

“The merit of all things lies in their difficulty.”

Alexandre Dumas (1802–1870) in *The Three Musketeers*

“Knowing is not enough; we must apply.

Willing is not enough; we must do.”

Johann Wolfgang von Goethe (1749–1832)

Food for Thought ...

A Call for a Human Exposome Project

Thomas Hartung^{1,2}

¹Center for Alternatives to Animal Testing (CAAT), Johns Hopkins University, Bloomberg School of Public Health, Baltimore, MD, USA;

²CAAT Europe, University of Konstanz, Konstanz, Germany

Abstract

Four decades of the Human Genome Project and its consequences have shown how the entrepreneurial state, through significant investment into science, can drive scientific progress and advance biomedicine. A certain fraction of diseases can now be explained as caused by genetics, and a more significant fraction as impacted by genetics. Besides another fraction caused by pathogens, the third and probably largest impactor is exposure, i.e., the many physicochemical and lifestyle factors. This article makes the case that it is time to start a Human Exposome Project, which systematically explores and catalogs the exposure side of human health and disease.

The envisioned Human Exposome Project needs to be more than a scaled exposomics approach, aiming to assess the totality of relevant exposures through -omics of human body fluids and forming exposure hypotheses. Exposomics is increasingly complemented by exposure science and biomonitoring to measure exposure, mechanistic understanding, human-relevant microphysiological systems, big data, and artificial intelligence (AI) to mine these data and integrate pieces of evidence. The potential impact of AI on a possible Human Exposome Project is so substantial that we should speak of exposome intelligence (EI) because this allows us to expand our limited current knowledge to the big unknown unknowns of threats to human health.

1 Introduction

The theoretical case for a Human Exposome Project is easy to make. Wouldn't it be wonderful to understand the contribution of exposure to disease? The question is, however, how and is it feasible? But first, why call it a Human Exposome Project and not a Human Toxome Project¹ (Hartung and McBride, 2011)? This has to do with the undeserved mediocre reputation of toxicology, which has already led us to preferentially talk of safety science, risk science, regulatory science, next generation risk assessment, etc. Undeserved, because toxicology is actually big science:

– Only a small percentage of disease is genetic, the rest is bad luck and exposure.

- We prevent disease (unfortunately nobody knows how much).
- Almost all products should undergo risk assessment.
- Toxicology pioneered quality (GLP) and validation (incl. relevance).
- Toxicology was the first preclinical science to adapt evidence-based approaches.

However, we are not doing a good job selling this. We are the nay-sayers in product development who delay or stop all these fabulous product launches and this typically at a very late stage close to marketing. We are the harbingers of bad news about our food and our environment, the risks that spoil enjoying them. And the discipline of toxicology seems so close to pharmacology that academic chairs are increasingly being

Received January 6, 2023;
© The Authors, 2023.

ALTEX 40(1), 4–33. doi:10.14573/altex.2301061

Correspondence: Thomas Hartung, MD, PhD
Center for Alternatives to Animal Testing (CAAT)
Johns Hopkins University
615 N Wolfe St., Baltimore, MD, 21205, USA
(THartun1@jhu.edu)

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is appropriately cited.

¹ In fact, there was a small Human Toxome Project funded as an NIH Transformative Research Grant (Bouhifd et al., 2015a).

combined and are becoming fewer, not more. To some extent, the bad reputation comes with the job, but it also has to do with how we do our job:

Toxicology as we do it

- ...takes too long,
- ...is too expensive,
- ...does not have the necessary throughput,
- ...is too animal-driven,
- ...does not really consider exposure,
- ...is not reproducible,
- ...focuses too much on regulation,
- ...has too many false-positives,
- ...struggles to integrate new tools.

Terence McKenna (1946-2000) is quoted, “*We are living in a state of constant scientific revolution. There is not a single area that you can name that is now seen as it was seen a hundred years ago. Nothing is left of the world view of one hundred years ago.*” He does not know toxicology, where repeated dose testing in rats was introduced a hundred years ago...

Let’s not lament about these today – we have done so often enough in this Food for thought ... series over 16 years. Let’s rather dream big about the vision of a Human Exposome Project. The idea to revive this came with the opportunity to develop a vision for the Department of Defense. More precisely, the Basic Research Office of the Office of the Under Secretary of Defense for Research and Engineering, OUSD(R&E), approached the author in August 2020 to organize and chair a Future Directions workshop, *Advancing the Next Scientific Revolution in Toxicology*. A vanguard of scientific and technical experts and agency observers developed a report of the meeting held April 28-29, 2022, laying out how recent developments can be embraced and set the direction of “*Toxicology for the 21st Century 2.0*” in the next decades. The report will be published shortly by the Department of Defense and a publication coauthored by the participants is in advanced stages. Here, a sneak-preview is given of the overall messages with a personal spin by the author. So, while it is an abrogated version, it is not a type of Reader’s Digest but a personal account of some roots and messages. The ultimate message of the workshop, however, was a call for a Human Exposome Project. This article expands on what we have and what it will take to get there.

On the way of supporting the ongoing revolution in toxicology (Hartung and Tsatsakis, 2021), many workshops were held and strategy papers were written. One of these workshops, *Determining Adverse Responses Using In Vitro Assays*, held in 2015, with its discussions helped the author a lot in focusing some of the principal ideas. Unfortunately, the respective report was never finalized. The surviving rudimentary report, redacted and revised by the author, is included as Section 4 of this article. It was not circulated again to the original discussants as this does not make much sense almost eight years later. Since the workshop was largely financed through the Center for Alternatives to Animal Testing (CAAT) and its t^4 workshop program, I feel some

ownership and responsibility to make these discussions public as originally intended. Still, I owe the participants for their stimulating ideas but take any blame for misrepresentations.

2 The challenge of toxicology for the 21st century

We cannot have a discussion on the future of toxicology without anchoring it at the watershed moment of US toxicology, i.e., the publication of the NRC report “*Toxicity Testing in the 21st Century: A Vision and a Strategy*” (NRC, 2007), aka Tox-21c. While not serving on the panel, the author presented to the Tox-21c panel and has had the privilege to work with many committee members on its implementation since then. It is tempting to cite Sir Isaac Newton, “*If I have seen further than others, it is by standing upon the shoulders of giants.*” However, a nice spin on this came from Benjamin F. Jones, Northwestern University and National Bureau of Economic Research in 2011, “*If one wants to stand on the shoulders of giants (taking Newton’s famous aphorism), then one must first climb up the giants’ backs. As knowledge accumulates, the harder this climb can become.*” The remarkable progress in implementing Tox-21 has been recently summarized (Krewski et al., 2020). The author suggested a roadmap for the implementation of Tox-21c in 2009 (Hartung, 2009a) and has pursued this route with CAAT since then. This article concluded that we need to map the pathways leading to toxicity, which one year later was introduced as adverse outcome pathways (AOP) (Ankley et al., 2010) and as pathways of toxicity in the context of the Human Toxome Project (Hartung and McBride, 2011; Kleensang et al., 2014). In the following, a quick resume is given on where we stand on the ten challenges identified in 2009 (Hartung, 2009a).

2.1 Challenge 1: Testing strategies instead of individual tests

The central call of Tox-21c to move to a mechanistic toxicology of perturbed pathways of toxicity implies the use of multiple tests and types of evidence. These integrated testing strategies (ITS), termed integrated approaches to testing and assessment (IATA) by the Organisation for Economic Cooperation and Development (OECD) (Tollefsen et al., 2014), and defined approaches, have been discussed in this series (Hartung et al., 2013a), a workshop (Rovida et al., 2015a), and a more recent review (Caloni et al., 2022). Some progress, especially the defined approaches on skin sensitization², is noted (Kleinstreuer et al., 2018; Kolle et al., 2020). However, overall, the advance towards a systematic composition and validation of testing strategies is meagre. The problem should be seen in the overall challenge of evidence integration (Linkov et al., 2015), which is the same whether carrying out a risk assessment or a systematic review from several combined evidence streams or constructing a prospective test strategy. Notably, this was the topic of two workshops held by our Evidence-based Toxicology Collaboration (EBTC) with the

² <https://www.oecd.org/chemicalsafety/guideline-no-497-defined-approaches-on-skin-sensitisation-b92879a4-en.htm>



European Food Safety Authority (EFSA) on *Evidence integration in risk assessment: the science of combining apples and oranges* (EFSA and EBTC, 2018) and the Risk Sciences Institute, Ottawa, on *Development of a Framework for Evidence Synthesis* (Krewski et al., in preparation).

Taken together, there is a notable advance toward integrated testing strategies, but we still fall short of making this the standard approach and composing and validating them systematically.

2.2 Challenge 2: Statistics and multiple testing

Closely linked to challenge 1, a statistical problem arises as stated in 2009: “When testing for multiple pathways, we will need to correct our statistics for multiple testing. We have to lower significance levels accordingly or we will run increasingly into false-positive findings. The proponents of the new approach assume more than a hundred, and less than a thousand such pathways. A lot of multiple testing... Assuming only 100 pathways, significance levels of $p = 0.05$ would have to be lowered to 0.006 using the most common Bonferroni correction. This – likely with sophisticated methods of high inherent variance – will result in an astronomic number of replicates necessary: For the example of $p = 0.05$ and stable noise/signal ratios, a 71-fold increase in sample size (e.g., number of animals or replicate cellular tests) is necessary to reach the same level of confidence.” This was written under the impression of the discussions on the genotoxicity test battery approach with its enormous accumulation of false-positive findings (Blakey et al., 2008; Kirkland et al., 2007; Basketter et al., 2012).

The new kid on the block is artificial intelligence (AI) employed as machine learning to make probabilistic predictions, which represents a different way of evidence integration. Both supervised and reinforcement learning train a neural network to deliver probabilities of a result. The different inputs are weighed in the iterations of deep learning, thereby reducing the dimensionality of input layers. Ultimately, this represents a probabilistic approach as discussed in this series (Maertens et al., 2022). It is the author’s strong belief that these are key components of a future toxicology based on multiple evidence streams.

Taken together, the potential of combining different pieces of evidence through AI to derive a probability instead of traditional significance testing with multiple testing correction needs to be explored.

2.3 Challenge 3: Threshold setting

A lot of the 2009 reasoning still holds today: “Where does a relevant effect start? Certainly not where we can measure a significant change. What is measurable depends only on our detection limits, and in the case of multiendpoint methods a lot on signal/noise relation and the inevitable number of false-positive results. If, for example, a toxicogenomics approach is taken, several thousand genes might be measured and, especially when low thresholds of fold induction are used, false-positive events will occur. Even if real-positive, the questions arising are then,

whether this is significant with the given number of replicates, or even more important, whether this is relevant (notably completely different questions). Although the former can be tested with replicate testing and statistics (see, however, problem of multiple testing), the relevance is more difficult to establish: The more remote we are in (sub-) cellular pathways, the more difficult to extrapolate to the overall organism. The NRC vision document is not really clear here, whether we talk of cells and their signal transduction pathways or the even more complicated physiological pathways in dynamic systems with compensatory mechanisms.

What does it mean if a pathway is triggered but if accompanied by some compensatory ones as well? We definitively have to overcome the mentality of ‘we see an effect, this is an effect level.’ Any method, which assesses only a certain level of the organism (e.g., the transcriptome when using genomics), will be questioned whether these changes are translated to the higher integration levels (proteome, metabolism, physiology). This argues for systems biology approaches where such considerations are taken into account, but complexity of modeling increases dramatically, with impacts on standardization, costs, feasible number of replicates etc. The greater the distance from the primary measurement to the overall result in a model, the more difficult threshold setting will become because of error propagation.

Setting of thresholds or other means of deriving a test result (data analysis procedure) is a most critical part of test development. It determines the sensitivity and specificity of the new test, that is, the proportion of false-positive and false-negative results.”

A major discussion in this context is the one on thresholds of toxicological concern (TTC), discussed earlier in this series (Hartung, 2017a). In 2008, Hartung and Leist first suggested to extend this concept to an internal TTC (iTTC), basically a refinement of the TTC concept based on plasma concentrations for the risk assessment of substances with a low absorption (by the oral or dermal route), as the internal exposure is in these cases more relevant than the external exposure. The iTTC concept has in the meantime been expanded (Partosch et al., 2015; Ellison et al., 2020). Notably, the concept is also key to the ongoing ONTOX project (Vinken et al., 2021) and its implementation of probabilistic risk assessment (Maertens et al., 2022; t⁴ workshop report, in preparation).

The central question of how to define where adversity on an organism level starts from *in vitro* systems was addressed in the workshop summarized in Section 4.

2.4 Challenge 4: What to validate against?

Validation is a demanding process, so demanding that shortcuts and redefinitions are often attempted. Ultimately, there is no alternative to the independent evaluation of the reliability and relevance of a test method and its scientific basis, which is the very definition of validation (Hartung, 2007a; Leist et al., 2012). Some improvements were already introduced in the modular approach (Hartung et al., 2004), but many of these were hardly ever implemented, such as lean design (Hoffmann and Har-

tung, 2006a) or retrospective validation (Corvi et al., 2008). We suggested earlier (Hoffmann et al., 2008) the concept of composite reference points, i.e., a consensus process of identifying the reference result attributed to a reference test substance. Currently, there are new attempts to refine the process of validation (Mondou et al., 2021; van der Zalm et al., 2022), and ECVAM has started to explore the validation of complex *in vitro* models through a survey³.

In the context of Tox-21c, the challenge to validation, however, goes beyond process: It needs to be human- not animal-based and mechanistic in nature. But what should we validate against? There is increasing understanding that animal studies do not serve well as gold standards. An ongoing US National Academy of Science committee works on *Variability and Relevance of Current Laboratory Mammalian Toxicity Tests and Expectations for New Approach Methods (NAMs) for use in Human Health Risk Assessment*⁴, which could be an important milestone, but the recent exclusion of some key committee members, leading to a predominance of animal experimentalists, is worrying.

We have developed suggestions for the validation of Tox-21c tools out of evidence-based toxicology (Hartung, 2010a) and for a mechanistic validation (Hartung et al., 2013b). The latter follows the concept that mechanistic assays should be evaluated primarily based on whether they reflect human pathomechanisms. This was implemented in Modafferi et al. (2021) as a case study where a model for developmental neurotoxicity (DNT) was shown to reflect metabolomic changes in human patients.

Taken together, validation, though often perceived as cumbersome and unduly time-consuming, is the clear path towards building trust in new approaches in the safety sciences and as such is indispensable. However, it is a permanent learning process, and the path to Tox-21c, or in extension to a Human Exposure Project, is different to that of replacing individual patches of the patchwork of toxicology. The big learning, which has become broadly accepted in recent years, is that traditional animal studies cannot serve as a general gold standard. Elements that could adapt the process, such as mechanistic validation, are on the table but are not yet embraced to a major extent.

2.5 Challenge 5:

How to open up regulators for change?

Isaac Asimov is quoted for “*The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.*” This dilemma is especially obvious in the regulatory arena, where we often could change but do not manage to

do so. How to convince regulators to change? The best way to convince regulators is strong evidence! Out of evidence-based medicine, a systematic approach to generate, accumulate, assess, and present an evidence-base has been developed. The Evidence-based Toxicology Collaboration (EBTC)⁵ (Hoffmann and Hartung, 2006b; Stephens et al., 2013) was established in 2011. The board of trustees includes members of leading regulatory agencies from the US and Europe. The collaboration has produced guidance on systematic reviews (Stephens et al., 2016; Hoffmann et al., 2017, 2022a), including case studies (Stephens et al., 2019; Hoffmann et al., 2021), quality scoring (Samuel et al., 2016), evidence integration, mechanistic frameworks (de Vries et al., 2021; Hoffmann et al., 2022b), etc. About 1,000 researchers have interacted with EBTC, and the number of systematic reviews in environmental health has increased greatly in the last decade.

In addition, the regulatory agencies have opened themselves up for change: Much of Tox-21c is driven by the alliance of EPA, FDA, and NIH⁶. Automated, “high-throughput” screening serves an important role, especially the overlapping EPA ToxCast⁷ activity. The resulting body of work goes beyond what can be discussed here and has created a toxicological data resource second to none. The Integrated Chemical Environment (ICE)⁸ (Abedini et al., 2021) by the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)⁹ further expands data and model access. The FDA is in a process of changing its attitude toward NAM¹⁰. With the lead role of pharmaceutical industry for biotech companies to market NAMs, this is of critical importance as the US with ~4% of the world population represented a stunning 49.1% of the world pharma market in 2021¹¹ and, even more extreme, 63.7% of the new medicines (introduced 2015-2020). This is why the position of FDA toward NAM is so critical for the NAM market (Meigs et al., 2018).

