial steps towards uncertainty quantification. More resources will need to be put in place to apply these concepts in a more routine manner, and the science of uncertainty assessment is likely to become an important support discipline of risk assessment.

Brief overview of case studies

Some of the project CSs were selected to test regulatory applicability of NAM-supported RAx. The design of a case study implied (1) the definition of a target chemical and a regulatory question, and (2) the delineation of data gaps and the approaches to fill them. A detailed overview and flow-chart have been presented elsewhere (Graepel et al. 2019; Escher et al. 2019).

CSs were based on assemblies of structurally similar compounds to demonstrate how NAMs can be used to substantiate a RAx hypothesis. CSs always contained some chemical analogues with *in vivo* endpoint data, so that predictivity and accuracy of the NAM data could be verified. For the same reason, these CSs also included structurally similar compounds that did not exhibit a shared toxicological effect patterns/AOPs in the *in vivo* data determining the RAx hypothesis. NAM data were subsequently used to better define the boundaries of the categories and to reduce uncertainties. The description of the individual CSs is beyond the scope of this editorial and will be highlighted in individual publications. Here we specifically address the overall learnings and impact of the CSs for RAx and their relevance for future next-generation risk assessment.

To illustrate the CS concept, we will describe a typical example: a CS focusing on the substance deguelin as a potential inducer of parkinsonian motor deficits. An AOP-based testing strategy was applied to support the RAx hypothesis. The AOP for inhibition of mitochondrial

complex I of nigrostriatal neurons leading to parkinsonian motor deficits (Terron et al. 2018) was used to assess the biological similarity of two different rotenoids: rotenone (source compound, data-rich) and deguelin (target compound, data-poor). For the source compound, there was compelling evidence for the induction of Parkinson's-like disorders in experimental models, and also epidemiological evidence for similar effects in human (Ntzani et al. 2013). Various NAMs were selected to represent the MIE and KEs of the AOP. Both, *in silico* and *in vitro* test methods were applied. For instance, *in silico* structural modelling was used to identify the common pharmacophore that determines the binding to complex I, being the MIE of the AOP. To determine effects on the KEs, several *in vitro* methods were used to detect effects on mitochondrial respiratory activity (oxygen consumption) and on proteostasis. High-content imaging methods were applied for measuring damage to human dopaminergic neurons, and to compare the effects of the two compounds. Finally, PBK modelling was performed to compare biokinetics and distribution of both compounds to the brain. The approach chosen in this case study demonstrated that an AOP-based testing strategy combining different test methods that cover the various KEs can be applied in supporting chemical safety assessment within a NAM-based RAx framework.

Feedback by the Regulatory Advisory Board

Given the overall EU-ToxRisk objectives to provide solutions for implementing NAMs in risk assessment, the NAM-enhanced RAx case studies were presented and discussed on several occasions for feedback from the regulatory community. For this purpose, the project's Scientific Advisory Board initiated close collaborations with regulators from national, European and international regulatory authorities, including BfR and RIVM, ECHA, EFSA and EMA, and NIEHS, respectively. This cooperation resulted in an improved mutual understanding of the requirements and pitfalls of RAx approaches supported by NAMs, from a scientific, academic, and regulatory perspective. The project's Regulatory Advisory Board (representing the main agencies/authorities) was set up in the course of the study as one of the first reactions to stakeholder feedback. It has since been key for the establishment of a reporting template for RAx cases. The extensive documentation and pre-validation of the EU-ToxRisk NAM toolbox is a step forward to increase the acceptability of regulatory dossiers based on NAM data. Overall regulatory learnings are summarized in Box 1.

Box 1: Regulatory learnings based on the NAM-based RAX framework assessment

I. Involve regulators from the start in the design of CSs so that they revolve around precisely phrased regulatory questions.

II. Meticulously define both regulatory context and hypothesis.

III. Provide a rationale for each step in the risk assessment argument that regulators can follow and thus evaluate.

IV. Provide a clear learning scenario. Sharply define the situation before the RAX and before the use of NAMs. Evaluate the situation afterwards and provide transparent conclusions of what each method contributed.

V. Compare established risk assessment approaches with NAM-based RAX approaches and identify advantages and shortcomings together with needs for future implementation training.

VI. Close interaction with the wider community of regulatory stakeholders is crucial to the establishment of the procedures. “Getting the science right” is very important, as is “getting the framing and reporting right” to match expectations and needs of the regulatory target audience.

Collected feedback from stakeholders and key opinion leaders on the EU-ToxRisk NAM-supported RAX framework

Following internal regulatory scrutiny, the project CSs and approaches were discussed in external forums. Learnings, achievements, and pitfalls of the approach were also analysed in detail in a dedicated workshop organized by EU-ToxRisk in May 2019 in Espoo, Finland. Over 60 international experts from industry, academia, and regulatory authorities explored five scientifically advanced project case studies developed within EU-ToxRisk, the OECD/Integrated Approach to Testing and Assessment (IATA) program, and NIHS Japan. The wide range of expertise and perspectives of the participants yielded valuable insights and answers to where and how NAMs can effectively support a RAX problem formulation, identify areas that contain data gaps, suggest how these gaps can be filled, and finally provide technical guidance on how those steps should be presented in a RAX regulatory dossier. Reports of three of the most mature EU-ToxRisk case studies, describing the RAX question, the approach taken (i.e. what NAM to select), the results obtained, and the conclusions derived, were further shared with the regulatory community for their review and feedback (Box 2).

Box 2: Learnings from stakeholder feedback

Scientific

I. Relevant aspects, such as xenobiotic metabolism and chemical or biological similarity considerations, should be addressed all at the same time to ensure a holistic coverage of potential contexts. Too much focus on one particular similarity context bears the danger of missing/neglecting important effects and relevant context.

II. The regulatory question is important for determining the scope of the scientific approach. In some cases, all potential toxicities (or metabolites, or aspects of the chemical structure) may need to be considered. In other cases, it may be well acceptable to focus only on one relevant toxicity (e.g. neurotoxicity) or on one specific toxicophore.

III. Anchoring a toxicity hypothesis to an AOP is desirable. If not available, one may still perform a weight-of-evidence approach using, for instance, a battery of NAM tests.

Regulatory

I. Qualitative expression of uncertainty is to be done in explicit categories (low/medium/high). Non-quantifiable uncertainties have to be made transparent. All underlying uncertainties of the assessment need to be covered, including category formation, metabolism predictions, choice of NAMs, etc.

II. Focus and limit testing. The goal should be to do as much testing as needed to achieve sufficient confidence in a RAX justification, but not more than that.

Case study reviews by the OECD IATA Case Studies Project group

In parallel, these same three project case studies plus an additional one were also submitted to the OECD for review in its 5th Meeting of the IATA Case Studies Project under the remit of the Working Party on Hazard Assessment. The IATA Case Studies Project aims to assess the practical applicability of NAMs, as a part of the IATA framework, for different aspects of regulatory decision-making to build assessment experience (OECD 2015). The submitted EU-ToxRisk project CSs underwent rigorous review by risk assessors from several countries, the OECD, and the International Council on Animal Protection in OECD (ICAPO). The reviewing had a broad scope, assessing not only the data but also identifying where further guidance would be helpful, e.g. the regulatory relevance and the reporting. The OECD consideration document is set to be published adjacent to the CSs after their finalization based on comments received. Final endorsement by the Working Party on

Hazard Assessment and the Joint Meeting members allows for their publication in September 2020. This represents a major milestone toward regulatory recognition of NAM-based RAX (Box 3).

Box 3: Learnings from OECD feedback

Scientific

- I. Provide argumentation for the selection of all compounds and the rationale for their selection (or exclusion).
- II. It is advantageous to quantitatively extrapolate NAM-based data to the (human) *in vivo* situation.
- III. A good toxicity hypothesis (e.g. AOP-based) supported by NAM data can serve as justification for RAX.
- IV. Inclusion of high-content data (e.g. toxicogenomics) can increase overall confidence (e.g. of not missing important adverse effects).

Regulatory

- I. Due to the lower level of validation of and experience with many NAMs in a regulatory context, it has to be ensured that recipients of RAX reports can understand all data provided, especially when methods are more complex and less well described in the literature.
- II. Strengths and weaknesses, as well as rigorous uncertainty assessment, should be clearly addressed in the reporting.

EU-ToxRisk advisory document on the proposed read-across framework

From the three different case study reviewing phases, it became evident that RAX justifications should be presented as logically structured, coherent “stories”, along the “as comprehensive as needed, as concise as possible” principle.

Here the need for an advisory document that would be complementary to RAAF (ECHA 2017) and the OECD guidance document for *in vitro* test method descriptions (OECD 2014), addressing regulatory acceptance from the point of view of registrants of NAM-supported RAX dossiers. The Read-Across Advisory Report was structured back-to-back with the EU-ToxRisk overall RAX framework (Escher et al. 2019) and includes common learnings from the above discussion.

The advisory document targets the broader toxicology community with practical instructions on the applications of NAM-based RAX in different regulatory contexts, including all relevant collected feedback and endorsements on taken approaches. Its application will improve the submission quality of RAX cases by registrants and thereby increase successful acceptance rates of non-animal approaches. For this purpose, an important component of the template already

now is guidance on how to fill out the different sections of the risk assessment reporting submission document—especially with NAM data. A web-based graphical user interface (GUI) is being produced to facilitate its use.

The repository of NAMs available for NAM-enhanced RAX

EU-ToxRisk started off with many NAMs available at partner laboratories. They were selected and refined with the purpose of ensuring applicability in different regulatory contexts. Based on the initial toolbox, existing test methods were continuously refined to enhance their applicability based on two parameters: test system complexity and test method throughput (Fig. 2). Simpler test systems were combined to obtain advanced complex test systems to better resemble the *in vivo* situation. These approaches include, for instance, the use of mixed cell cultures (Gutbier et al. 2018), 3D organoids and neurospheres (Brull et al. 2020; Hiemstra et al. 2019; Kobolak et al. 2020) and microphysiological systems (MPS), e.g. four-organ-chips developed to interconnect miniaturized human intestine, liver, brain and kidney equivalents (Ramme et al. 2019) and improved iPSC cell differentiation to hepatocytes (Boon et al. 2020). These advancements were fuelled by the establishment of novel differentiation protocols that produce high-quality cellular test systems (Ballester et al. 2019; Coll et al. 2018; Dreser et al. 2020; Gu et al. 2018; Rauch et al. 2018).

In parallel, the availability of advanced high-throughput approaches was exploited to increase the applicability domains and throughput levels of the test methods. Screening transcriptomics data were provided from hepatic (Albrecht et al. 2019; Campos et al. 2020; Copple et al. 2019; Ramirez et al. 2018), renal (Limonciel et al. 2018a, b) and neuronal (Delp et al. 2019) cell systems. High-content imaging was expanded to complex toxicological endpoints, such as DART (Dreser et al. 2020; Nyffeler et al. 2018) and systemic RDT with the establishment of novel fluorescent protein reporter cell lines (Bischoff et al. 2019; Schimming et al. 2019; Wink et al. 2018; Yang et al. 2020).

The toolbox was further enriched by novel *in silico* approaches, supporting and complementing *in vitro* cell-based NAMs. An important area here was the extrapolation of NAM concentrations to *in vivo* doses (IVIVE, PBPK) (Fisher et al. 2019; Simeon et al. 2020; Toma et al. 2018). Additionally, novel QSAR tools and machine learning approaches were put in place to assess chemical similarity among test compounds (Gadaleta et al. 2018a, b; Hemmerich et al. 2020; Luechtefeld et al. 2018; Toropova et al. 2018; Toropov and Toropova 2017; Troger et al. 2020) and to predict toxicological properties from their structure.