On the European side, the driving work by EFSA (EFSA et al., 2022; Escher et al. 2022) and the starting European Partnership for the Assessment of Risks from Chemicals (PARC)¹² must be mentioned. PARC is a 7-year partnership under Horizon Europe, the EU’s 2021-2027 framework program for the funding of research and innovation, with 200 partners in 28 countries and at EU level, national agencies and research organizations working in the areas of the environment or public health, the European Chemicals Agency (ECHA), EFSA, and the European Environment Agency (EEA). A budget of €400 million, 50% funded by the European Union and 50% by EU

³ <https://op.europa.eu/en/publication-detail/-/publication/3824b024-5621-11eb-b59f-01aa75ed71a1/language-en>

⁴ <https://www.nationalacademies.org/our-work/variability-and-relevance-of-current-laboratory-mammalian-toxicity-tests-and-expectations-for-new-approach-methods-nams-for-use-in-human-health-risk-assessment>

⁵ <https://www.ebtox.org>

⁶ <https://ntp.niehs.nih.gov/whatwestudy/tox21/index.html>

⁷ <https://www.epa.gov/chemical-research/toxicity-forecasting>

⁸ <https://ntp.niehs.nih.gov/whatwestudy/niceatm/comptox/ct-ice/ice.html>

⁹ <https://ntp.niehs.nih.gov/whatwestudy/niceatm/index.html>

¹⁰ <https://www.fda.gov/media/144891/download>

¹¹ <https://www.efpia.eu/publications/data-center/the-pharma-industry-in-figures-economy/world-pharmaceutical-market/>

¹² <https://www.anses.fr/en/content/european-partnership-assessment-risks-chemicals-parc>



member states, is foreseen. Further, the German Federal Institute for Risk Assessment has recently voiced some remarkably progressive positions¹³ (Tralau et al., 2015; Tralau and Luch, 2015; Herzler et al., 2021).

Altogether, this shows how much the attitude of some regulatory agencies has changed toward new approaches since the beginning of Tox-21c. However, we should be clear that this attitude has not always reached to parts of the agencies handling registrations or to some other key agencies and countries outside the transatlantic space. Education of regulators about the challenges and opportunities of NAMs appears to be the best way forward, especially when it involves regulators who have made this mental transition.

2.6 Challenge 6: The global dimension

Science is global, companies market their products globally, and chemicals essentially threaten all humans similarly – but regulation is regional. The globalization of industries sometimes creates opportunities but often hurdles (Bottini et al., 2007). An impressive example was the initiative to abandon the one-year-dog study for pesticides: Six major studies showed the limited usefulness of this test (Spielmann, 2019), resulting in changes to regulatory requirements that struck the study in the EU in 2009 and made it required in the US, Canada, and Australia only when dogs are shown to be more sensitive than rodents in 90-day studies. Still, this meant no change for agrochemical companies until most recently Brazil, Korea, and Japan took similar positions, following lobbying activities of several partners.

This stresses the importance of international harmonization, e.g., with the OECD (Browne et al., 2019). In the pharmaceutical field, the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH)¹⁴ between the US, Europe, and Japan, originally founded in 1990¹⁵ and reformed as a non-profit legal entity under Swiss Law in 2015, with one purpose being the reduction of unnecessary animal testing without compromising safety and effectiveness¹⁶, has produced 100+ guidelines on technical requirements. Quite remarkably, guidance agreed in August 2022¹⁷ and released by FDA in November 2022 on *Testing for carcinogenicity of pharmaceuticals S1B(R1)* opens up for alternative methods for the cancer bioassay “*Application of this integrative approach reduces the use of animals in accordance with the 3R (reduce/refine/*

replace) principles and shifts resources to focus on generating more scientific mechanism-based carcinogenicity assessments, while continuing to promote safe and ethical development of new pharmaceuticals.” Even in the pesticide sector, which relies heavily on animal testing of its products, progress is being made internationally toward the 3Rs (Stucki et al., 2022).

An important part of international harmonization is the creation of discussion fora such as:

- The series of *World Congresses on Alternatives and Animal Use in the Life Sciences* started by CAAT in 1993; the 12th conference is planned for Niagara Falls, Canada, Aug 27-31, 2023¹⁸.
- International Cooperation on Cosmetics Regulation¹⁹ created as a result of our discussions between ECVAM and DG Enterprise in 2005.
- The series of *Pan-American Conferences for Alternative Methods* started by CAAT in 2016^{20,21} – the series was paused due to the pandemic and will resume as part of WC12 in Canada in 2023; the next conference in South America is foreseen for 2025.
- The series of *Microphysiological Systems World Summits*, started in 2022 by CAAT in New Orleans²²; the 2nd summit will take place in Berlin, Germany, June 26-30, 2023²³.
- The International Collaboration on Cosmetics Safety (ICCS)²⁴ is a new global multistakeholder collaboration involving more than 40 cosmetics companies, trade associations, academia, NGOs, and regulatory bodies to achieve widespread use of next generation approaches to replace animal testing.
- Ongoing efforts to create an Asian federation of alternatives to animal experiments societies.

Taken together, international harmonization is progressing, but there is still room for improvement. Especially, the global standard for validation, as enshrined in OECD GD 34, is eroding in a variety of activities to revise validation standards and processes (see Section 2.4). Varying regional acceptance is sometimes the biggest hurdle for the introduction of new approaches, as the more progressive regulatory processes typically allow companies to stick to traditional approaches as long as important markets require them. It would be a step forward if acceptance of a new method meant that it became the required method independent of requirements in other regions. In consequence, a market for the respective NAM would be formed, unleashing the marketing powers of the (biotech) companies. Where regions

¹³ https://www.bfr.bund.de/en/press_information/2021/42/innovative_non_animal_methods_for_risk_assessment_of_chemicals_in_the_centre_of_interest-287549.html

¹⁴ <https://www.ich.org>

¹⁵ Expanded to regulators from Switzerland, Canada, Brazil, Singapore, China, South Korea, Turkey and Taiwan and observers such as WHO

¹⁶ For summary of 3Rs-relevant activity see presentation by Paul C. Brown, FDA from 2020: https://ntp.niehs.nih.gov/ntp/about_ntp/sacatm/2020/september/present/2a-2-brown-508.pdf

¹⁷ https://database.ich.org/sites/default/files/S1B-R1_FinalGuideline_2022_0719.pdf

¹⁸ <https://www.wc12canada.org>

¹⁹ <https://www.iccr-cosmetics.org>

²⁰ <https://caat.jhsph.edu/programs/workshops/PanAmerican/>

²¹ <https://caat.jhsph.edu/programs/workshops/PanAmerican2018/>

²² <https://mpsworldsummit.com/mps-world-summit-2022-3/>

²³ <https://mpsworldsummit.com/mps-world-summit-2023/>

²⁴ https://www.mayerbrown.com/-/media/files/perspectives-events/events/2022/11/fragfs2022_program.pdf

Tab. 1: Good practices and reporting standards in the life sciences, especially toxicology

<i>In silico</i>	<i>In vitro</i>	<i>In vivo</i>
(Q)SAR – no Good Read-Across Practice (GRAP)	Good Cell Culture Practice (GCCP) Good In Vitro Method Practice (GIVIMP)	FELASA/AALAS guidance Good Laboratory Practice (GLP)
(Q)SAR Model Reporting Format (QMRF) Read-Across Assessment Framework (RAAF)	Good In Vitro Reporting Standards (GIVReSt), in preparation	Animal Research: Reporting of In Vivo Experiments (ARRIVE)

are not yet harmonized, parallel data of old and new methods would be obtained and could be used for further evaluations of the NAM, and an interest of the regulated company would be created to make other regions change as well.

2.7 Challenge 7:

Quality assurance for the new approach

Henry Ford (1863-1947) is often quoted for “*Quality means doing it right when no one is looking.*” In science, especially in toxicology, we often have to assess quality when integrating often contradictory pieces of evidence. This quality scoring is a crucial part of evidence-based approaches (see Section 2.5), but it is best when work has already been carried out according to quality standards and is then reported appropriately. As stated already in the introduction, toxicology is far ahead of other pre-clinical sciences with the development of good laboratory practice (GLP). The OECD Principles of GLP²⁵ ensure the generation of high-quality and reliable test data related to the safety of industrial chemical substances and preparations. The principles were created in the context of harmonizing testing procedures for the mutual acceptance of data (MAD). Noteworthy, at the time, these were primarily animal test guidelines. For these, laboratory animal guidance was established by organizations such as the Institute for Laboratory Animal Research (ILAR)^{26,27} of the US National Academies of Sciences, Engineering, and Medicine (NAS), the American Association for Laboratory Animal Science (AALAS)²⁸, and the Federation of European Laboratory Animal Science Associations (FELASA) (Guillen, 2012) as well as the ARRIVE (Animal Research: Reporting of In Vivo Experiments) reporting guidelines²⁹ (Tab. 1).

For *in vitro* and *in silico* work, the same level of quality assurance guidance has not yet been established. With the rise of NAMs, Good In Vitro Method Practices (GIVIMP)³⁰ were developed to complement GLP. This started with an ECVAM workshop on GLP in 1998 (Cooper-Hannan et al., 1999). Parallel initiatives in 1996 in Germany organized by the author and in 1999 in Bologna at the Third World Congress on Alternatives and Animal Use in the Life Sciences led to a declaration toward

good cell culture practice (GCCP) (Gstraunthaler and Hartung, 1999). A GCCP task force was established, which produced two reports (Hartung et al., 2002, Coecke et al., 2005). The Coecke et al. (2005) document is paraphrased strongly in GIVIMP and annexed in full. With the increasing use of stem cells and commercial human primary cells and the bioengineering of MPS, a need for an update of GCCP formed. In two workshops (Pamies et al., 2017, 2018), the necessary adaptation was sketched out. Notably, the first of these workshop reports (Pamies et al., 2017) is also annexed in GIVIMP in full. A taskforce was formed, which published draft guidance (Pamies et al., 2020) for stakeholder discussion and then full guidance (Pamies et al., 2022). GCCP differs from GLP as it is geared toward non-regulated communities, especially academia, as a minimum standard for quality assurance, as GLP cannot normally be implemented in academia on the grounds of costs and its lack of flexibility. For example, GLP requires that fully trained personnel carry out any experiment, while “learning on the job” is academic practice. GCCP wants to provide guidance for journals and funding bodies. A large part of GCCP is about documentation needs for experimental work to make it reproducible. Excellent and well-documented work can still be reported badly, making it useless for other researchers, irreproducible, and unsuitable for evidence integration. With this in mind, the development of Good In Vitro Reporting Standards (GIVReSt) was started by CAAT in 2016 (Hartung et al., 2019). Apparently, NC3Rs in the UK has recently started a similar activity, but there is nothing yet in the public domain.

In the *in silico* field, the situation is even more dire. For (Q)SAR (Muratov et al., 2020), single-authored recommendations are available (Tropsha, 2010) but no broadly endorsed guidance. However, guidance is available both for validation³¹ and reporting (Piir et al., 2018; Nantassenamat, 2020). Noteworthy, Piir et al. (2018), analyzing 1,533 QSAR articles, found “*strikingly poor documentation of QSARs*” where “*42.5% of the reviewed articles were found to be potentially reproducible*”. The (Q)SAR Model Reporting Format (QMRF)³² is a template developed at the European Union Reference Laboratory

²⁵ <https://www.oecd.org/chemicalsafety/testing/good-laboratory-practiceglp.htm>

²⁶ <https://www.nationalacademies.org/ilar/institute-for-laboratory-animal-research>

²⁷ <https://nap.nationalacademies.org/catalog/12910/guide-for-the-care-and-use-of-laboratory-animals-eighth>

²⁸ <https://www.aalas.org>

²⁹ <https://arriveguidelines.org>

³⁰ <https://www.oecd.org/env/guidance-document-on-good-in-vitro-method-practices-givimp-9789264304796-en.htm>

³¹ <https://www.oecd.org/chemicalsafety/risk-assessment/validationofqsarmodels.htm>

³² <https://qmrf.sourceforge.net>



for Alternatives to Animal Testing (EURL ECVAM). QMRF is intended for summarizing and reporting key information on (Q)SAR models, including the results of any validation studies. The information is structured according to the OECD (Q)SAR validation principles. Read-across together with (Q)SAR represent the non-test methods, and read-across fully becomes a QSAR when automated, e.g., as read-across-based structure activity relationships (RASAR) (Hartung, 2016a; Luechtefeld et al., 2018) or generalized read-across (Helman et al., 2019). As read-across turned out to be initially the most used NAM in REACH, we initiated good read-across practices (GRAP) (Ball et al., 2016; Zhu et al., 2016) and worked toward the international acceptance of read-across (Chesnut et al., 2018; Rovida et al., 2020). However, ECHA's publication of a Read Across Assessment Framework (RAAF)³³ around the time of the publication of GRAP limited GRAP's usefulness as it was not aligned with RAAF. The standards set by RAAF are very high, so that in fact it is often less effort to carry out the animal tests, and the windows of opportunity for pragmatic shortcuts have become very small. Nevertheless, revamping the GRAP program is currently under discussion.

Taken together, the critical role of good practices and other QA measures including validation is central to Tox-21c. With GCCP 2.0 and other ongoing developments, we are on the right path, but QA needs to be implemented at every step. Aristotle (384–322 BC) put it, “*Quality is not an act, it is a habit.*”

Table 1 lists prominent quality assurance schemes for carrying out (upper row) and reporting (lower row) studies relevant to toxicology. For references and abbreviations, see above.

2.8 Challenge 8:

How to change with step-by-step developments becoming now available?

A fundamental challenge to Tox-21c is the integration of different traditional and new evidence streams. How to handle contradictory findings? When to trust negative ones? We have recently laid out (Maertens et al., 2022) how a transition to probabilities of hazard and risk, respectively, might aid this process. If hazard is no longer seen as a black-or-white property but as a probability (shades of grey with uncertainty), we can accept more easily any piece of evidence as a shift in grey and uncertainty. As discussed in Maertens et al. (2022), AI lends itself to such analyses as the machine learns to weigh different pieces of evidence to make a prediction. The move to explainable AI provides a type of sensitivity analysis that will enable us to understand what different methods contribute to our judgements and, the other way around, allows an information economy, where we can calculate whether the investment into further testing is actually likely to change our assessments to a major extent. A workshop on *Probabilistic Risk Assessment* was held in Ranco, Italy, in July 2022, to expand on these concepts, and the report is in preparation.

2.9 Challenge 9:

How to organize transition?

The WHO Constitution (1946) envisages “...the highest attainable standard of health as a fundamental right of every human being.” From a public health perspective, this implies the prevention of disease through avoidance of threatening exposures. Risk sciences have helped us a lot in this direction; however, the progress of the past is becoming a hinderance to moving forward. We fail to renovate our approaches. There is no other science where fundamental experimental approaches have not changed for some sixty years. John F. Kennedy (1917 – 1963) said, “*And our liberty, too, is endangered if we pause for the passing moment, if we rest on our achievements, if we resist the pace of progress.*” I would like to make the case that the methods for Tox-21c are largely available, and the bigger challenge is bringing them into use. Several disruptive technologies are currently changing the way we discover, innovate, produce, and live. This is most evident in information technologies, artificial intelligence, or sensor technologies. These are mostly democratizing technologies, which are giving more and more people access to their benefits. However, biotechnology is often a different story as bioengineering and its products are costly. Their benefits come at a price and have the potential to increase inequality within and between nations. The open access movement of information technologies applies here, if at all, to sharing of knowledge through open access publishing. The enormous acceleration of innovation with an abundance of new technologies raises questions, for example, whether we should revise the patent system. The first-mover advantage already often outweighs legal protections, which too often hinder the field and the spread of knowledge.

Biotechnologies are currently revolutionizing biomedical research, e.g., with microphysiological systems (MPS), allowing to produce more human-relevant information faster and faster. They are also allowing the production of novel foods and drugs. However, they also play enormous roles in identifying environmental threats, mitigating contaminations, or cleaning wastewater. Social justice requires equality and equity in benefiting from these technologies (Aschner et al., 2021). Often these innovations are driven by the entrepreneurial state (Mazzucato, 2013), i.e., the investment of public money with a long-term vision as contrasted by the short-term shareholder value-driven perspective of industry. This allows coupling such development programs to goals of social justice. If all are paying for it, all should benefit from it. Open access to knowledge on innovation plays a key role here. However, engagement, participation, and collaboration are needed to allow all to keep pace with the accelerated generation of disruptive technologies.

How we assess and manage risks to our health is a primary governmental role, weighing these needs against societal costs and economic developments. Therefore, it is critical to assess

³³ https://echa.europa.eu/documents/10162/13628/raaf_en.pdf/614e5d61-891d-4154-8a47-87efebd1851a

economical (Meigs et al., 2018) and societal (von Aulock et al., 2022) aspects of a transition to Tox-21c. The Tox21 alliance of US agencies or the European PARC initiative, which actually has some similarities to the European Safety Sciences Institute (ESSI) we called for earlier (Busquet and Hartung, 2017), might be fertile ground to further grow the implementation, but the case is made below that we need a global Human Exposome Project at the basis of such transition.

2.10 Challenge 10:

Making it a win/win/win situation

Stakeholders from academia, regulators, and the regulated communities in industry need to collaborate on an international level to renovate safety assessments. Chemicals do not have geographical or industrial sector borders. Some chemicals are used in more than 50,000 products. The EU aims for a one-substance-one-assessment (OSOA) approach as laid out in a joint position paper of EFSA and ECHA³⁴. A big question is whether this will make global harmonization more challenging?

In principle, all stakeholders agree that toxicology needs to be based on sound science. Unfortunately, this sound science is often the science of 30 years ago. On the one hand, this is the science from the time when today's decision-makers studied. On the other hand, it reflects the time it takes for standardization, validation, and acceptance, as illustrated by the author's involvement in alternatives to the rabbit pyrogen test earlier (Hartung, 2015, 2021a). In toxicology, we still discuss the regulatory use of methods developed around the turn of the century. So, a system that embraces more recent developments faster will win, as it diminishes the disconnect between modern toxicology and regulatory practices.

The disconnect between regulating and regulated community is less about identifying risks – no company wants to poison their consumers. The dissent is typically about uncertainty and necessary precaution. It is also about the relevance of animal data for humans: A large part of investigative toxicology (Beilmann et al., 2019) in pharmaceutical industry is about “de-risking” of substances, i.e., showing that product development can proceed despite flags from animal testing. Making these uncertainties and their consequences more obvious (probabilistic risk assessment) and treating existing data more objectively (evidence-based toxicology) should help these discussions.

Overall, the possible advantages of Tox-21c to regulation lie mostly in their human relevance and the throughput of NAMs, the latter also possibly translating to accelerated evaluations, i.e., time to market for industry and more exhaustive evaluation from the viewpoint of regulators. These are the benchmarks to show stakeholders the win/win/win of a new approach.

3 The exposome and exposomics

Exposure is one of three principal pathogenic paths (Fig. 1). Living beings are biochemical systems in action. Millions of our bodies' own chemicals meet millions of substances in the environment, in its totality forming the human exposome as summarized earlier in this series (Sillé et al., 2020) and by others lately (Miller, 2020; Zhang et al., 2021; Price et al., 2022). Some of these exogenous chemicals such as nutrients are essential, some such as pharmaceuticals and remedies are beneficial, many are inert, and some are detrimental as toxicants – and for most we simply do not know. Even for those we believe to know, we often must revise our judgments, or it simply depends on the circumstances, the individual, and most importantly, the amount. Coffee is a good example, where our critical views had to be revised lately. For nutrients such as cholesterol and sugar our assessments are shifting in opposite directions, making it difficult to adapt our lifestyle. Nassim N. Taleb (1960 –) nicely summarized, “*The world has changed too fast for our genetic makeup.*”

Scientific advice on avoiding certain exposures or lifestyles is remarkably difficult, also because commercial interests interfere. In the meantime, we must take decisions based on imperfect tools applied to far too few chemicals. AI is emerging as a tool to integrate existing evidence and extrapolate to those chemicals for which we have no data. This allows focusing resources, but the integration of such information into politics and public health action has not even started. On top of this, lifestyle, cultural, and value decisions must be taken into account. Public health focuses on longevity and health but struggles with personal choices such as drug abuse, alcohol, and tobacco. Cannabis legalization, the weighing of positive versus negative effects of alcohol (Grønbæk, 2009; Sayette, 2017), and the steering of smoking toward much less damaging non-combustion products (Hartung, 2016b,c,d; Fowle et al., 2017) are challenging debates. These debates are already difficult in the transatlantic divergence of political, legal, and regulatory systems. They cannot be simply exported to other parts of the world, which have different cultural choices, economic opportunities, and exposure challenges. However, the novel tools promise strongly broadened risk characterization as the basis for informed risk management tailored to the needs of different individuals and nations.

After 40 years of the Human Genome Project, we can explain about 5% of diseases as caused by genetics, and ca. 40% have a genetic component (40% of 560 diseases had a genetic component in Lakhani et al., 2019). However, an estimated 70-80% of diseases are caused or aggravated by exposures (Rappaport and Smith, 2010), using a broad definition of exposure. For the much narrower exposure to chemicals, a systematic review of the global burden of disease by Prüss-Ustün et al. (2011) showed for the year 2004, in total, 4.9 million deaths (8.3% of total) and 86 million disability-adjusted life years (DALYs)

³⁴ https://www.efsa.europa.eu/sites/default/files/corporate_publications/files/EFSA-ECHA-position-paper-OSOA.pdf



(5.7% of total) attributable to environmental exposure and management of selected chemicals. In 2016, the World Health Organization (WHO) published *The public health impact of chemicals: knowns and unknowns*³⁵ with an addendum,³⁶ updating the main data tables and figures for the year 2019. The 2021 data addendum estimates that 2 million lives and 53 million DALYs were lost in 2019 due to exposures to selected chemicals. This is higher than the estimate of the previous data addendum of 1.6 million lives and 45 million DALYs lost in 2016. “Nearly half of deaths attributable to chemical exposures in 2019 were due to lead exposure and resulting cardiovascular diseases. The other largest contributors were chronic obstructive pulmonary disease (COPD) from occupational exposure to particulates and cancers from occupational exposure to carcinogens. ... Data are however only available for a small number of chemical exposures and people are exposed to many more chemicals every day.” Grandjean and Bellanger (2017) calculated that environmental chemical exposures contribute costs that may exceed 10% of the global domestic product and that current DALY calculations substantially underestimate the economic costs associated with preventable environmental risk factors. This shows the potential of a Human Exposome Project, which would bring toxicology to another level.

At the same time, a Battelle report from 2011, updated 2013³⁷, is titled *Economic impact of the Human Genome Project – how a \$3.8 billion investment drove \$796 billion in economic impact, created 310,000 jobs and launched the genomic revolution*. The title says it all. The report appraised the large and widespread economic and functional impacts³⁸: “Between 1988 and 2010 the human genome sequencing projects, associated research and industry activity – directly and indirectly – generated an economic (output) impact of \$796 billion, personal income exceeding \$244 billion, and 3.8 million job-years of employment. In the 2013 update, these numbers increased to economic (output) impact of \$965 billion, personal income exceeding \$293 billion, and 4.3 million job-years of employment.” Peter Diamandis³⁹ recently summarized: “The first human genome cost about \$3 billion. The next about \$100 million. Since then, the cost has been dropping at 5 times the speed of Moore’s Law. Genome sequencing has led to multiple advances in medicine: from blood tests that can detect cancer early and genetically-targeted drugs, to rare disease diagnosis and even the Covid-19 vaccines.” These comments in his newsletter were prompted by Illumina just unveiling its newest genome sequencing machines⁴⁰: NovaSeq X series. The machines are the company’s most cost-efficient and fastest yet and can sequence a human genome for \$200 (compared to \$10,000 a decade ago and \$600 today) and produce a readout twice as fast. Illumina says the NovaSeq X series machines will cost around \$1 million and generate 20,000 whole genomes per year. This illus-

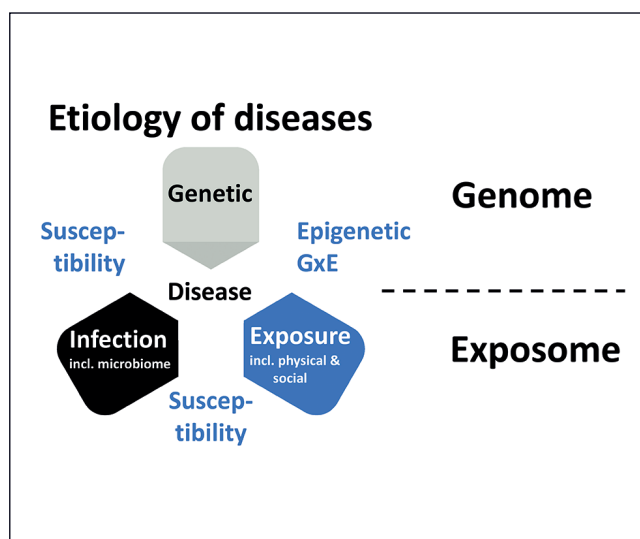


Fig. 1: Etiology of disease – genes x environment x infection

The three principal paths to disease, genetics (genome), and infection and exposure (exposome) are depicted. The interplay of genes x environment manifests as epigenetics, and both exposure and genetics can modify the susceptibility to infection.

trates nicely how technologies progress once there is a demand and a scientific opportunity is identified.

The transformation of toxicology envisioned here as a Human Exposome Project (see Section 8) represents a similar opportunity, promising to identify an even larger fraction of causes of disease and opening up new opportunities for prevention and cure. However, there are several challenges to address for toxicology to embrace an exposure-driven approach:

- Exposure is not a single chemical, but multiple chemical agents are present in multiple media (air, water, food, products...).
- Utilizing real-life exposures not maximum tolerated doses
- Physical agents: noise, light, climate interact with these.
- Social agents: racism, economic deprivation, social support
- Multiple mechanistic pathways at multiple “nodes” in the network
- Considering multiple health outcomes
- Set exposure thresholds from new approaches
- Ultimately informing the identification and evaluation of effective interventions to improve overall health.

The next two sections will address two disruptive technologies to identify hazards, i.e., the determination of adversity of exposures *in vitro* and through AI.

³⁵ WHO/FWC/PHE/EPE/16.01: <https://www.who.int/publications/i/item/WHO-FWC-PHE-EPE-16-01>, <https://apps.who.int/iris/rest/bitstreams/916484/retrieve>

³⁶ <https://apps.who.int/iris/rest/bitstreams/1354311/retrieve>

³⁷ https://web.ornl.gov/sci/techresources/Human_Genome/publicat/BattelleReport2011.pdf

³⁸ https://web.ornl.gov/sci/techresources/Human_Genome/project/economics.shtml

³⁹ Peter Diamandis, email newsletter “Top 22 Breakthroughs of 2022 vs. 1922” from 28 Dec 2022

⁴⁰ <https://www.wired.com/story/the-era-of-fast-cheap-genome-sequencing-is-here/>

4 Determining adverse responses using *in vitro* assays

The central question of how to define where adversity on an organism level starts from *in vitro* systems was addressed in a workshop in June 2015 at Brown University, organized by Kim Boekelheide and sponsored by Brown University, the Human Toxome Project, and CAAT and its Transatlantic Think Tank for Toxicology (t⁴). There were 15 workshop participants⁴¹ from academia, government, and industry. Participants were present as individuals with their own opinions, and not as representatives of any entity. Unfortunately, the workshop report was never completed. In the following, some parts of the remnant document are summarized by the author in order to preserve some of these discussions; important to note, these were not formally authorized by the workshop participants.

In the context of biology/toxicology and risk assessment, the *in vitro* determination of adverse responses serves two related but distinct purposes. On the one hand, an *in vitro* adverse response identifies a cell autonomous mechanism of injury indicative of an irreversible alteration in cellular homeostasis. On the other hand, an adverse response in an *in vitro* assay can be used to satisfy information requirements for hazard and safety/risk assessment purposes by being predictive of a toxicological endpoint or health effect of regulatory importance (e.g., cancer or reproductive toxicity). A fundamental requirement to fulfill these purposes is the need to distinguish an adverse response, defined as a reversible or irreversible cellular response that may progress to a pathophysiological alteration, from an adaptive response, defined as a protective cellular response that is activated by a stressor and returns to baseline upon removal of the stressor (Blaauboer et al., 2012). Computational systems biology approaches (Hartung et al., 2012, 2017) are used to identify the molecular and cellular perturbations associated with such irreversible alterations in cellular homeostasis, distinguishing between adaptive and adverse responses to identify points of departure (PoD).

The workshop was held to define and identify the relationship between adaptive and adverse responses, to explore the characteristics of upstream impairments of fundamental biological function associated with adversity (many of which are common across anatomical/physiological sites/systems), and to contemplate how to use this information in a regulatory framework defined by apical pathological alterations and disease endpoints. The workshop discussion was guided by the following key questions:

- 1) How are *in vitro* adaptive and adverse responses defined, identified, and distinguished?
- 2) How do toxicity pathway and adverse outcome pathway responses inform the determination of *in vitro* points of departure? Included in the discussion of this question: How does the relative non-selectivity/promiscuity of many industrial

and environmental chemicals impact the determination of *in vitro* points of departure?

- 3) How can *in vitro* approaches that measure alterations in cellular homeostasis facilitate the evaluation of potential adverse effects for the purpose of safety/risk assessment?

4.1 Key question #1⁴²: How are *in vitro* adaptive and adverse responses defined, identified, and distinguished?

4.1.1 Historical context: The cell is the fundamental reactive unit in disease

Pathologists traditionally recognize and diagnose disease based on altered morphology of cells and tissues using light microscopy. The founder of modern pathology, Rudolf Virchow (1821 – 1902), first articulated that the cell is the fundamental reactive unit in disease (Virchow and Chance, 1860). During the 20th century, understanding of the pathogenesis of disease was expanded to encompass structural, biochemical, and molecular abnormalities occurring at cellular and subcellular levels. Integration of these abnormalities provides a fundamental understanding of disease pathogenesis. This integrated model of the cellular responses to injury (Boekelheide and Campion, 2010) can be applied to toxicity testing based on identification of intracellular toxicity pathways that are perturbed in a dose and time-dependent manner in response to chemicals, drugs, and environmental toxicants (Bhattacharya et al., 2011). Toxicity pathways are defined as biological pathways that can lead to adverse health effects when sufficiently perturbed (Krewski et al., 2010). A quantitative mechanistic understanding of toxicity pathways is fundamental for alternative toxicity testing in the 21st century.

4.1.2 Defining and characterizing cellular-level responses

A key challenge in characterizing the nature of cellular-level responses to chemical exposures is the potential for many factors, both endogenous and exogenous, to modulate the progression to tissue-, organism-, and population-level effects. Clearly, a one-to-one correspondence between cellular- and organismal-level effects does not exist. On the one hand, organisms are by and large resilient and therefore have multiple mechanisms for adapting to and repairing cellular-level injury so that the physiological functions of cells or tissues are maintained (Smirnova et al., 2015). Conversely, cellular-level alterations that do not manifest as injury *per se* may be sufficient to reduce resiliency to further injury, thereby increasing an organism's vulnerability to other endogenous and exogenous stressors. Additionally, the same cellular response that may be innocuous in some biological contexts may have adverse consequences in other contexts (e.g., life-stage susceptibility).

⁴¹ Kim Boekelheide (Brown University), Tara Barton-Maclaren (Health Canada), Mounir Bouhifd (CAAT), Paul Carmichael (Unilever), Weihsueh Chiu (Texas AM University), Rebecca Clewell (The Hamner Institutes of Health), Thomas Hartung (CAAT), Daland Juberg (Dow AgroSciences), Agnes Kane (US EPA), Patrick McMullen (The Hamner Institutes of Health), Andy Nong (Health Canada), Imran Shah (US EPA), Russell Thomas (US EPA), Maurice Whelan (European Commission, Joint Research Centre), James D. Yager (Johns Hopkins Bloomberg School of Health)

⁴² Group members: Boekelheide, Chiu, Kane, McMullen, Yager

To navigate the complexity of characterizing cellular-level responses, the following three definitions are adopted:

- **Adaptive cellular response:** A protective cellular response that is activated by a stressor and returns to baseline upon removal of the stressor.
- **Reversibly vulnerable cellular response:** A response that is reversible at the cellular level with cessation of exposure but may have irreversible consequences at the organism level due to additional stressor exposures that occur concurrently with the response.
- **Adverse cellular response:** A reversible or irreversible cellular response that may progress to a pathophysiological alteration.