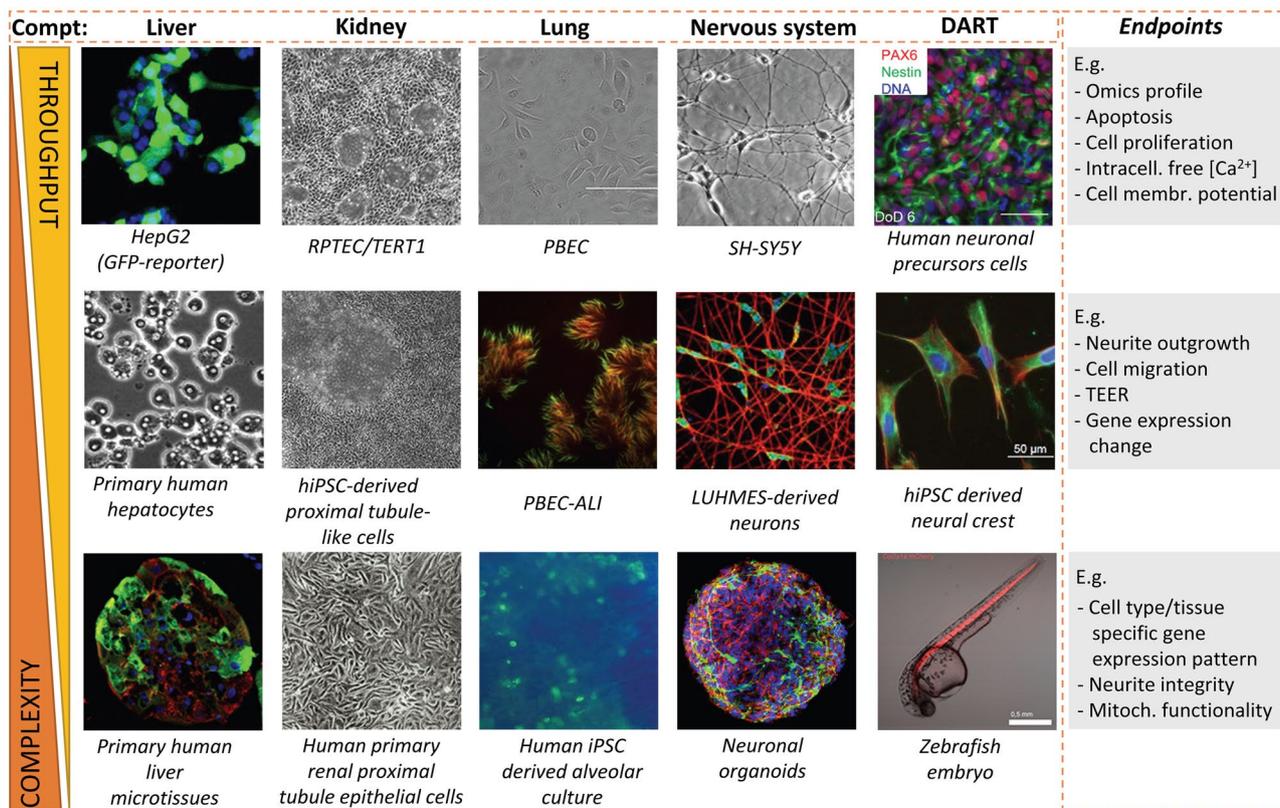


Fig. 2 The EU-ToxRisk method toolbox. The EU-ToxRisk method toolbox includes test systems using cells from four major organ systems. In addition, models allowing readouts on DART have been included. The toolbox comprises both simpler cell models (2D, monoculture, etc.) and complex systems (3D, co-cultures, zebrafish embryos). The characteristics of each model determine their through-

put and their use at different stages of case studies. For each test system, various endpoints have been established, so that assays can be run to assess effects on cell viability in parallel with functional, biochemical, and toxicogenetics endpoints. Figure adapted from the EU-ToxRisk project's website (https://www.eu-toxrisk.eu/page/media_items/test-methods8.php)

Finally, the project defined new AOPs to seamlessly integrate data generated by alternative methods or in vivo testing in a mechanistic and quantitative manner (Maertens et al. 2018; Terron et al. 2018; Zgheib et al. 2019).

All this work contributed not only to the regulatory implementation of NAM-enhanced RAX, but also presented a scientific advance recognized by the academic community. This output is exemplified here by a selection of publications according to toxicological topics (Supp. Table 1), and high-ranking lists of papers in terms of citation and journal impact factor (Supp. Table 2; Supp. Table 3).

Knowledge infrastructure: sharing of data, test methods, results and knowledge

The scientific progress described above opens up new possibilities for reducing uncertainties associated with RAX and other risk assessment approaches. Irrespective, there was valuable feedback from regulatory specialists on the relative importance of high-end science vs the more basic quality

control and documentation aspects. An important notion was that it is critical for regulators to quickly understand a new method; its information value, its purpose and its quality status. Similar considerations applied to data generated via NAMs, and it was enlightening to many contributing partners that the level of information sufficing for publication did not scale with regulatory requirements. The project team responded to this essential notion from the Regulatory Advisory Board by ensuring that all NAMs are described in a maximally transparent way to enable regulatory toxicologists to assess quality, reliability, applicability, and relevance of method outputs (Pamies et al. 2020; Bal-Price et al. 2018; Krebs et al. 2019). This approach was addressed by the use of comprehensive test method descriptions that also allow for the assessment of readiness levels of individual tools, combined with an efficient procedure for data upload (Krebs et al. 2020b). An efficient data management system was developed as a hub to transparently enable connecting methods to data, integrating data derived from different methods, and linking the integrated data to confirm or reject an overall testing hypothesis (like RAX).

In the context of a cross-system testing case study, detailed protocols for FAIR data handling were established, including the provision of sufficiently rich and transparent metadata explaining how data was produced. This includes the experimental setup and the various processing steps from raw to summary data such as benchmark concentration (Krebs et al. 2020a) or no-observed-adverse-effect-levels (NOAELs) used in risk assessment. EU-ToxRisk, together with its stakeholders, created test method descriptions and provided guidance on data processing (Krebs et al. 2018). A knowledge infrastructure was programmed that allows for efficient management of this information, using structured but still flexible input. The starting point is a detailed test method description including a standard operating procedure (SOP), and also more specific details on the validation status and relationship of the assay to AOPs.