In defining “adaptive” and “adverse” responses, it became clear that these domains of potential biological response were not mutually exclusive and that a third category of “*reversibly vulnerable*” was needed to describe responses that could be either adaptive or adverse, depending upon context. As illustrated in Figure 2, while “purely” adaptive effects will always ultimately have no effect at the organismal level, reversibly vulnerable effects fit the definition of adaptive response, but could in some cases contribute to injury or to disease depending on the presence of other stressor exposures. Additionally, there are clearly adverse effects defined at the cellular level (e.g., cellular apoptosis), even though in some cases these may ultimately resolve into “no effect” at the organismal level due to repair or adaptation at the tissue- or organismal-level. Taking a probabilistic point of view, *in vitro* adaptive cellular effects are defined as those that lead to no increase in the probability of an organism-level adverse effect, *in vitro* reversibly vulnerable cellular effects are those that increase the probability of an organism-level adverse effect conditional upon the concurrent presence of some specific additional stressor, and *in vitro* adverse effects are those that contribute to an increase in the probability of an organism-level adverse effect in a broad set of circumstances.

A more fine-grained characterization of these responses beyond these three categories can be based on the following three key characteristics:

- **Physiologic function:** Does the cell maintain (or enhance its) physiologic functions, or have cellular functions been reduced or impaired?

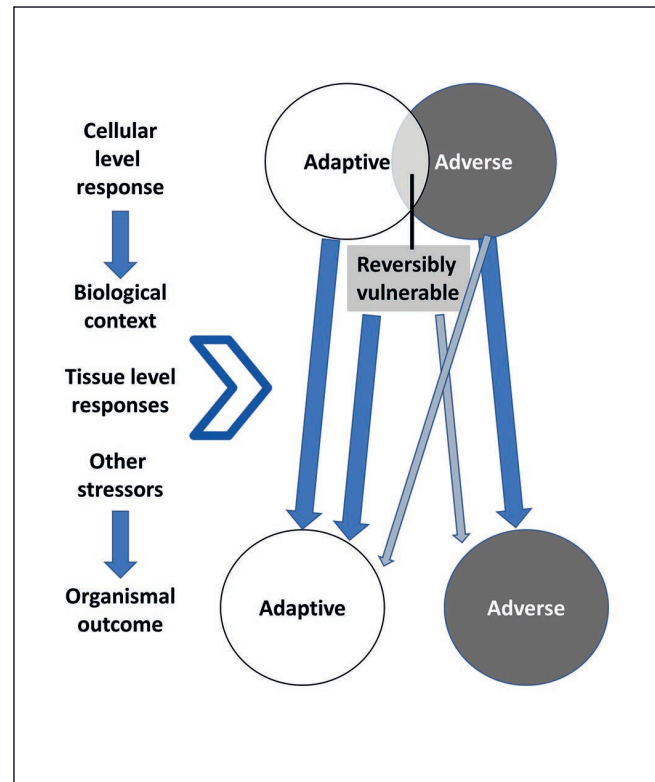


Fig. 2: The complex relationship between cellular, tissue, and organismal adaptive and adverse responses

An intermediate response category, reversibly vulnerable, is introduced to explain adaptive responses that might lead to injury or disease after co-exposure to a subsequent stressor.

- **Resiliency to additional stressors:** Does the cell maintain (or enhance its) resiliency to additional stressors, or has its ability to adapt or maintain its functions in the presence of additional stressors been impaired?
- **Reversibility with cessation of exposure:** Does the cell return to its initial homeostatic state if exposure is removed, or does it enter an altered homeostatic or dysregulated state?

Cellular response characteristic	Adaptive	Reversibly vulnerable	Adverse				
Physiologic function ("+" for up, "-" for down)	+	+	-	-	+	-	-
Resiliency to additional stressors ("+" for up, "-" for down)	+	-	+	-	-	+	-
Reversibility with cessation of exposure ("+" for yes, "-" for no)	+	+	+	+	-	-	-

Fig. 3: Adaptive versus adverse reactions

The cellular characteristics of physiologic function, resiliency, and reversibility are used to characterize further *in vitro* adaptive (white), reversibly vulnerable (grey), and adverse (black) responses.

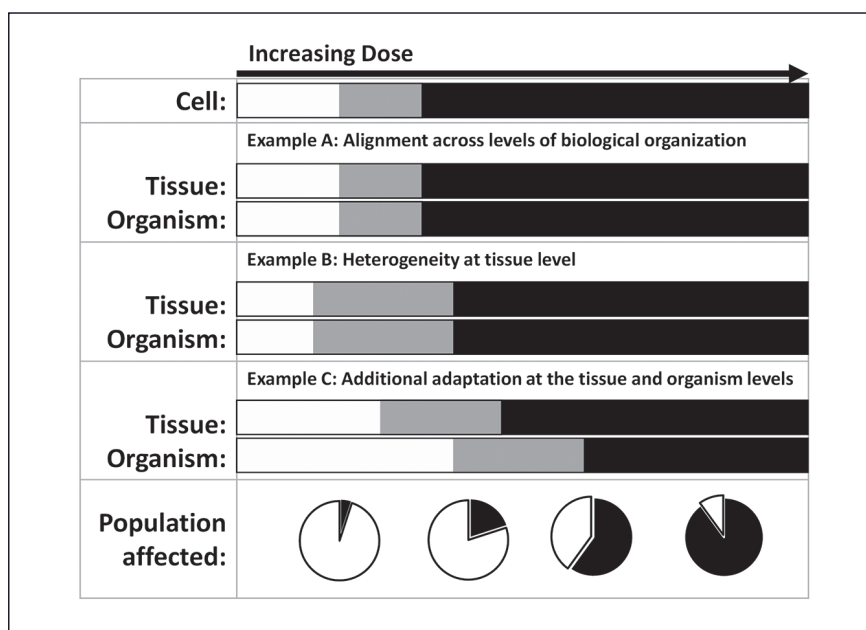


Fig. 4: Extrapolation from cell to tissue, organism, and population responses involves consideration of heterogeneity and compensation

From left to right is shown the effects of increasing dose causing a transition from an adaptive (white) to reversibly vulnerable (grey) to adverse (black) response. The rows depict responses for different levels of organization (cell, tissue, organism, population) in the context of the examples discussed in the text. A key point is the expectation that responses are stochastic, with an increasing likelihood of an adverse effect with increasing dose.

Based on these three characteristics, the niche characteristic “reversibly vulnerable” is clearly evident in the transition between adaptive and adverse responses (Fig. 3). Specifically, the transitional category of reversibly vulnerable responses is defined by the situation in which a stressor has temporarily impaired the resiliency of the cell to additional stressors, thereby making it more vulnerable to injury or irreversible alterations.

4.1.3 Extrapolation from cell to tissue, organismal, and population responses

The progression of responses from cells to tissues to organisms and to populations can proceed in multiple ways (Fig. 3). In some cases, there may be concordance or alignment across levels of biological organization, so that adaptive/reversibly vulnerable/adverse effects observed at the cellular level correspond to the same outcomes at higher levels of biological organization (Fig. 4, Example A). Or, the region of reversible vulnerability may occur over a wider range of doses at the “tissue” or “organism” levels due, for example, to heterogeneity of the cellular response (Fig. 4, Example B). Finally, consider the possibility that additional adaptation at the tissue level may produce greater resiliency at higher levels of biological organization, causing the transition to adverse effects to occur at progressively higher doses going to the tissue and then the organismal level (Fig. 4, Example C). An additional feature of this progression is that, as dose increases, the fraction of the population anticipated to manifest an adverse effect also increases (Fig. 4, bottom row). It should also be noted that for reversibly vulnerable responses, only a fraction of the population is likely to have concurrent stressors that could lead to an adverse effect at the organismal level, since the vulnerabilities in this transition region will tend to be specific to the elicited cellular response. As the dose increases and cellular responses become adverse, the risk of an adverse organismal-level effect is

no longer contingent upon other stressors, so a broader population would be expected to be affected.

4.1.4 Examples of reversible vulnerability resulting in adverse effects

Co-exposure to additional toxicants that interact with adaptive cellular responses is one mode by which such responses may increase an organisms’ vulnerability. To explore the nature and scope of how reversible vulnerability could result in adverse effects, specific examples were considered, including xenobiotic metabolizing enzyme induction, arrhythmias following inhibition of the hERG ion channel, and epigenetic effects.

Metabolism of chemicals

At the cellular level, many chemicals at low doses and at early times following exposure may elicit an adaptive response such as induction of Phase I P450s and/or Phase II enzymes that enhance their metabolism and excretion at the organismal level. However, while this process, reversible upon withdrawal of the chemical, may initially reduce the toxicity of the chemical and allow the cell/tissue/organ/organism to maintain its normal physiological functions, it may also render the cells, and thus the tissue/organ and ultimately the organism, more susceptible to concomitant or subsequent exposure to another chemical. In this situation, the cell could be considered to be in a ‘reversibly vulnerable’ state. This increased sensitivity could be due to its increased metabolism of a second toxicant, reduced protective capacity such as reduced glutathione level, increased reactive oxygen species resulting from the metabolites, the metabolism process itself, or other mechanisms.

At the organism level, this phenomenon is well documented for exposures to haloalkanes such as ethanol, chloroform, isopropanol, ketone and others where prior exposure to one chemical



leads to enhanced toxicity upon exposure to another in the same chemical class. Another well-known chemical-chemical interaction is that between ethanol and acetaminophen. Acute ethanol exposure is protective against acetaminophen toxicity, whereas the toxicity of acetaminophen is potentiated by chronic ethanol consumption (Lieber, 1980). Saturation of phase II sulfation/glucuronidation pathways, induction of CYP 2E1, and depletion of glutathione levels by exposure to ethanol or other chemicals or drugs can result in increased hepatotoxicity of acetaminophen.

hERG ion channel blockage

Numerous chemicals are known to lengthen the QT interval via blockage of the hERG ion channel. Based on pharmacokinetic/pharmacodynamics model-based analysis, even small responses (5-10%) in an *in vitro* hERG assay are predictive of measurable lengthening of QT (Jonker et al., 2005). While such limited inhibition of hERG ion channel transport would not impair cardiac function at the cellular level, this effect does have potential adverse consequences. In particular, multiple studies have demonstrated that cardiovascular risks are associated with QT interval on a continuum, so that even in the “normal range,” QT prolongation is associated with increased cardiovascular-related mortality (Nielsen et al., 2014; Beinart et al., 2014; Zhang et al., 2011). Moreover, such increases are observed even in individuals without existing cardiovascular disease. Thus, small amounts of hERG channel blockage represent a reversible, cellular-level response that does not impair cellular function but may lead to an adverse organismal consequence in the presence of additional stressors by increasing the likelihood that such stressors result in torsade de pointes, thereby triggering myocardial infarction. Stressors relevant to this increased vulnerability are not limited to exogenous chemical exposure, since increased adrenaline (e.g., from exercise or stress) can trigger an infarction.

Epigenetic effects

A reversibly vulnerable state could arise after a chemical exposure that causes epigenetic changes that result in altered gene expression. Epigenetic changes include alterations in DNA methylation, histone posttranslational modifications, or changes in microRNA expression (Aguilera et al., 2010). Two settings that illustrate this concern are epigenetic alterations that modify the expression of protooncogenes or tumor suppressor genes in proliferating somatic cells, increasing the risk of cancer (Jones and Baylin, 2007; Ting et al., 2006), and chemical exposures *in utero* resulting in epigenetic alteration of the expression of genes in developing gonadal tissue that produce transgenerational effects (Aguilera et al., 2010; Skinner, 2014; Manikkam et al., 2014). While the mechanisms for transgenerational effects are not well understood, a growing literature implicates epigenetic pathways as responsible. In a study using the yellow Agouti mouse, it was demonstrated that maternal dietary exposure to bisphenol A altered the F1 coat color distribution from pseudo-agouti to yellow. This was accompanied by a reduction in CpG methylation in the

upstream promoter region that controls expression of the Agouti gene (Dolinoy et al., 2007). This effect could be reversed by simultaneous exposure to dietary genistein. These observations indicate a need to develop high-throughput screening assays designed to detect epigenetic effects of chemicals and their potential to cause adverse transgenerational health outcomes.

4.2 Key question #2⁴³: How do toxicity pathway and adverse outcome pathway responses inform the determination of *in vitro* points of departure? How does the relative non-selectivity/promiscuity of many industrial and environmental chemicals impact the determination of *in vitro* points of departure?

4.2.1 Regulatory decision context

Very different risk assessment frameworks are in place for different product categories and different geographical regions. This workshop focused primarily on the context of industrial chemicals. Very different to the rather prescriptive frameworks for drugs or pesticides, where a catalogue of tests has to be completed for registration with relatively little leeway for variation from standard regimens, industrial chemicals are regulated with more room for maneuver. In the US, under the Toxic Substance Control Act (TSCA), the majority of substances brought to the market before 1982 were grandfathered with respect to safety assessments and fall under the liability of producers and marketers. Most (7 of 8) required premarketing notifications came without safety data, and the EPA as administrator had just 90 days to demand testing, which was rather rare. This changed with the 2016 TSCA reauthorization. Beside this, a rather prescriptive testing scheme for endocrine disruptors was put into place, which foresees defined batteries of tests for screening and definitive testing.

Very differently, in Europe, the previous scheme under the Dangerous Substance Directive was rather prescriptive with a certain set of tests depending on production volumes. Since 2007, the European regulation on chemicals, REACH, requires registration of all new and existing chemicals above a production or marketing volume of one ton per year with a reduced test requirement scheme. However, most importantly, the legislation foresees many opportunities for deviations from standard testing requirements, promoting the use of non-animal and non-test methods as well as existing data (Hartung, 2010c). The integral use of this information should limit the traditional testing demand in animal tests. Noteworthy, substances used exclusively for cosmetics are exempted from REACH and fall under an animal testing ban since the entry into force of the 7th amendment of the Cosmetics Directive; since March 2013, this ban is complete for all hazards independent of the availability of alternative methods.

The different and emerging regulatory frameworks for industrial chemicals on both sides of the Atlantic allow considerable freedom to use new approaches for testing and priority setting. On the EU side, the focus is at this moment largely on the side of replacing traditional testing, while the US approach focuses

⁴³ Group members: Barton-Maclaren, Clewell, Hartung, Thomas, Whelan

more on priority setting and defining testing needs. They have in common an increasing unease about the economical and animal burden of traditional tests and increasing understanding of their limitations in predicting human health hazards and risks.

With growing understanding of the modes of action, adverse outcome pathways, and molecular pathways of toxicity, there is the expectation that these can inform and optimize or even ultimately replace current approaches. Challenges lie in the deduction of the mechanistic processes, their annotation, their quality control (validation), and putting them into use for regulatory decision-making. Here, specifically the aspect of deciding on thresholds of adversity must be addressed, i.e., how mechanistic information can be translated into a hazard / no-hazard decision and how the respective tipping points of chemical concentrations can be translated into anchors of a dose paradigm, similar to lowest observed effect levels (LOEL).

4.2.2 Chemical/data matrix characteristics – limited mechanistic knowledge

The basis for the deduction of mechanistic information and its integration into pathways is the production of respective data. The availability of new information-rich technologies (high-content and high-throughput) can facilitate this. Curating large datasets by compiling available data from many sources (e.g., PubChem, ToxRefDB), from high-throughput *in vitro* testing (e.g., ToxCast, Tox21) or omics approaches (e.g., Maertens et al., 2015, 2017; Pendse et al., 2017), and even *in silico* predictions (Luechtefeld et al., 2018) and making them publicly available promises the deduction of such mechanistic knowledge. This has led to the expectation that mining these data will allow the identification of pathways, at least signatures of features such as structure and/or biological activities, which correspond to hazard manifestation.

Very often this means that known or supposed pathways are remodeled, heavily biased by animal model-derived understanding. Another key problem is that the organism or cell model collapsing upon intoxication shows multiple derangements, and it is difficult to figure out, what the Achilles heel is, i.e., the most sensitive target or pace-maker, as these different derangements take place in short sequence and at similar dose levels. The fact that most substances are studied at high dose levels and using rather acute exposure scenarios further biases this analysis. We might be looking too often for mechanisms that are not relevant for the low-dose chronic exposure we typically should worry about.