The developed test method documentation (ToxTemp) covers mostly transparency and validation aspects. It was designed in alignment with the OECD Guidance Document 211, which provides a template for “descriptions of non-guideline in vitro methods”. The many types of essential information can be grouped into four sections: (1) the overall test method description, (2) the technical test procedure (as outlined in an SOP, e.g. defined labware, consumables and pipetting steps), (3) the characterization of test and reference materials/chemicals, and (4) all issues relating to data processing and archiving. Finally, additional paragraphs address the test purpose, the test limitations (like information on its applicability), and the criteria to be used for interpreting test results. ToxTemp was endorsed by the project’s external advisory boards and over 30 experts from industry, regulatory bodies, and academia (Krebs et al. 2019). The public version of the ToxTemp method repository is available at <https://eutoxrisk.edelweissconnect.com/public/>. It is additionally used to provide metadata for the associated datasets as computer-readable annotations.

All EU-ToxRisk metadata linked to all project data output will become available from the BioStudies database at the European Bioinformatics Institute (EBI) once the data have been publicly released (Sarkans et al. 2018). The EU-ToxRisk knowledge-sharing platform accesses data from there and provides it to the consortium in a structured way for searching, browsing, visualization, and modelling purposes and, ultimately, for decision-making in risk assessment as exemplified in project case studies (Fig. 3). Collecting test method information and metadata can be a time-consuming and error-prone task if done manually. Therefore, automated procedures were designed to collect data and metadata from experimental equipment and processing software and provide automated file validation tools to check data consistency, complemented by human curation by multiple individuals as a final quality check (Box 4).

Box 4: Learnings on data management

- I. Use harmonized templates for data and methods right from the project onset.
- II. Invest time and resources into early training of researchers on data handling and management.
- III. Provide and flexibly adapt data conversion and input tools.

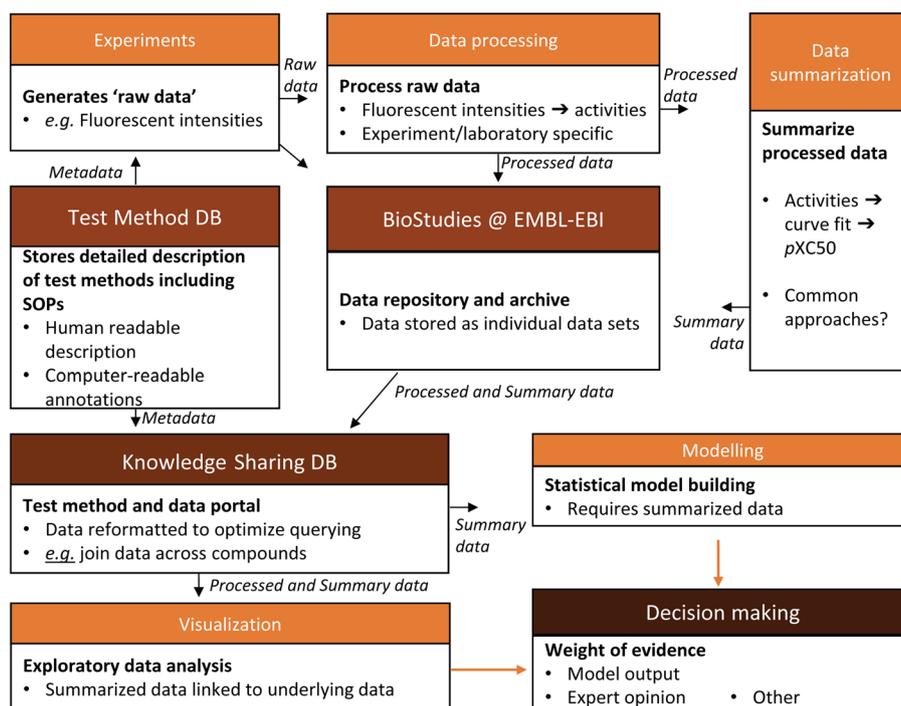
Gaps and future requirements

Despite significant progress, there are still gaps that need to be addressed before standard risk assessment can be performed entirely with animal-free methods. A comprehensive gap analysis would require an extensive vision document on the future of toxicology. This is beyond the scope of this short overview, but for an overview of past and current ideas, we refer to various visionary reports and documents (Cote et al. 2016; Hartung and Leist 2008; Khadka et al. 2020; Krewski et al. 2020; Leist et al. 2014, 2008; Lupu et al. 2020; Sauer et al. 2015; Thomas et al. 2019, 2018).

Here, we outline nine key aspects that need further attention. This itemized summary may help in structuring work packages and defining strategies of future large risk assessment projects (Fig. 4):

1. *Method shop* The width of in vitro and in silico NAMs—and of strategies using them—is very large compared to traditional animal models. New concepts and business models are required to make broad panels of new models available to all stakeholders.
2. *Human anchoring* Method validation has traditionally relied on the correlation of NAM data with animal study data. Even though animal data was useful for the development of NAMs, they also have shortcomings towards chronic/long-term human health consequences. Therefore, the development and calibration of NAMs should also be anchored to human (patho)biology and not be calibrated exclusively by animal data.
3. *Metabolism and transport* More quantitative methods are required to predict the influence of metabolism and transport on the hazard of novel compounds.
4. *Ab initio assessment* Strategies need to be worked out to quickly arrive at a mechanistic hypothesis for the potential hazard of compounds with little or no prior safety information.
5. *Non-toxicants* A comprehensive strategy needs to be established on how to define compounds of low or no toxicity. This also requires measures for uncertainty. In particular, the risk of false-negative assessment needs to be evaluated.

Fig. 3 The EU-ToxRisk knowledge-sharing infrastructure. Data and metadata follow different flows to be processed and deposited into the respective data infrastructure. Raw data are processed and summarized before being deposited into the Biostudies database at EMBL-EBI. Metadata are generated for each data set and deposited together into the knowledge-sharing database run by Edelweiss Connect. The coupled information is used for statistical model building and can be explored via visualization tools. The collection of information allows for weight-of-evidence-based decision-making processes



6. *Immune system* Both native and adaptive immune responses can be involved in toxicity (Benedetti et al. 2013; Fasbender et al. 2020; Fredriksson et al. 2011, 2014; Leist et al. 1997; Monshi et al. 2013). This notion needs further exploration to predict idiosyncratic reac-

tions and chronic health consequences for more susceptible individuals.