4.2.3 Assay development and design for maximal biological coverage

For practical reasons, the number of assays and endpoints needs to be limited. This is not only a problem of economics but also of signal/noise ratio and over-fitting. Too many non-meaningful data dilute the signal and add noise; large datasets of many measured variables allow fitting any assumption from a point of reference. Thus, selection of the meaningful assays and parameters has to restrict the dataset, feature elimination is needed, and steps of confirmation by external challenges to the forming hypothesis are required in an iterative process to optimize and reduce data acquisition. So, while broad biological coverage is desirable

when screening for candidate biomarkers and pathways, a selective approach is needed for actual testing.

Broad biological coverage is very desirable when establishing biological similarity between substances, which is increasingly considered to strengthen read-across, i.e., data gap-filling by using toxicity data from similar substances with test results. Such early phase broad biological coverage is achieved by two means – broad biological systems (cell types and development stages, i.e., the targets such as molecular initiating events and the machinery of the initiated pathways, the key events, need to be present in a functional manner) and the broad measurement by high-content methods or high-throughput testing. Importantly, relevant concentrations of test substances, redundant chemistry, sufficient replicate measurements, and functional endpoints need to be safeguarded.

4.2.4 Unsupervised analysis (signatures or bioactivity)

There are two principally different approaches to interpret the large toxicological dataset, i.e., supervised (targeted) or unsupervised (untargeted). The former approach uses a hypothesis from former knowledge and carries the respective limitations and biases – it is difficult to find something new. The latter carries the problem of too many possible associations between the multitude of endpoints while the number of measurements is always limited. This is further impaired by the noise of these methods, i.e., both reproducibility and systematic errors. Both methods have their strengths and weaknesses, and in the end, they need to be combined.

Typically, untargeted methods, which are often belittled as fishing expeditions, lend themselves especially in the beginning as they allow to find unknown connections, the candidate mechanisms and biomarkers. They then need to be sorted out by targeted approaches to see whether they make biological sense and have predictive value.

4.2.5 Adversity versus safety prediction

The traditional approach in toxicity testing is testing for the potential of a substance to elicit a given hazard, often rendered more sensitive by high-dose testing and use of the most sensitive species. This has been mirrored in *in vitro* toxicity tests, which usually aim to replace the traditional animal test one-by-one. This has turned out to be especially difficult for the complex systemic and chronic toxicity tests, where a multitude of mechanisms and endpoints can be involved. An *in vitro* adversity prediction therefore would require a test battery that reflects this complexity. However, in the best case, this leaves us at the level of hazard (classification and labeling), with all problems of over-sensitive animal tests, species-differences, and high-dose bias of the reference data from animals. The vision for many researchers in the field is to identify the most meaningful of these assays and establish a quantitative *in vitro* to *in vivo* extrapolation (QIVIVE). This always bears the risk of missing something important. The major challenge then is to enable a quantitative risk assessment. What is the relevant concentration, what is the translation to a relevant health effect?

A completely opposite approach is to not worry about the completeness of the test battery but establish at which concentration a substance starts to impact biological systems. This is based on the observation that many biological functions in different cell

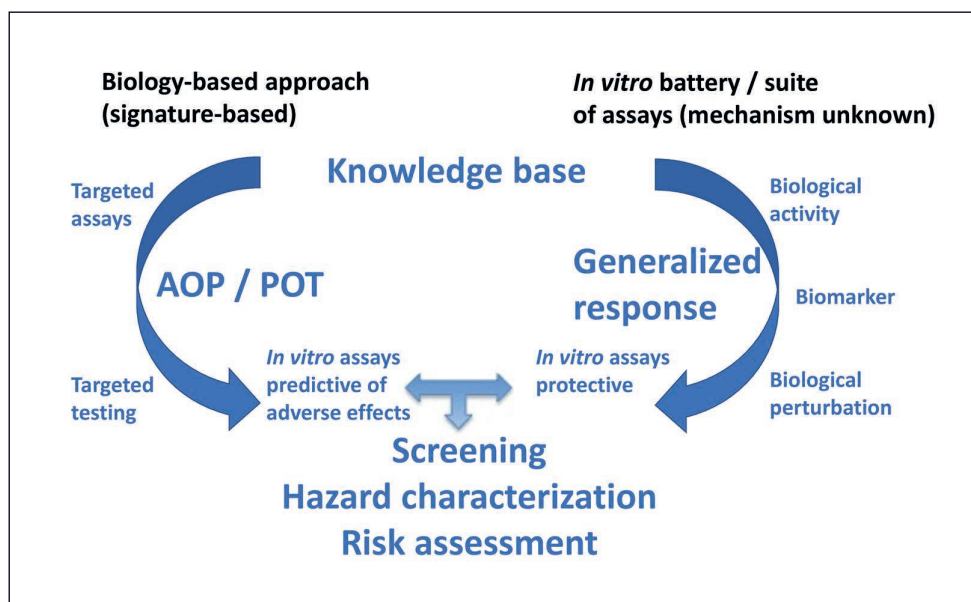


Fig. 5: A biology-based approach versus a safety-based approach to *in vitro* hazard and risk assessment

The figure depicts two different approaches to using *in vitro* assays to establish risk versus safety. (1) Targeted assays based on adverse outcome pathways (AOP) / pathways of toxicity (PoT) are combined to identify possible adverse reactions. (2) A battery of tests is used to identify in an untargeted manner the perturbation of biology based on the observation that for most chemicals perturbation affects many pathways in a narrow concentration range, and thus a point of departure can be defined.

systems start to be disturbed in a narrow range of concentrations. Using some safety factors, this allows establishing a minimum concentration reflecting a tissue concentration that would need to be achieved to exert any biological (toxicological) effect. If this can be linked by QIVIVE to an exposure necessary to achieve this concentration, we can estimate whether such an exposure can actually be achieved or, the other way around, whether the given use is safe. The two concepts are depicted in Figure 5.

The question is, how can we design a workflow, i.e., what is the purpose of the *in vitro* battery? Do we aim to predict an apical endpoint, i.e., a specific hazard (which often does not translate between species) or be protective by defining a point of departure to establish safety. The latter means that the data generated have biological relevance (relevant concentrations) but do not have an explicit mechanistic underpinning (the specific test turning positive). The question is, ultimately, are we predicting a specific hazard or the starting dose of a biological response? In other words, do we establish a dose-response for a specific apical endpoint versus dose-response for the most sensitive biological activity? Is the difference data-driven (left), where the knowledge on pathways establishes the relevant assays in a supervised manner, versus knowledge-driven (right), which accepts uncertainties (knowledge gaps) and uses concentrations required to perturb biology. In the end, this is no different to risk assessment using animal studies, where no-effect levels serve to identify a benchmark dose and the specific manifestation (hazard) is not really considered.

4.2.6 How can AOP improve these approaches?

AOP represent a form of organizing toxicological knowledge in a structured consensus process. It allows for their quality control when stored in general repositories. The obvious benefit lies

in structuring any test battery. Coverage of the AOP safeguards completeness and representation of relevant molecular initiating events and key events. For this reason, AOP, first of all, serve a targeted and adversity-driven approach. It is expected that completeness of AOP coverage improves the prediction of adversity as well as the establishment of a chemical's similarity with respect to the hazard, which may be utilized for read-across (see above) (Fig. 5). AOP are especially useful when moving from a battery of tests to an integrated testing strategy. AOP knowledge can inform how different elements need to be combined, i.e., which are crucial, redundant, augmenting, or even protective. The initiating events will inform about the point of departure of a possible concentration response curve. In combination with exposure information, this can be used to establish safety, for example, using the TTC concept (Hartung, 2017a). Via mechanistic validation (Hartung et al., 2013b), AOP can also link to animal or human reference data (Fig. 6); this concept suggests using systematic review of the literature to establish the relevance of a given mechanism covered in an *in vitro* test instead of a black-box correlation of results.

4.3 Key question #3⁴⁴: How can *in vitro* approaches that measure alterations in cellular homeostasis facilitate the evaluation of potential adverse effects for the purpose of safety/risk assessment?

4.3.1 Improving cellular assays

Major limitations of current *in vitro* toxicity assays are the artificial context of growing cells on hard plastic in conventional 2D monolayer cultures and that acute exposures and acute toxicity endpoints may not predict the subchronic or chronic responses that are common *in vivo* (Astashkina and Grainger, 2014). In ad-

⁴⁴ Group members: Bouhifd, Carmichael, Juberg, Nong, Shah

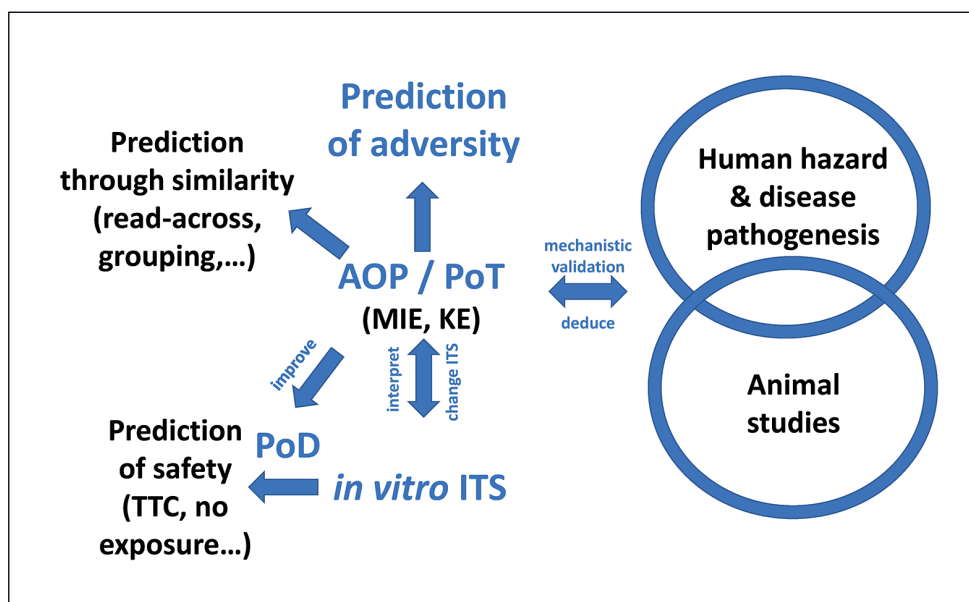


Fig. 6: Adverse outcome pathways and pathways of toxicity as the basis for risk assessment

The figure depicts how AOP / PoT can be deduced from human disease pathogenesis and animal studies, which makes them the basis of mechanistic validation. They enable the prediction of adversity, the establishment of similarity based on perturbing the same pathway, and the design of integrated testing strategies (ITS) and the identification of relevant points of departure (PoD) to establish safety.

dition, rodent *in vivo* models are not reliable in predicting toxicity and disease outcomes in humans (Heinonen, 2015). Human primary cells, especially patient-derived cells as well as immortalized human cell lines are available for toxicity testing; however, there is significant variability between donors, and primary cells have a limited lifespan *in vitro*. Induced pluripotent stem cells are emerging as an approach to obtain uniform populations of differentiated cells for drug and chemical toxicity testing (Engle et al., 2014).

As an alternative to conventional static 2D monocultures, organotypic 3D cultures have the potential to be more predictive of *in vivo* toxicity. Various commercial platforms for 3D cell cultures are available including cell micro-carrier cultures, hanging drop spheroids, rotating wall vessel bioreactors, and pre-cast native extracellular matrices or crosslinked synthetic gels (reviewed in Astashkina et al., 2014). These commercial platforms have been shown to be applicable for toxicity testing, especially for chemical-induced hepatotoxicity (LeCluyse et al., 2012; Godoy et al., 2013; Persson et al., 2014). The advantages of 3D organotypic models are better cellular differentiation, maintenance of spatial tissue organization, and recapitulation of physiological, metabolic, and mechanical functions. For example, in 3D cultures hepatocytes form bile canaliculi, maintain bile and albumin secretion, and express phase I and phase II metabolic capacity for longer times than 2D cultures (Astashkina et al., 2014). Human lung epithelial cells cultured at the air-liquid interface alone or as co-cultures with monocyte-derived macrophages and dendritic cells are better differentiated than submerged liquid cultures, although prolonged viability and reproducibility are technical challenges. Human lung organoids based on differentiation of human pluripotent stem cells may provide a new model for lung toxicity testing (Dye et al., 2015).

Advances in tissue engineering have led to the development of next-generation organotypic models for drug and toxicity

testing. Human liver co-cultures of hepatocytes, Kupffer cells, stellate cells, and endothelial cells can be maintained on a 3D scaffold for months (Kostadinova et al., 2013). Alternatively, micro-patterned co-cultures of primary hepatocytes surrounded by fibroblasts have been used for long-term culture and high-content single-cell imaging of morphological and functional assays (Trask et al., 2014). Blood flow and mechanical stimuli have been incorporated into organ-on-chip constructs to recapitulate complex tissue architecture and physiology, although these bioengineered tissues have not yet been adapted for high-throughput screening of drugs and chemicals (Astashkina et al., 2014).

High-content imaging is a quantitative technique based on automated microscopy and computerized image analysis that allows simultaneous assessment of multiple functional and morphological endpoints at a single-cell level (Tolosa et al., 2015). This technology provides a link between diagnosis of disease based on conventional light microscopy and a predictive systems biology approach to identify toxicity response pathways to chemicals, drugs, and environmental toxicants (van Vliet et al., 2014; Wink et al., 2014). Quantitative morphological alterations based on high-content imaging can be integrated with cellular and molecular endpoints and computer modeling (Tolosa et al., 2012; Kim et al., 2012). These morphological assays can be combined with live-cell gene expression studies based on molecular beacons that can monitor specific mRNA expressions over time (Alexander et al., 2011; Desai et al., 2014). These integrated morphological, functional, and molecular endpoints based on single cell assays are anticipated to be more predictive of *in vivo* responses and development of disease than population-based assays using transcriptomic or proteomic approaches (O'Brien, 2014).

A significant challenge for toxicity testing is the assessment of time-dependent changes resulting from repeated doses, inclusion of multiple exposures and stressors, and the evaluation

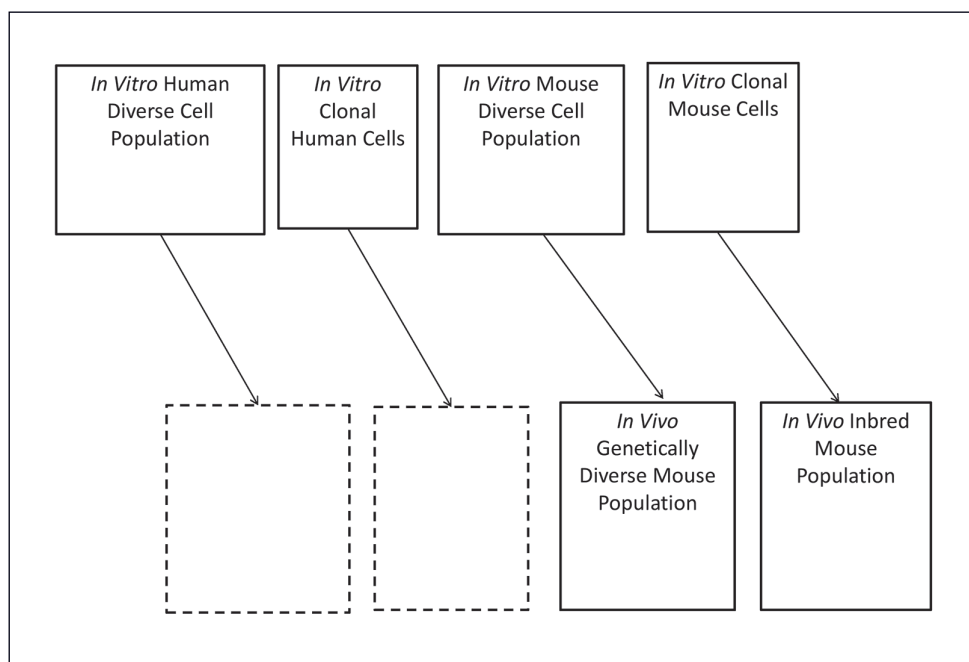


Fig. 7: A parallelogram approach to include genetic heterogeneity into a quantitative toxicity testing framework

This approach takes advantage of the extensive animal database of toxicity testing and new models of animal genetic diversity to estimate the variability in human response due to genetic heterogeneity.

of reversibility of effects. While traditional *in vivo* animal testing has addressed these challenges with subchronic and chronic models of exposure, such testing is expensive, time-intensive, and very low-throughput. *In vitro* assays have an opportunity to tackle these challenging issues through creative and thoughtful test design. One approach to developing an *in vitro* screen for assessment of the response of susceptible cells to a subsequent toxicant challenge would be to develop a standard mix comprised of chemicals known to activate key toxicity pathways (e.g., DNA damage/damage response/repair, oxidative stress response/ROS levels, heat shock response, mitochondrial permeability transition, apoptosis, and key receptor-mediated transcriptional effects). Initial exposure of the cells to this activating mixture of pathway toxicants across a concentration-response would render the cells susceptible and prime them for determining the cellular response to a second chemical or stressor in a defined, vulnerable state.