7. *Protection perspective* Various past efforts attempted to substitute traditional hazard assessment with modern NAMs, following the concept that defined adverse

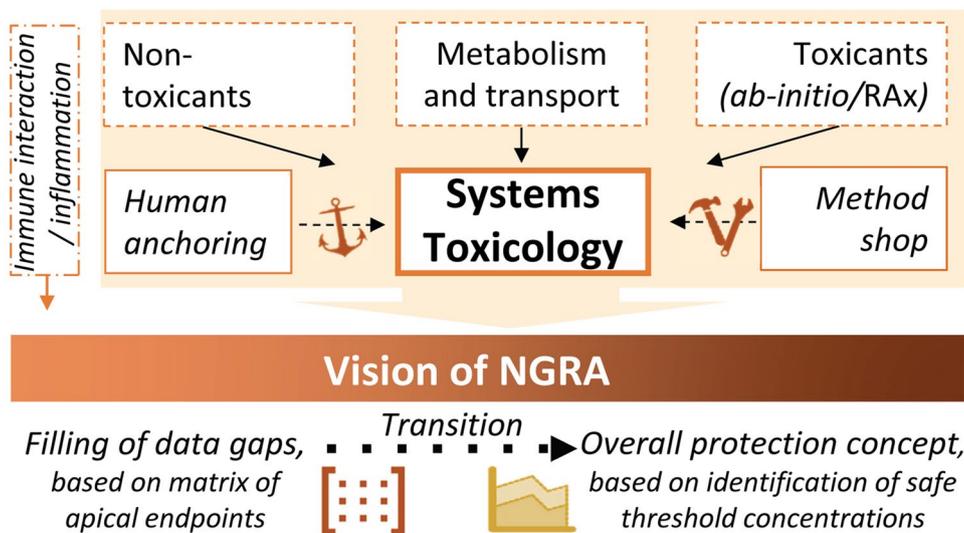


Fig. 4 Elements of next-generation risk assessment (NGRA) framework. NGRA should see the shift from mainly using NAMs for filling demarcated data gaps to a human-centric overall protection concept underpinned by NAM-based hazard quantification. Immune responses need to be taken into account to predict idiosyncratic reactions and chronic health consequences. To quantify risk and while accounting for associated uncertainties, NGRA will rely on integra-

tive systems toxicology-based modelling approaches, anchoring safety testing to in vivo human biology. Considering metabolism and transport early on will be key to arrive at hazard scenarios for new chemical compounds. To cater for the wider field of applied toxicology, NGRA should be developed with both toxic and non-toxic, and both data-rich and data-poor compounds in mind

events need to be identified and the subsequent animal-based method to be substituted one-on-one by NAM(s). This strategy will fail for complex endpoints. It is even doubtful whether such an approach can cover all potential target organs and tissues in the human body, be it practically or at an acceptable cost. A future safety science concept may require a more revolutionary approach that is not anchored to traditional adverse outcomes, but rather to the identification of safe concentrations. Staying below such concentrations would protect the population from toxic effects by ensuring that no biological events linked to toxicity are activated. While some CSs of EU-ToxRisk started exploring this concept, future projects need to address this systematically.

8. *Systems toxicology* Points 3–6 need to be successfully incorporated into a comprehensive systems toxicology model, which will then eventually form the basis for point 9. This will involve outputs not only on concentration thresholds associated with hazard (and extrapolations to respective doses/exposures), but also quantify the associated uncertainties.
9. *Next-generation risk assessment* Similar to the work of EU-ToxRisk on hazard assessment, the next step is to progress towards NGRA, using NAM-based hazard quantification in combination with defined exposure scenarios and exposure models.

An important future step to increase the impact of EU-ToxRisk approaches will be the uptake of its tools and strategies in upcoming initiatives (e.g. the Horizon Europe Partnership for the Assessment of Risk from Chemicals; PARC) and in collaboration with the private sector. Already some CSs in partnership with external industry stakeholders were initiated. Such joint CSs help to (1) inform strategic decisions, (2) prioritize chemicals within a group, and (3) support the problem-solving process in investigative toxicology (Beilmann et al. 2019). Furthermore, the sustainability of the approaches will be supported by a commercialization platform that consolidates NAMs and NAM data from different partner organizations, and use the integrated results in both safety assessment and investigative toxicology. In this context, activities aiming to increase the international acceptance of the NAMs are crucial. This requires continued interactions amongst all relevant stakeholders groups, including industry, regulatory agencies, academia, policy-makers, and NGOs (Busquet et al. 2020). For this reason, EU-ToxRisk provided many discussion platforms in the form of stakeholder meetings, conference sessions, and closed workshops with regulators and parallel projects. Recent examples of such activities are the workshops on microphysiological systems and their regulatory application (Marx et al. 2020) and on international acceptance of read-across approaches (Rovida et al. 2020a, b) (Box 5).

Box 5: General learnings for future risk assessment projects:

- I. Establish a Regulatory Advisory Board right from the onset (application phase).
- II. Structure the project along CSs rather than work packages.
- III. Anchor each CS to a strictly defined regulatory question.
- IV. Use fit-for-purpose and well-structured risk assessment reporting templates (compatible with regulatory requirements).
- V. Ensure a strong emphasis on data management and establish consensus method description templates and data upload templates and provide a project data integration platform.
- VI. Build and maintain networks with other toxicology programs. Maintain good personal relationships. Joint events and activities are necessary to make exchanges work, so reserve budget and resources to this end.
- VII. For projects with a long runtime (≥ 5 years), allow for some re-organization possibilities and the realignment of scientific goals as technologies progress and regulatory practice evolves.
- VIII. Reserve budget for test method validation and sustainability measures.
- IX. For all CSs, establish starting knowledge (pre-registration); then evaluate the gain of knowledge after CS finalization.
- X. For larger projects, finance and recruit a full-time scientific manager (in addition to administrative management and overall project coordination).

Acknowledgements This work was supported by the European Union's Horizon 2020 research and innovation programme under Grant agreement no. 681002 (EU-ToxRisk).