A desire for simplicity and statistical robustness has meant that toxicity testing, both *in vivo* and *in vitro*, has by-and-large been conducted with genetically and epigenetically homogeneous populations of animals and cells. However, it is clear that organisms and cells respond very differently to chemical exposures depending upon their genetic and epigenetic backgrounds (Schmidt, 2015). The importance of this heterogeneity in determining the outcome of a chemical exposure raises the question of how best to incorporate this feature of biology into the *in vitro* toxicity testing framework. One tactic to consider takes advantage of emerging animal models of genetic diversity, such as the Diversity Outbred (Schmidt, 2015) and the Collaborative Cross (Threadgill and Churchill, 2012) mouse projects, taking a parallelogram approach (Fig. 7). Taking such an approach would provide quantitative insight into the range of responses that might

be expected in making the extrapolations from cellular to organismal responses, and from genetically homogenous to heterogeneous populations of cells and organisms.

Note by the author: Section 4 essentially reflects the notes and summaries of the workshop. They have been edited slightly, a few own references added, but most has been left as summarized eight years ago, even if some newer references would be desirable.

5 The role of AI for a Human Exposome Project

Readers of this series of articles will not be surprised to see that I place many hopes in AI. Expectations are high, but many are challenged to understand its impact and most of all the permanent acceleration of its development. Over the last ten years, AI capacity has doubled every 3 months! The American decision theory and artificial intelligence researcher Eliezer Yudkowsky (1979-) coined it, “By far, the greatest danger of artificial intelligence is that people conclude too early that they understand it.” Computational approaches, especially AI, play a key role for Tox-21c 2.0 and a future Human Exposome Project:

1. *A central role of exposomics is to change to more exposure-driven toxicology, with AI enabling us to make sense of ~omics (big) data*

The central tool of exposomics is the use of ~omics of body fluids to identify patterns associated with exposure and adversity and for an exposure hypothesis. This big data effort lends itself to machine learning.

2. *Predictive toxicology through automated read-across such as read-across-based structure-activity relationships (RASAR)*

In 2018, we demonstrated that automated read-across through

machine learning for nine OECD test guidelines can outperform animal test reproducibility in classifying chemical hazards (Luechtefeld et al., 2018). Given the enormous number of untested chemicals on the market, entering the market newly, or being considered for synthesis and product development, a RASAR approach is most promising for extending our knowledge to untested substances. To illustrate the potential, in a recent study (Fu et al., 2023) we made 38,250 predictions for 4,729 food-relevant substances. The respective animal studies would take years and cost about \$250 million. A small validation exercise showed 83% accurate results ($n = 139$). Noteworthy, these are acute and topical as well as environmental hazards. However, the approach is not limited to these: Preliminary work (Luechtefeld et al., in preparation) gave reproductive toxicity 82% accurately predicted (balanced accuracy (BAC) for 1152 REACH-registered chemicals) and 75% accurate for carcinogenicity ($n = 950$). Noteworthy, this also worked to predict *in vitro* tests out of the Tox21 alliance, i.e., androgenic effects 98% accurate ($n = 8492$) and estrogenic transactivation 80% accurate ($n = 1660$). Noteworthy, the EU ONTOX project (\$20 million, 2021-2026) is currently expanding the RASAR approach to liver, kidney, and developing brain (Vinken et al., 2021, see below).

3. The computational modeling of *in vitro* tests and MPS

We have made the case earlier (Smirnova et al., 2018) that computational modeling of MPS might allow carrying out virtual experiments and then verifying them experimentally to further improve the computational model.

4. Digital pathology through image analysis

The critical role of pathology for toxicology and medicine as a whole cannot be overstated. The field is transitioning to a digital pathology (Jahn et al., 2020; Baxi et al., 2022; Dawson, 2022) with enormous advantages for standardizing, storing, comparing, and analyzing digital images of histopathological samples.

5. Prediction of all protein 3D structures

AI just predicted ALL known protein structures in work of Google-owned DeepMind and Meta (formerly Facebook): The Grand Challenge of computer modeling since the 1960's was the "protein folding problem," in which a program must predict the 3D structure of a protein based solely on an amino acid sequence. An AI program called AlphaFold by DeepMind solved this for roughly 800,000 proteins in February 2022 and then expanded to the more than 200 million proteins known to science by July 2022^{45,46}. The structures and underlying code are freely available for use. In November, researchers from Meta AI predicted the structures of roughly 617 million proteins from bacteria, viruses, and other microorganisms that have not been categorized in just two weeks (Callaway, 2022). Meta AI's database, the ESM Metagenomic Atlas, will allow scientists to quickly achieve protein structures using an appli-

cation programming interface (API). The potential for modeling drug and toxicant interactions with biological targets must still be explored.

6. Information extraction by natural language processing (NLP) of scientific literature and the grey information of the internet as well as curated databases of legacy data

The enormous progress in NLP^{47,48} (Nelson et al., 2022) is mind-blowing. For the purpose of this paper, let's simply say that NLP is close to reading scientific articles and extracting information comparable to a PhD student, but doing so millions at a time. And the computer does not forget... This means that the knowledge of the past becomes available and computable. Obviously, access to scientific literature is critical here, stressing the need for open access publishing (Hartung, 2021b).

7. Evidence integration of different evidence streams allows probabilistic risk assessment

Data mining does not always provide the answer we are seeking. The answer needs to be in the data. Cathy O'Neil wrote in her book *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, "Sometimes the job of a data scientist is to know when you don't know enough." In fact, very often AI enables us to identify data needs. For example, we should not expect to receive answers for potency or chemical interactions if we do not feed (enough of) them in in good quality. Nassim N. Taleb in *Antifragile: Things That Gain from Disorder* said it nicely, "More data means more information, but it also means more false information."

The EU project ONTOX is working toward the implementation of some of these goals. ONTOX aims to deliver NAMs for probabilistic risk assessment in toxicology (Vinken et al., 2021). To this end, physiological and toxicological data is collected and aggregated into physiological maps. These maps constitute current knowledge on physiological and toxicological perturbations caused by chemicals. The maps are meant as input for the establishment of new or to improve existing AOPs. From this, quantitative AOPs will be developed to quantitatively model compound-biology interactions. Next to the AOPs and physiological maps, ONTOX aims to develop a big data approach for performing probabilistic risk assessment (Maertens et al., 2022) based on read-across-based quantitative relationships (RASAR) (Luechtefeld et al., 2018). This artificial intelligence approach is ultimately purposed as an information toolbox for performing chemical toxicological risk assessment.

6 Relevant data versus big data – two ways of evidence integration

The traditional path of science is that beliefs, experiences, and observations lead us to hypotheses, which are explored experimentally, ultimately leading to mechanistic understanding and

⁴⁵ <https://www.technologyreview.com/2022/02/23/1044957/ai-protein-folding-deepmind/>

⁴⁶ <https://www.nature.com/articles/d41586-022-02083-2>

⁴⁷ <https://hbr.org/2022/04/the-power-of-natural-language-processing>

⁴⁸ <https://www.sorcero.com/blog/measuring-progress-natural-language-understanding>

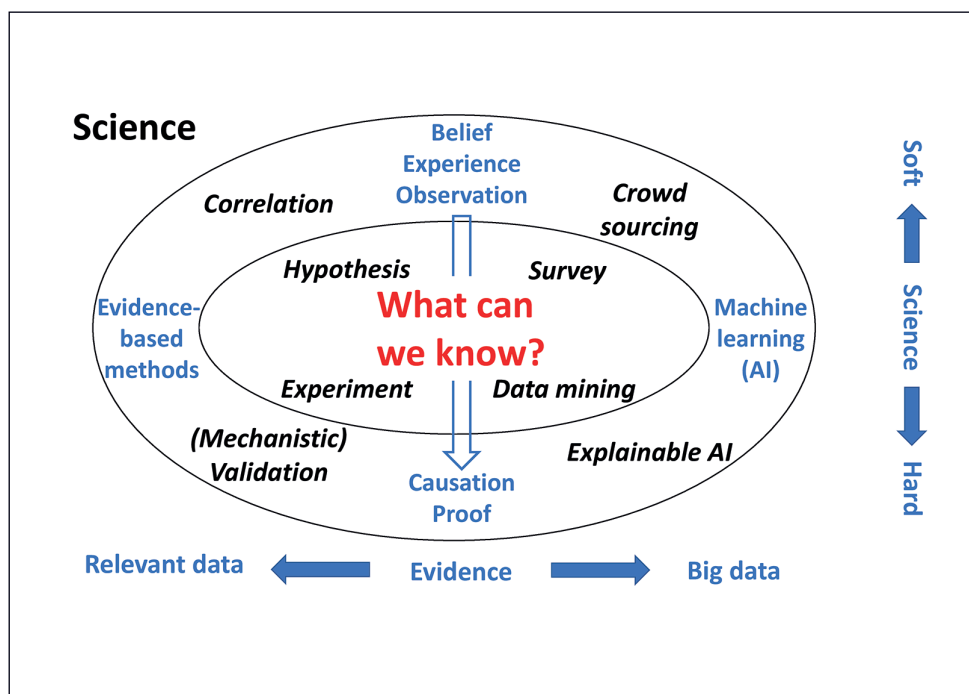


Fig. 8: Evidence-based methods and big data / machine learning approaches extend scientific methodology

The traditional approach of generating scientific knowledge through hypothesis testing has been complemented by evidence-based approaches, which reduce the available evidence through pre-defined objective criteria to allow meta-analysis of the relevant evidence. With the advent of big data technologies, machine learning, aka AI, allows automated evidence integration.

proof. This leads us from soft knowledge to hard science. Two very different variations of this process have come forward in the last decades:

1. The evidence-based methodologies with their strict, transparent, and objective handling of evidence aim to identify the relevant high-quality data and focus on this information (see Section 2.5). This approach is reductionistic in nature.
2. The big data approach lets machine learning do the data mining to integrate information (see Section 5). This approach is holistic, at least at the start.

Figure 8 shows these different approaches and some of their principal tools. To some extent, they complement each other. The problem may be the human element. In evidence-based approaches, enormous effort is spent to define admissible evidence, and the typically little remaining high-quality evidence is integrated. In machine learning, usually supervised learning is employed, i.e., the computer is told how a training set is classified (e.g., toxic or not). This is to distinguish from non-supervised machine learning, which clusters patterns and is normally less useful for toxicology, though it can help to form hypotheses. Notably, more recently, reinforcement learning as a third approach has emerged with tremendous success: In essence, the computer tries to optimize its algorithms and occasionally gets feedback on how it is performing. To illustrate this, chess computers in the past were trained based on human matches of the past. Now, Deep Mind's AlphaZero⁴⁹ starts without knowing

any chess rules and learns by trial and error. After half a day, it outperforms the best human player. Fascinatingly, as DeepMind's CEO Demis Hassabis explained⁵⁰, the computer does not play like a typical player, for example, it sacrifices many more pieces to gain other advantages, to an extent that human chess players have now started to study the computer's play. This means, we should expect that machine learning can in fact not only substitute for humans but might find ways of doing things better than we have so far.

7 Toward Tox-21 c 2.0

The 2007 NRC report on *Toxicity Testing for the 21st Century – A Vision and a Strategy* (Tox-21c) was a watershed moment for US toxicology, changing the discussion from whether to change to when and how? With knowledge in the life sciences vastly increased since 2007, technologies such as MPS and AI have emerged, which were hardly covered in the original report. Exposure-driven assessments were only covered in a parallel NRC report, but the needs for their integration into toxicology, for example, through exposomics, are increasingly evident. A key challenge is the integration of data and methods (evidence streams) in test strategies, systematic reviews, and risk assessments. Evidence-based toxicology and probabilistic risk assessments are emerging here.

⁴⁹ <https://en.wikipedia.org/wiki/AlphaZero>

⁵⁰ <https://www.youtube.com/watch?v=jocWJiztxYA&list=WL&index=74&t=974s>

In order to embrace these developments and move Tox-21c toward implementation, the Basic Research Office of the Office of the Under Secretary of Defense for Research and Engineering, OUSD(R&E), hosted Tox-21c scientists⁵¹ and agency observers⁵² for a Future Directions workshop *Advancing the Next Scientific Revolution in Toxicology* on April 28-29, 2022, at the Basic Research Innovation Collaboration Center (BRICC), in Arlington, VA, to lay out how recent developments can be embraced and Tox-21c implemented in the next decades.

The workshop was conceptualized and coordinated by the author together with Ana Navas-Acien (Columbia University), and Weihsueh Chiu (Texas A&M University). It developed to a call for Tox-21c 2.0 in the US and other countries. The workshop identified three aims: (1) precision health, (2) targeted public health interventions and environmental regulations, and (3) safer drugs and chemicals. Precision health aims for individual, personalized preventive interventions and pharmaceutical and non-pharmaceutical therapies. Targeted public health interventions and environmental regulations have to address population and spatial-temporal variability in the genome and epigenome as well as past and present exposome. Three threads were elaborated:

7.1 Exposure-driven toxicology

Tox-21c 2.0 must reflect real-world exposure. Population-scale measurements based on biobanks and ecobanks that inform on the distribution of thousands of chemical and non-chemical stressors in relevant populations are needed. In the mid-term, technologies leveraging longitudinal studies and biobanks retrospectively and prospectively, ensuring “FAIR”ness (Findability, Accessibility, Interoperability, and Reuse of digital assets), and linking the exposome with health outcomes are required. In the long-term, exposome – disease predictions and exposome-targeted prevention and treatment solutions will become part of the toxicology landscape.

7.2 Technology-enabled toxicology

MPS can play a key role to support Tox-21c 2.0 with advanced cell culture engineering and tissue-specific architecture and functionality emerging over the last 10-15 years. Technological capabilities of MPS to be further developed include a variety of models of increasing architectural complexity for different stages of drug/chemical development, representation of healthy and diseased populations by a personalized multi-verse of possible futures, platform standardization, increase in throughput, validation against human *in vivo* outcomes, incor-

poration of biosensors with near real-time outputs, and automated fabrication.

The central role of computational toxicology in Tox-21c 2.0 envisions the role of computational methods, especially machine learning, enabling the structure upon which the next scientific revolution in toxicology is based. AI has emerged as a key technology for data mining, predictive modeling, hypothesis generation, and evidence interpretation. Data-sharing following the FAIR principles⁵⁴ is key to unleash these opportunities. For the technological and biological capabilities, a need for comparable, compatible, integrable multi-omics databases, QIVIVE, and the development of *in silico* “digital twins” of *in vitro* and *in vivo* systems is anticipated.

7.3 Evidence-integrated toxicology

Evidence integration for regulatory use of Tox-21c 2.0 is key for regulatory implementation. Evidence integration across evidence streams (epidemiological, animal toxicology, *in vitro*, *in silico*, non-chemical stressors, etc.) plays a key role in translating evidence into knowledge that can inform decision-making. The vision is to conduct complex rapid/real-time evidence integration by combining advances made in data-sharing and application of AI with the transparency and rigor tenets of systematic review. There is a need for collaborative, open platform(s) to transparently collect, process, share, and interpret data, information, and knowledge on chemical and non-chemical stressors.