Funding Open Access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Albrecht W, Kappenberg F, Brecklinghaus T et al (2019) Prediction of human drug-induced liver injury (DILI) in relation to oral doses and blood concentrations. *Arch Toxicol* 93:1609–1637. <https://doi.org/10.1007/s00204-019-02492-9>
- Bal-Price A, Hogberg HT, Crofton KM et al (2018) Recommendation on test readiness criteria for new approach methods in toxicology: exemplified for developmental neurotoxicity. *Altex* 35:306–352. <https://doi.org/10.14573/altex.1712081>
- Ballester M, Bolonio M, Santamaria R et al (2019) Direct conversion of human fibroblast to hepatocytes using a single inducible polycistronic vector. *Stem Cell Res Ther* 10:317. <https://doi.org/10.1186/s13287-019-1416-5>
- Beilmann M, Boonen H, Czich A et al (2019) Optimizing drug discovery by investigative toxicology: current and future trends. *Altex* 36:289–313. <https://doi.org/10.14573/altex.1808181>
- Benedetti G, Ramaiahgaris S, Herpers B et al (2013) A screen for apoptotic synergism between clinical relevant nephrotoxicant and the cytokine TNF-alpha. *Toxicol In Vitro* 27:2264–2272. <https://doi.org/10.1016/j.tiv.2013.09.004>
- Bischoff LJM, Kuijper IA, Schimming JP et al (2019) A systematic analysis of nrf2 pathway activation dynamics during repeated xenobiotic exposure. *Arch Toxicol* 93:435–451. <https://doi.org/10.1007/s00204-018-2353-2>
- Boon R, Kumar M, Tricot T et al (2020) Amino acid levels determine metabolism and CYP450 function of hepatocytes and hepatoma cell lines. *Nat Commun* 11(1):1393. <https://doi.org/10.1038/s41467-020-15058-6>
- Brull M, Spreng AS, Gutbier S et al (2020) Incorporation of stem cell-derived astrocytes into neuronal organoids to allow neuroglial interactions in toxicological studies. *Altex*. <https://doi.org/10.14573/altex.1911111>
- Busquet F, Hartung T, Pallocca G et al (2020) Harnessing the power of novel animal-free test methods for the development of COVID-19 drugs and vaccines. *Arch Toxicol*. <https://doi.org/10.1007/s00204-020-02787-2>
- Campos G, Schmidt-Heck W, De Smedt J et al (2020) Inflammation-associated suppression of metabolic gene networks in acute and chronic liver disease. *Arch Toxicol* 94:205–217. <https://doi.org/10.1007/s00204-019-02630-3>
- Coll M, Perea L, Boon R et al (2018) Generation of hepatic stellate cells from human pluripotent stem cells enables in vitro modeling of liver fibrosis. *Cell Stem Cell* 23(101–113):e107. <https://doi.org/10.1016/j.stem.2018.05.027>
- Copple IM, den Hollander W, Callegaro G et al (2019) Characterisation of the nrf2 transcriptional network and its response to chemical insult in primary human hepatocytes: implications for prediction of drug-induced liver injury. *Arch Toxicol* 93:385–399. <https://doi.org/10.1007/s00204-018-2354-1>
- Cote I, Andersen ME, Ankley GT et al (2016) The next generation of risk assessment multi-year study-highlights of findings, applications to risk assessment, and future directions. *Environ Health Perspect* 124:1671–1682. <https://doi.org/10.1289/EHP233>
- Daneshian M, Kamp H, Hengstler J et al (2016) Highlight report: launch of a large integrated European *in vitro* toxicology project: EU-ToxRisk. *Arch Toxicol* 90:1021–1024. <https://doi.org/10.1007/s00204-016-1698-7>
- Dearfield KL, Gollapudi BB, Bemis JC et al (2017) Next generation testing strategy for assessment of genomic damage: a conceptual framework and considerations. *Environ Mol Mutagen* 58:264–283. <https://doi.org/10.1002/em.22045>
- Delp J, Funke M, Rudolf F et al (2019) Development of a neurotoxicity assay that is tuned to detect mitochondrial toxicants. *Arch Toxicol* 93:1585–1608. <https://doi.org/10.1007/s00204-019-02473-y>
- Dempster AP (1967) Upper and lower probabilities induced by a multivalued mapping. *Ann Math Stat* 38:325–339
- Dent M, Teixeira Amaral R, Amores Da Silva P (2018) Principles underpinning the use of new methodologies in the risk assessment of cosmetic ingredients. *Comput Toxicol* 7:20–26. <https://doi.org/10.1016/j.comtox.2018.06.001>
- Dresner N, Madjar K, Holzer AK et al (2020) Development of a neural rosette formation assay (RoFa) to identify neurodevelopmental toxicants and to characterize their transcriptome disturbances. *Arch Toxicol* 94:151–171. <https://doi.org/10.1007/s00204-019-02612-5>
- ECHA (2017) Read-across assessment framework (RAAF). ECHA-17-R-01-EN. <https://doi.org/10.2823/619212>
- Escher SE, Kamp H, Bennekou SH et al (2019) Towards grouping concepts based on new approach methodologies in chemical hazard assessment: the read-across approach of the EU-ToxRisk project. *Arch Toxicol* 93:3643–3667. <https://doi.org/10.1007/s00204-019-02591-7>
- Fasbender F, Obholzer M, Metzler S et al (2020) Enhanced activation of human NK cells by drug-exposed hepatocytes. *Arch Toxicol* 94:439–448. <https://doi.org/10.1007/s00204-020-02668-8>
- Fisher C, Simeon S, Jamei M et al (2019) Vivd: virtual in vitro distribution model for the mechanistic prediction of intracellular concentrations of chemicals in in vitro toxicity assays. *Toxicol In Vitro* 58:42–50. <https://doi.org/10.1016/j.tiv.2018.12.017>
- Fredriksson L, Herpers B, Benedetti G et al (2011b) Diclofenac inhibits tumor necrosis factor- α -induced nuclear factor- κ B activation causing synergistic hepatocyte apoptosis. *Hepatology* 53(6):2027–2041. <https://doi.org/10.1002/hep.24314>
- Fredriksson L, Wink S, Herpers B et al (2014a) Drug-induced endoplasmic reticulum and oxidative stress responses independently sensitize toward TNF α -mediated hepatotoxicity. *Toxicol Sci* 140(1):144–159. <https://doi.org/10.1093/toxsci/kfu072>
- Gadaleta D, Lombardo A, Toma C et al (2018a) A new semi-automated workflow for chemical data retrieval and quality checking for modeling applications. *J Chem Inform* 10:60. <https://doi.org/10.1186/s13321-018-0315-6>
- Gadaleta D, Manganelli S, Roncaglioni A et al (2018b) Qsar modeling of toxcast assays relevant to the molecular initiating events of AOPs leading to hepatic steatosis. *J Chem Inf Model* 58:1501–1517. <https://doi.org/10.1021/acs.jcim.8b00297>
- Graepel R, ter Braak B, Escher SE et al (2019) Paradigm shift in safety assessment using new approach methods: the EU-ToxRisk strategy. *Curr Opin Toxicol* 15:33–39. <https://doi.org/10.1016/j.cotox.2019.03.005>
- Gu X, Albrecht W, Edlund K et al (2018) Relevance of the incubation period in cytotoxicity testing with primary human hepatocytes. *Arch Toxicol* 92:3505–3515. <https://doi.org/10.1007/s00204-018-2302-0>
- Gutbier S, Spreng AS, Delp J et al (2018) Prevention of neuronal apoptosis by astrocytes through thiol-mediated stress response modulation and accelerated recovery from proteotoxic stress. *Cell Death Differ* 25:2101–2117. <https://doi.org/10.1038/s41418-018-0229-x>
- Hartung T, Leist M (2008) Food for thought ... On the evolution of toxicology and the phasing out of animal testing. *Altex* 25:91–102. <https://doi.org/10.14573/altex.2008.2.91>
- Hemmerich J, Troger F, Fuzi B et al (2020) Using machine learning methods and structural alerts for prediction of mitochondrial toxicity. *Mol Inform* 39:e2000005. <https://doi.org/10.1002/minf.202000005>
- Hiemstra S, Ramaiahgari SC, Wink S et al (2019) High-throughput confocal imaging of differentiated 3D liver-like spheroid cellular stress response reporters for identification of drug-induced liver injury liability. *Arch Toxicol* 93:2895–2911. <https://doi.org/10.1007/s00204-019-02552-0>