In summary, the workshop started a roadmap toward implementing the new concepts. Future Tox-21c 2.0 must be driven by the identification of negligible exposure (e.g., TTC) to deprioritize risk assessments and be guided by the identification of relevant exposures through exposomics. The adaptation to technical progress, especially MPS and AI, requires harmonization of reporting and quality assurance. The key challenge lies in integration of these different evidence streams. Evidence-based medicine (EBM) and probabilistic risk assessment are key. Major challenges are the validation of such new approaches and training, communication, and outreach. The workshop report will be published by OUSD(R&E) and a peer-reviewed version co-authored by the participants is in preparation.

8 Toward a Human Exposome Project

When calling for a Human Exposome Project, the vision goes beyond a large exposomics project, which is to some extent on the way as the European Human Exposome Network⁵³,

⁵¹ Tony Atala (WakeForest University), Dana Dolinoy (University of Michigan), Lauren Heine (ChemForward), Salman Khetani (University of Illinois at Chicago), Marianthi Kioumertzoglou (Columbia University), Nicole Kleinstreuer (NIEHS NTP), Koren Mann (McGill University), Uwe Marx (TissUse), Patrick McMullen (Scitovation), Gary Miller (Columbia University), Katie Paul-Friedman (US EPA), Jennifer Sass (NRDC), Kris Thayer (US EPA), Cavin Ward-Caviness (US EPA), Cheryl Walker (Baylor University), Katrina Waters (Pacific Northwest National Laboratory), Hao Zhu (Rutgers University)

⁵² OUSD(R&E), Basic Research Office: Bindu Nair, Jean Luc Cambrier, Shanni Silberberg, Daniel Osburn and Betsy Melebrink; Environmental Protection Agency: Anna Lowit and Louis Scarano; US Army Public Health Center: Mark Johnson; US Army Edgewood Chemical Biological Center: Rabi Jabbour; US Army Engineer Research and Development Center: Natalie Vinas; Department of Homeland Security, Chemical Security Analysis Center: Rachel Gooding

⁵³ <https://www.humanexposome.eu>

⁵⁴ <https://www.go-fair.org/fair-principles/>



launched in 2020, which includes 126 research groups from 24 countries in 9 large-scale projects funded with €106 million by the European Commission. In the US, exposome research has been funded mainly by NIEHS⁵⁵, but interest is expanding to NIOSH⁵⁶, NINDS⁵⁷, NIA⁵⁸, and NSF.

Exposomics has the beauty and the shortcoming of focusing on relevant exposures that can be identified from body fluids, i.e., blood and urine. We must assume, however, that a number of relevant exposures cannot be identified in this way, for example when the exposure was in the past and/or its imprint is on organs and not on blood. It will have to be shown whether mixture effects can be identified this way. Probably, large scale data generation is necessary to explore the principles of mixture toxicology. Also, epigenetic effects do not necessarily reflect on blood cells.

The title “Human Exposome Project” is intended to reflect its complementarity to the Human Genome Project, which was based largely on one technology, i.e., sequencing, but is expanding steadily to include additional aspects such as the transcriptome, gene methylation, epigenetics, etc. We will also have to clarify whether infections should be seen as part of the exposome as they are very difficult to distinguish from the microbiome, which is typically included in the exposome. This question opens another can of worms...

As a biomedical (not just toxicological) endeavor to elucidate causes of disease that are clearly beyond the typical toxicological hazards, a community must form that includes the clinical sciences, epidemiology, chemical analytics, data sciences, and others. The growing interest of NIH institutes beyond NIEHS may be taken as some appetite to go into this direction. The Human Genome Project was most remarkable also because of the international collaboration behind it, as such collaboration would be highly desirable for a Human Exposome Project.

The central role of AI in mining the exposome and expanding to the many untested and undetected chemicals, etc., prompts to call for:

$$\text{EI (exposome intelligence)} = \text{exposome} + \text{AI}$$

EI promises to make sense of the data that is generated but also to enable the fast expansion to unknown unknowns of possible associations of (mixtures of) chemicals and diverse health effects. Possible expansions to biologicals, (pathogenic) microorganisms, physical stressors like heat and radiation, social stressors, and others will have to be carefully considered. In the AI field, where capabilities are doubling every three months, it is difficult to predict what EI will look like, but here is a wish list:

1. High-throughput data generation for relevant big data that answers important safety questions, e.g., realistic exposures, potencies, mixtures, interindividual differences.

2. Combining biomonitoring and exposomics approaches.
3. Increasing combination of high-throughput and high-content analyses.
4. Improved data mining, e.g., applying signal detection theory (McNicol, 2016) or Dempster-Shafer theory⁵⁹, a combination of evidence obtained from multiple sources, and the modeling of conflict between them.
5. Explainable AI, i.e., the emerging use of tools and frameworks to help understand and interpret predictions made by machine learning models. Currently, these come with lower accuracy of prediction, but this will be a matter of time only. This is a type of principal component analysis or sensitivity analysis for AI, i.e., showing what informed a prediction most.
6. Adaptation to technical AI progress. Currently, this means the use of large, existing models to integrate new data only and taking advantage of the superior architecture and transfer learning from past data analyses, NLP (transformers) to mine the knowledge of the past, image analysis, reinforcement learning, attention layers, distributed agents, federated systems, etc. For those not familiar with these terms, it's time to learn about them (or wait for an upcoming review on AI in toxicology). Just as AlphaZero playing binary games led to the development of AlphaFold, which can predict 3D structures of proteins, muZero⁶⁰, which excels at all types of games without knowing their rules at the start, could be developed to master the game of solving our toxicological riddles.
7. Concepts of AI-based latent class analysis (LCA). LCA is a form of unsupervised learning: “*Latent class analysis is more statistically principled than either of the standard nonhierarchical and hierarchical clustering techniques, in that the statistical inference is built from a probability model assumed to hold in the data.*” (Bunge and Judson, 2005). This means that no gold standard is employed (i.e., we know what is toxic and what is not), but we assume the correct result is held by all different measurements together. This is used, for example, when introducing new diagnostic tools, where no accepted method can serve for comparison.

Tox-21c implementation and the preparation of Tox-21c 2.0 were driven by a conceptual discussion, which took place in part in this journal over the last 16 years (Tab. 2). Such discussions will need to be continued toward a Human Exposome Project. Is it possible? I think so. Arthur C. Clarke (1917–2008) formulated three adages that are known as Clarke’s three laws. One of them is, “*When a distinguished but elderly scientist states that something is possible, he is almost certainly right. When he states that something is impossible, he is very probably wrong.*” Take this from someone who is starting to fit at least the “elderly” category.

⁵⁵ <https://www.niehs.nih.gov/research/supported/exposure/bio/index.cfm>

⁵⁶ <https://www.cdc.gov/niosh/topics/exposome/default.html>

⁵⁷ <https://www.ninds.nih.gov/current-research/research-funded-ninds/translational-research/onetox-neural-exposome-and-toxicology-programs/onetox-neural-exposome>

⁵⁸ <https://www.nia.nih.gov/research/milestones/epidemiology-population-studies/milestone-1-b>

⁵⁹ <https://www.osti.gov/servlets/purl/800792-s9WKeP/native/>

⁶⁰ <https://en.wikipedia.org/wiki/MuZero>

Tab. 2: Conceptual articles in 16 years of the Food for Thought ... series and Transatlantic Think Tank for Toxicology (t⁴) for transforming toxicology

70 articles were published in the Food for Thought ... series in *ALTEX* since 2007. The articles involving the author are labeled in **blue**. A few additional articles similar in style to the previous ones by the author are labeled in **red**. In the same time frame, 42 t⁴ (workshop) reports were published and are labeled in **green**. Only partially related to toxicology: [Kang et al., 2021](#) (use of MPS for COVID research); [Hartung, 2015, 2021a](#); [Hasiwa et al., 2013](#) (pyrogen testing); [Hartung, 2017d](#) (on the Food for Thought ... series itself). Tbd, to be done

	Current toxicology	Tox-21c	Tox-21c 2.0	Human Exposome Project
Goal	3Rs Shortcomings of <i>in vivo</i>	Mechanistic toxicology and IVIVE	Exposure-driven, technology-enabled, evidence-integrated	Exposure leading to disease
Tools	3Rs a) Reduce Steger-Hartmann et al., 2020 b) Replace Andersen et al., 2019 ; Jacobs et al., 2018 ; de Souza Santos et al., 2017 ; Gordon et al., 2015 ; Patlewicz et al., 2014 ; Juberg et al., 2014 ; Leist et al., 2014 ; Smirnova et al., 2014 ; Hartung and Corsini, 2013 ; Parks Saldutti et al., 2013 ; Rossini and Hartung, 2012 ; Basketter et al., 2012 ; Kuegler et al., 2010 ; Hartung and Hoffmann, 2009 ; Leist et al., 2008a ; Leist et al., 2008b ; Hartung, 2007b c) Refine Herrmann et al., 2019 ; Zurlo and Hutchinson, 2014 Shortcomings of <i>in vivo</i> Eisenman, 2016 ; Hartung, 2013 ; Paparella et al., 2013 ; Hartung, 2008b	a) Mechanistic Hoffmann et al., 2022b ; Thomas et al., 2018 ; Leist et al., 2017 ; Smirnova et al., 2015 ; Kleensang et al., 2014 ; Ramirez et al., 2013 ; Hartung et al., 2012 ; Blaauboer et al., 2012 ; Andersen et al., 2011 ; Hartung and McBride, 2011 ; Hartung and Leist, 2008 b) IVIVE Hartung, 2018c ; Tsaioun et al., 2016	Exposure-driven Sillé et al., 2020 ; Hartung, 2017b ; Escher et al., 2017 Technology enabled a) MPS Roth et al., 2019 ; Marx et al., 2016, 2020 ; Smirnova et al., 2018 ; Alépée et al., 2014 ; Andersen et al., 2014 b) AI Briggs et al., 2021 ; Madia et al., 2020 ; Luechtefeld and Hartung, 2017 ; Hartung, 2018b ; Hartung, 2016a c) High-content imaging van Vliet et al., 2014 Evidence integration Krewski et al., 2022 ; Maertens et al., 2022 ; Farhat et al., 2022 ; Caloni et al., 2022 ; Hoffmann et al., 2022a,b ; Hartung, 2017c, 2021b ; Hoffmann et al., 2017 ; Linkov et al., 2015 ; Hartung, 2009b	Systems biology and toxicology Hartung et al., 2017 Big data from biomonitoring and exposomics Tbd Exposome intelligence Tbd
Applications	Food Hartung, 2018a ; Hartung and Koëter, 2008 Drugs Cassotta et al., 2022 ; Beilmann et al., 2019 Chemicals Taylor et al., 2014 ; Rovida, 2010 ; Hartung, 2010c Cosmetics Hartung, 2008a ; Hartung et al., 2011 Nanoparticles Hartung, 2010b ; Silbergeld et al., 2011 Biocides Ferrario and Rabbit, 2012 Medical devices Kerecman Myers et al., 2017 Countermeasures Hartung and Zurlo, 2012 Marine biotoxins Daneshian et al., 2013	Chemicals Carmichael et al., 2022 E-cigarettes Hartung, 2016b Drugs and food Hartung, 2017c ; Rovida et al., 2015b		One-substance-one-assessment Tbd
Quality and validation	Validation Bal-Price et al., 2018 ; Wilcox and Goldberg, 2011 ; Hartung, 2007a Reporting standards Leist et al., 2010 Read-across Ball et al., 2016 ; Zhu et al., 2016 Terminology Ferrario et al., 2014	Validation Hartung et al., 2013b Good Cell Culture Practice Pamies et al., 2017, 2018, 2020, 2022 ; Pamies and Hartung, 2017 Reporting standards Hartung et al., 2019 ; Krebs et al., 2019 Ontology Hardy et al., 2012	Metabolomics Lippa et al., 2022 ; Evans et al., 2020 ; Beger et al., 2019 ; Bouhifd et al., 2015b Exposomics Tbd	



	Current toxicology	Tox-21c	Tox-21c 2.0	Human Exposome Project
Goal	3Rs Shortcomings of <i>in vivo</i>	Mechanistic toxicology and IVIVE	Exposure-driven, technology- enabled, evidence-integrated	Exposure leading to disease
Strategic	Animal use Busquet and Hartung, 2020 ; Daneshian et al., 2015 ; Rovida et al., 2011 ; Rovida and Hartung, 2009 Policy Knight et al., 2021 ; Locke and Myers, 2011 ; Hartung, 2010d ; Dog use Hasiwa et al., 2011 ; Zurlo et al., 2011 Read-across Rovida et al., 2020	Economic Meigs et al., 2018 ; Bottini and Hartung, 2009 Globalization Bottini et al., 2007 Green toxicology Maertens et al., 2022, 2014 ; Maertens and Hartung, 2018 Integrated testing Rovida et al., 2015a ; Hartung et al., 2013a Education Daneshian et al., 2011 ; Hartung et al., 2009	Societal von Aulock et al., 2022 Scientific Hartung and Tsatsakis, 2021 ; Busquet and Hartung, 2017 Publishing Whaley et al., 2021	Vision Hartung, 2023 (this article)

References

- Abedini, J., Cook, B., Bell, S. et al. (2021). Application of new approach methodologies: ICE tools to support chemical evaluations. *Comput Toxicol* 20, 100184. doi:10.1016/j.comtox.2021.100184
- Aguilera, O., Fernandez, A. F., Munoz, A. et al. (2010). Epigenetics and environment: A complex relationship. *J Appl Physiol* 109, 243-251. doi:10.1152/japplphysiol.00068.2010
- 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
- Alexander, J. C., Pandit, A., Bao, G. et al. (2011). Monitoring mRNA in living cells in a 3D in vitro model using TAT-peptide linked molecular beacons. *Lab Chip* 11, 3908-3914. doi:10.1039/c1lc20447e
- Andersen, M. E., Clewell, H. J. III, Carmichael, P. L. et al. (2011). Can case study approaches speed implementation of the NRC report: Toxicity testing in the 21st century: A vision and a strategy? *ALTEX* 28, 175-182. doi:10.14573/altex.2011.3.175
- Andersen, M. E., Betts, K., Dragan, Y. et al. (2014). Developing microphysiological systems for use as regulatory tools – Challenges and opportunities. *ALTEX* 31, 364-367. doi:10.14573/altex.1405151
- Andersen, M. E., McMullen, P. D., Phillips, M. B. et al. (2019). Developing context appropriate toxicity testing approaches using new alternative methods (NAMs). *ALTEX* 36, 523-534. doi:10.14573/altex.1906261
- 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
- Aschner, M., Ceccatelli, S., Daneshian, M. et al. (2017). Reference compounds for alternative test methods to indicate developmental neurotoxicity (DNT) potential of chemicals: Example lists and criteria for their selection and use. *ALTEX* 34, 49-74. doi:10.14573/altex.1604201
- Aschner, M., Paoliello, M. M. B., Tsatsakis, A. et al. (2021). Social injustice in environmental health: A call for fortitude. *Environ Res* 194, 110675. doi:10.1016/j.envres.2020.110675
- Astashkina, A. and Grainger D.W. (2014). Critical analysis of 3-D organoid in vitro cell culture models for high-throughput drug candidate toxicity assessments. *Adv Drug Deliv Rev* 69-70, 1-18. doi:10.1016/j.addr.2014.02.008
- Ball, N., Cronin, M. T. D., Shen, J. et al. (2016). Toward good read-across practice (GRAP) guidance. *ALTEX* 33, 149-166. doi:10.14573/altex.1601251
- Bal-Price A., Hogberg, H. T., Crofton, K. M. et al. (2018). Recommendation on test readiness criteria for new approach methods in toxicology: Exemplified for developmental neurotoxicity. *ALTEX* 35, 306-352. doi:10.14573/altex.1712081
- Basketter, D. A., Clewell, H., Kimber, I. et al. (2012). A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing. *ALTEX* 29, 3-89. doi:10.14573/altex.2012.1.003
- Baxi, V., Edwards, R., Montalto, M. et al. (2022). Digital pathology and artificial intelligence in translational medicine and clinical practice. *Mod Pathol* 35, 23-32. doi:10.1038/s41379-021-00919-2
- Beger, R. D., Dunn, W. B., Bandukwala, A. et al. (2019). Towards quality assurance and quality control in untargeted metabolomics studies. *Metabolomics* 15, 4. doi:10.1007/s11306-018-1460-7
- Beilmann, M., Boonen, H., Czich, A. et al. (2019). Optimizing drug discovery by investigative toxicology: Current and future trends. *ALTEX* 36, 3-17. doi:10.14573/altex.1808181
- Beinart, R., Zhang, Y., Lima, J. A. et al. (2014). The QT interval is associated with incident cardiovascular events: The MESA study. *J Am Coll Cardiol* 64, 2111-2119. doi:10.1016/j.jacc.2014.08.039
- Bhattacharya, S., Zhang, Q., Carmichael, P. L. et al. (2011). Toxicity testing in the 21 century: Defining new risk assessment approaches based on perturbation of intracellular toxicity pathways. *PLoS One* 6, e20887. doi:10.1371/journal.pone.0020887
- Blakey, D., Galloway, S. M., Kirkland, D. J. et al. (2008). Regulatory aspects of genotoxicity testing: From hazard identification to risk assessment. *Mutat Res* 657, 84-90. doi:10.1016/j.mrgentox.2008.09.004