- Khadka KK, Chen M, Liu Z et al (2020) Integrating adverse outcome pathways (AOPs) and high throughput in vitro assays for better risk evaluations, a study with drug-induced liver injury (DILI). *Altex* 37:187–196. <https://doi.org/10.14573/altex.1908151>
- Kobolak J, Teglas A, Bellak T et al (2020) Human-induced pluripotent stem cell-derived 3D-neurospheres are suitable for neurotoxicity screening. *Cells*. <https://doi.org/10.3390/cells9051122>
- Krebs A, Nyffeler J, Rahnenfuhrer J et al (2018) Normalization of data for viability and relative cell function curves. *Altex* 35:268–271. <https://doi.org/10.14573/1803231>
- Krebs A, Waldmann T, Wilks MF et al (2019) Template for the description of cell-based toxicological test methods to allow evaluation and regulatory use of the data. *Altex* 36:682–699. <https://doi.org/10.14573/altex.1909271>
- Krebs A, Nyffeler J, Karreman C et al (2020a) Determination of benchmark concentrations and their statistical uncertainty for cytotoxicity test data and functional in vitro assays. *Altex* 37:155–163. <https://doi.org/10.14573/altex.1912021>
- Krebs A, van Vugt-Lussenburg BMA, Waldmann T et al (2020b) The EU-ToxRisk method documentation, data processing and chemical testing pipeline for the regulatory use of new approach methods. *Arch Toxicol*. <https://doi.org/10.1007/s00204-020-02802-6>
- Krewski D, Westphal M, Andersen ME et al (2014) A framework for the next generation of risk science. *Environ Health Perspect* 122:796–805. <https://doi.org/10.1289/ehp.1307260>
- Krewski D, Andersen ME, Tyshenko MG et al (2020) Toxicity testing in the 21st century: progress in the past decade and future perspectives. *Arch Toxicol* 94:1–58. <https://doi.org/10.1007/s00204-019-02613-4>
- Leist M, Gantner F, Naumann H et al (1997) Tumor necrosis factor-induced apoptosis during the poisoning of mice with hepatotoxins. *Gastroenterology* 112:923–934. <https://doi.org/10.1053/gast.1997.v112.pm9041255>
- Leist M, Hartung T, Nicotera P (2008) The dawning of a new age of toxicology. *Altex* 25:103–114
- Leist M, Hasiwa N, Rovida C et al (2014) Consensus report on the future of animal-free systemic toxicity testing. *Altex* 31:341–356. <https://doi.org/10.14573/altex.1406091>
- Limonciel A, Ates G, Carta G et al (2018a) Comparison of base-line and chemical-induced transcriptomic responses in HepaRG and RPTEC/TERT1 cells using tempo-seq. *Arch Toxicol* 92:2517–2531. <https://doi.org/10.1007/s00204-018-2256-2>
- Limonciel A, van Breda SG, Jiang X et al (2018b) Persistence of epigenomic effects after recovery from repeated treatment with two nephrocarcinogens. *Front Genet* 9:558. <https://doi.org/10.3389/fgene.2018.00558>
- Luechtefeld T, Marsh D, Rowlands C et al (2018) Machine learning of toxicological big data enables read-across structure activity relationships (RASAR) outperforming animal test reproducibility. *Toxicol Sci* 165:198–212. <https://doi.org/10.1093/toxsci/kfy152>
- Lupu D, Andersson P, Bornehag CG et al (2020) The ENDpoiNTs Project: novel testing strategies for endocrine disruptors linked to developmental neurotoxicity. *Int J Mol Sci* 21(11):E3978. <https://doi.org/10.3390/ijms21113978>
- Maertens A, Tran V, Kleensang A et al (2018) Weighted gene correlation network analysis (WGCNA) reveals novel transcription factors associated with bisphenol A dose-response. *Front Genet* 9:508. <https://doi.org/10.3389/fgene.2018.00508>
- Marx U, Akabane T, Andersson TB et al (2020) Biology-inspired microphysiological systems to advance patient benefit and animal welfare in drug development. *Altex*. <https://doi.org/10.14573/altex.2001241>
- Monshi MM, Faulkner L, Gibson A et al (2013) Human leukocyte antigen (HLA)-b*57:01-restricted activation of drug-specific T cells provides the immunological basis for flucloxacillin-induced liver injury. *Hepatology* 57:727–739. <https://doi.org/10.1002/hep.26077>
- Ntzani EE, Chondrogiorgi M, Ntritsos G et al (2013) Literature review on epidemiological studies linking exposure to pesticides and health effects. *EFSA Support Publ* 10(10):1–159. <https://doi.org/10.2903/sp.efsa.2013.EN-497>
- Nyffeler J, Chovancova P, Dolde X et al (2018) A structure-activity relationship linking non-planar PCBs to functional deficits of neural crest cells: new roles for connexins. *Arch Toxicol* 92:1225–1247. <https://doi.org/10.1007/s00204-017-2125-4>
- OECD (2014) Guidance document for describing non-guideline in vitro test methods
- OECD (2015) Report on considerations from case studies on integrated approaches for testing and assessment (IATA)
- Pamies D, Leist M, Coecke S et al (2020) Good cell and tissue culture practice 2.0 (gccp 2.0)—draft for stakeholder discussion and call for action. *Altex* 37:490–492
- Ramirez T, Strigun A, Verlohner A et al (2018) Prediction of liver toxicity and mode of action using metabolomics *in vitro* in HepG2 cells. *Arch Toxicol* 92:893–906. <https://doi.org/10.1007/s00204-017-2079-6>
- Ramme AP, Koenig L, Hasenberg T et al (2019) Autologous induced pluripotent stem cell-derived four-organ-chip. *Future Sci OA* 5:FSO413. <https://doi.org/10.2144/fsoa-2019-0065>
- Rathman JF, Yang C, Zhou H (2018) Dempster–Shafer theory for combining in silico evidence and estimating uncertainty in chemical risk assessment. *Comput Toxicol* 6:16–31
- Rauch C, Feifel E, Kern G et al (2018) Differentiation of human iPSCs into functional podocytes. *PLoS ONE* 13:e0203869. <https://doi.org/10.1371/journal.pone.0203869>
- Rovida C, Barton-Maclaren T, Benfenati E et al (2020a) Internationalization of read-across as a validated new approach method (NAM) for regulatory toxicology. *Altex*. <https://doi.org/10.14573/altex.1912181>
- Rovida C, Escher SE, Herzler M et al (2020b) Workshop report: NAM-supported read-across: from case studies to regulatory guidance in safety (**in preparation**)
- Sarkans U, Gostev M, Athar A et al (2018) The biostudies database—one stop shop for all data supporting a life sciences study. *Nucleic Acids Res* 46:D1266–D1270. <https://doi.org/10.1093/nar/gkx965>
- Sauer JM, Hartung T, Leist M et al (2015) Systems toxicology: the future of risk assessment. *Int J Toxicol* 34:346–348. <https://doi.org/10.1177/1091581815576551>
- Schimming JP, Ter Braak B, Niemeijer M et al (2019) System microscopy of stress response pathways in cholestasis research. *Methods Mol Biol* 1981:187–202. https://doi.org/10.1007/978-1-4939-9420-5_13
- Shafer G (1976) A mathematical theory of evidence. Princeton University Press, Princeton (**ISBN 978-0691100425**)
- Schultz TW, Richarz A-N, Cronin MTD (2018) Assessing uncertainty in read-across: questions to evaluate toxicity predictions based on knowledge gained from case studies. *Comput Toxicol* 9:1–11
- Simeon S, Brotzmann K, Fisher C et al (2020) Development of a generic zebrafish embryo PBPK model and application to the developmental toxicity assessment of valproic acid analogs. *Reprod Toxicol* 93:219–229. <https://doi.org/10.1016/j.reprotox.2020.02.010>
- Terron A, Bal-Price A, Paini A et al (2018) An adverse outcome pathway for parkinsonian motor deficits associated with mitochondrial complex I inhibition. *Arch Toxicol* 92:41–82. <https://doi.org/10.1007/s00204-017-2133-4>
- Thomas RS, Paules RS, Simeonov A et al (2018) The US federal tox21 program: a strategic and operational plan for continued leadership. *Altex* 35:163–168. <https://doi.org/10.14573/altex.1803011>
- Thomas RS, Bahadori T, Buckley TJ et al (2019) The next generation blueprint of computational toxicology at the U.S.

- Environmental protection agency. *Toxicol Sci* 169:317–332. <https://doi.org/10.1093/toxsci/kfz058>
- Toma C, Gadaleta D, Roncaglioni A et al (2018) Qsar development for plasma protein binding: Influence of the ionization state. *Pharm Res* 36:28. <https://doi.org/10.1007/s11095-018-2561-8>
- Toropov AA, Toropova AP (2017) The index of ideality of correlation: a criterion of predictive potential of QSPR/QSAR models? *Mutat Res* 819:31–37. <https://doi.org/10.1016/j.mrgentox.2017.05.008>
- Toropova AP, Toropov AA, Marzo M et al (2018) The application of new hard-descriptor available from the coral software to building up NOAEL models. *Food Chem Toxicol* 112:544–550. <https://doi.org/10.1016/j.fct.2017.03.060>
- Troger F, Delp J, Funke M et al (2020) Identification of mitochondrial toxicants by combined in silico and in vitro studies—a structure-based view on the adverse outcome pathway. *Comput Toxicol* 14:100–123. <https://doi.org/10.1016/j.comtox.2020.100123>
- Wink S, Hiemstra SW, 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. <https://doi.org/10.1007/s00204-018-2178-z>
- Yang H, Niemeijer M, van de Water B et al (2020) Atf6 is a critical determinant of chop dynamics during the unfolded protein response. *iScience* 23:100860. <https://doi.org/10.1016/j.isci.2020.100860>
- Zgheib E, Gao W, Limonciel A et al (2019) Application of three approaches for quantitative AOP development to renal toxicity. *Comput Toxicol* 11:1–13. <https://doi.org/10.1016/j.comtox.2019.02.001>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.