- 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
- Boekelheide, K. and Campion, S. N. (2010). Toxicity testing in the 21st century: Using the new toxicity testing paradigm to create a taxonomy of adverse effects. *Toxicol Sci* 114, 20-24. doi:10.1093/toxsci/kfp307
- Bottini, A. A., Amcoff, P. and Hartung, T. (2007). Food for thought... on globalization of alternative methods. *ALTEX* 24, 255-261. doi:10.14573/altex.2007.4.255
- Bottini, A. A. and Hartung, T. (2009). Food for thought ... on the economics of animal testing. *ALTEX* 26, 3-16. doi:10.14573/altex.2009.1.3
- Bouhifd, M., Andersen, M. E., Baghdikian, C. et al. (2015a). The human toxome project. *ALTEX* 32, 112-124. doi:10.14573/altex.1502091
- Bouhifd, M., Beger, R., Flynn, T. et al. (2015b). Quality assurance of metabolomics. *ALTEX* 32, 319-326. doi:10.14573/altex.1509161
- Briggs, K., Bosc, N., Camara, T. et al. (2021). Guidelines for FAIR sharing of preclinical safety and off-target pharmacology data. *ALTEX* 38, 187-197. doi:10.14573/altex.2011181
- Browne, P., Delrue, N. and Gourmelon, A. (2019). Regulatory use and acceptance of alternative methods for chemical hazard identification. *Curr Op Toxicol* 15, 18-25. doi:10.1016/j.cotox.2019.02.003
- Bunge, J. A., and Judson, D. H. (2005). Data Mining. In K. Kempf-Leonard (ed.), *Encyclopedia of Social Measurement* (617-624). Elsevier. doi:10.1016/B0-12-369398-5/00159-6
- Busquet, F. and Hartung, T. (2017). The need for strategic development of safety sciences. *ALTEX* 34, 3-21. doi:10.14573/altex.1701031
- Busquet, F., Kleensang, A., Rovida, C. et al. (2020). New European Union statistics on laboratory animal use – What really counts! *ALTEX* 37, 167-186. doi:10.14573/altex.2003241
- Callaway, E. (2022). AlphaFold's new rival? Meta AI predicts shape of 600 million proteins. *Nature* 611, 211-212. doi:10.1038/d41586-022-03539-1
- Caloni, F., De Angelis, I. and Hartung, T. (2022). Replacement of animal testing by integrated approaches to testing and assessment (IATA): A call for in vivitrosi. *Arch Toxicol* 96, 1935-1950. doi:10.1007/s00204-022-03299-x
- Carmichael, P. L., Baltazar, M. T., Cable, S. et al. (2022). Ready for regulatory use: NAMs and NGRA for chemical safety assurance. *ALTEX* 39, 359-366. doi:10.14573/altex.2204281
- Cassotta, M., Geerts, H., Harbom, L. et al. (2022). The future of Parkinson's disease research: A new paradigm of human-specific investigation is necessary... and possible. *ALTEX* 39, 694-709. doi:10.14573/altex.2203161
- Chesnut, M., Yamada, T., Adams, T. et al. (2018). Regulatory acceptance of read-across: Report from an international satellite meeting at the 56th annual meeting of the society of toxicology. *ALTEX* 35, 413-419. doi:10.14573/altex.1805081
- Coecke, S., Balls, M., Bowe, G. et al. (2005). Guidance on good cell culture practice – A report of the second ECVAM task force on good cell culture practice. *Altern Lab Anim* 33, 261-287. doi:10.1177/026119290503300313
- Cooper-Hannan, R., Harbell, J. W., Coecke, S. et al. (1999). The principles of good laboratory practice: Application to in vitro toxicology studies – The report and recommendations of ECVAM workshop 37. *Altern Lab Anim* 27, 539-577. doi:10.1177/026119299902700410
- Corvi, R., Albertini, S., Hartung, T. et al. (2008). ECVAM retrospective validation of in vitro micronucleus test (MNT). *Mutagenesis* 23, 271-283. doi:10.1093/mutage/gen010
- 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
- Daneshian, M., Botana, L. M., Dechraoui Bottein, M. Y. et al. (2013). A roadmap for hazard monitoring and risk assessment of marine biotoxins on the basis of chemical and biological test systems. *ALTEX* 30, 487-545. doi:10.14573/altex.2013.4.487
- Daneshian, M., Busquet, F., Hartung, T. et al. (2015). Animal use for science in Europe. *ALTEX* 32, 261-274. doi:10.14573/altex.1509081
- Dawson, H. (2022). Digital pathology – Rising to the challenge. *Front Med* 9, 888896. doi:10.3389/fmed.2022.888896
- de Souza Santos, R., Frank, A. P., Palmer, B. F. et al. (2018). Sex and media: Considerations for cell culture studies. *ALTEX* 35, 435-440. doi:10.14573/altex.1806151
- de Vries, R. B. M., Angrish, M., Browne, P. et al. (2021). Applying evidence-based methods to the development and use of adverse outcome pathways. *ALTEX* 38, 336-347. doi:10.14573/altex.2101211
- Desai, H. V., Voruganti, I. S., Jayasuriya, C. et al. (2014). Live-cell, temporal gene expression analysis of osteogenic differentiation in adipose-derived stem cells. *Tissue Eng Part A* 20, 899-907. doi:10.1089/ten.tea.2013.0761
- Dolinoy, D.C., Huang, D. and Jirtle, R. L. (2007). Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. *Proc Natl Acad Sci U S A* 104, 13056-13061. doi:10.1073/pnas.0703739104
- Dye, B. R., Hill, D. R., Ferguson, M. A. et al. (2015). In vitro generation of human pluripotent stem cell derived lung organoids. *Elife* 4, e05098. doi:10.7554/eLife.05098
- EFSA – European Food Safety Authority and EBTC – Evidence-Based Toxicology Collaboration (2018). EFSA scientific colloquium 23: Evidence integration in risk assessment: The science of combining apples and oranges. *EFSA Supporting Publications* 15, 1396E. doi:10.2903/sp.efsa.2018.EN-1396
- EFSA, Tarazona, J., Kass, G. et al. (2022). New approach methodologies. *EFSA Supporting Publications* 19, e200502. doi:10.2903/sp.efsa.2022.e200502
- Eisenman, S. F. (2016). Criticizing animal experimentation, at my peril. *ALTEX* 33, 3-12. doi:10.14573/altex.1509282
- Ellison, C. A., Api, A. M., Becker, R. A. et al. (2020). Internal threshold of toxicological concern (iTTC): Where we are today and what is possible in the near future. *Front Toxicol* 2, 621541. doi:10.3389/ftox.2020.621541
- Engle, S. J. and Vincent, F. (2014). Small molecule screening in human induced pluripotent stem cell-derived terminal cell types. *J Biol Chem* 289, 4562-4570. doi:10.1074/jbc.r113.529156



- Escher, B. I., Hackermüller, J., Polte, T. et al. (2017). From the exposome to mechanistic understanding of chemical-induced adverse effects. *Environ Int* 99, 97-106. doi:10.1016/j.envint.2016.11.029
- Escher, S. E., Partosch, F., Konzok, S. et al. (2022). Development of a roadmap for action on new approach methodologies in risk assessment. *EFSA Supporting Publication* 19, 7341E. doi:10.2903/sp.efsa.2022.EN-7341
- Evans, A. M., O'Donovan, C., Playdon, M. et al. (2020). Dissemination and analysis of the quality assurance (QA) and quality control (QC) practices of LC-MS based untargeted metabolomics practitioners. *Metabolomics* 16, 113. doi:10.1007/s11306-020-01728-5
- Farhat, N., Tsaioun, K., Saunders-Hastings, P. et al. (2022). Systematic review in evidence-based risk assessment. *ALTEX* 39, 463-479. doi:10.14573/altex.2004111
- Ferrario, D. and Rabbit, R. R. (2012). Analysis of the proposed EU regulation concerning biocide products and its opportunities for alternative approaches and a toxicology for the 21st century. *ALTEX* 29, 157-172. doi:10.14573/altex.2012.2.157
- 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
- Fowle, J. J. R. III, Curren, R. D., Hartung, T. et al. (2017). Twenty-first century in vitro toxicology testing methods and the assessment of e-cigarettes. *Appl In Vitro Toxicol* 3, 3-9. doi:10.1089/aivt.2017.29011.rtl
- Fu, Y., Luechtefeld, T., Karmaus, A. et al. (2023). Chapter 39 – The use of artificial intelligence and big data for the safety evaluation of US food-relevant chemicals. In M. E. Knowles, L. E. Anelich, A. R. Boobis et al., *Present Knowledge in Food Safety* (575-589). Academic Press. doi:10.1016/B978-0-12-819470-6.00061-5
- Godoy, P., Hewitt, N. J., Albrecht, U. et al. (2013). Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. *Arch Toxicol* 87, 1315-1530. doi:10.1007/s00204-013-1078-5
- Gordon, S., Daneshian, M., Bouwstra, J. et al. (2015). Non-animal models of epithelial barriers (skin, intestine and lung) in research, industrial applications and regulatory toxicology. *ALTEX* 32, 327-378. doi:10.14573/altex.1510051
- Grandjean, P. and Bellanger, M. (2017). Calculation of the disease burden associated with environmental chemical exposures: Application of toxicological information in health economic estimation. *Environ Health* 16, 123. doi:10.1186/s12940-017-0340-3
- Grønbaek, M. (2009). The positive and negative health effects of alcohol- and the public health implications. *J Intern Med* 265, 407-420. doi:10.1111/j.1365-2796.2009.02082.x
- Gstraunthaler, G. and Hartung, T. (1999). Bologna declaration toward good cell culture practice. *Altern Lab Anim* 27, 206.
- Guillen, J. (2012). FELASA guidelines and recommendations. *J Am Assoc Lab Anim Sci* 51, 311-321.
- Hardy, B., Apic, G., Carthew, P. et al. (2012). Food for thought ... a toxicology ontology roadmap. *ALTEX* 29, 129-137. doi:10.14573/altex.2012.2.129
- Hartung, T., Balls, M., Bardouille, C. et al. (2002). Good cell culture practice. ECVAM good cell culture practice task force report 1. *Altern Lab Anim* 30, 407-414. doi:10.1177/026119290203000404
- 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. doi:10.1177/026119290403200503
- Hartung, T. (2007a). Food for thought ... on validation. *ALTEX* 24, 67-72. doi:10.14573/altex.2007.2.67
- Hartung, T. (2007b). Food for thought ... on cell culture. *ALTEX* 24, 143-147. doi:10.14573/altex.2007.3.143
- Hartung, T. (2008a). Food for thought ... on alternative methods for cosmetics safety testing. *ALTEX* 25, 147-162. doi:10.14573/altex.2008.3.147
- Hartung, T. (2008b). Food for thought ... on animal tests. *ALTEX* 25, 3-16. doi:10.14573/altex.2008.1.3
- Hartung, T. and Leist, M. (2008). Food for thought ... on the evolution of toxicology and the phasing out of animal testing. *ALTEX* 25, 91-102. doi:10.14573/altex.2008.2.91
- Hartung, T. and Koeter, H. (2008). Food for thought ... on food safety testing. *ALTEX* 25, 259-264. doi:10.14573/altex.2008.4.259
- Hartung, T. (2009a). A toxicology for the 21st century: Mapping the road ahead. *Tox Sci* 109, 18-23. doi:10.1093/toxsci/kfp059
- Hartung, T. (2009b). Food for thought ... on evidence-based toxicology. *ALTEX* 26, 75-82. doi:10.14573/altex.2009.2.75
- Hartung, T. and Hoffmann, S. (2009). Food for thought ... on in silico methods in toxicology. *ALTEX* 26, 155-166. doi:10.14573/altex.2009.3.155
- Hartung, T., Blaauboer, B. and Leist, M. (2009). Food for thought ... on education in alternative methods in toxicology. *ALTEX* 26, 255-263. doi:10.14573/altex.2009.4.255
- Hartung, T. (2010a). Evidence based-toxicology – The toolbox of validation for the 21st century? *ALTEX* 27, 241-251. doi:10.14573/altex.2010.4.253
- Hartung, T. (2010b). Food for thought ... on alternative methods for nanoparticle safety testing. *ALTEX* 27, 87-95. doi:10.14573/altex.2010.2.87
- Hartung, T. (2010c). Food for thought ... on alternative methods for chemical safety testing. *ALTEX* 27, 3-14. doi:10.14573/altex.2010.1.3
- Hartung, T. (2010d). Comparative analysis of the revised Directive 2010/63/EU for the protection of laboratory animals with its predecessor 86/609/EEC. *ALTEX* 27, 285-303. doi:10.14573/altex.2010.4.285
- 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., Blaauboer, B. J., Bosgra, S. et al. (2011). An expert consortium review of the EC-commissioned report “Alternative (non-animal) methods for cosmetics testing: Current status and future prospects – 2010”. *ALTEX* 28, 183-209. doi:10.14573/altex.2011.3.183
- Hartung, T. and Zurlo, J. (2012). Alternative approaches for medical countermeasures to biological and chemical terrorism and warfare. *ALTEX* 29, 251-260. doi:10.14573/altex.2012.3.251
- 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. (2013). Look back in anger – What clinical studies tell us about preclinical work. *ALTEX* 30, 275-291. doi:10.14573/altex.2013.3.275
- Hartung, T. and Corsini, E. (2013). Immunotoxicology: Challenges in the 21st century and in vitro opportunities. *ALTEX* 30, 411-426. doi:10.14573/altex.2013.4.411
- Hartung, T., Luechtefeld, T., Maertens, A. et al. (2013a). Integrated testing strategies for safety assessments. *ALTEX* 30, 3-18. doi:10.14573/altex.2013.1.003
- Hartung, T., Stephens, M. and Hoffmann, S. (2013b). Mechanistic validation. *ALTEX* 30, 119-130. doi:10.14573/altex.2013.2.119
- Hartung, T. (2015). The human whole blood pyrogen test – Lessons learned in twenty years. *ALTEX* 32, 79-100. doi:10.14573/altex.1503241
- Hartung, T. (2016a). Making big sense from big data in toxicology by read-across. *ALTEX* 33, 83-93. doi:10.14573/altex.1603091
- Hartung, T. (2016b). E-Cigarettes and the need and opportunities for alternatives to animal testing. *ALTEX* 33, 211-224. doi:10.14573/altex.1606291
- Hartung, T. (2016c). Can we wait for e-cigarette trials? *Chemistry World*. <http://www.rsc.org/chemistryworld/2016/04/electronic-cigarettes-trials-vaping-safety-data>
- Hartung, T. (2016d). E-cigarettes – The ugly duckling of public health? *Sci Am* 315, 9.
- Hartung, T. (2017a). Thresholds of toxicological concern – Setting a threshold for testing where there is little concern. *ALTEX* 34, 331-351. doi:10.14573/altex.1707011
- Hartung, T. (2017b). Opinion versus evidence for the need to move away from animal testing. *ALTEX* 34, 193-200. doi:10.14573/altex.1703291
- Hartung, T. (2017c). 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. (2017d). Food for thought ... the first ten years. *ALTEX* 34, 187-192. doi:10.14573/altex.1703311
- Hartung, T., FitzGerald, R., Jennings, P. et al. (2017). Systems toxicology – Real world applications and opportunities. *Chem Res Toxicol* 30, 870-882. doi:10.1021/acs.chemrestox.7b00003
- Hartung, T. (2018a). Rebooting the generally recognized as safe (GRAS) approach for food additive safety in the US. *ALTEX* 35, 3-25. doi:10.14573/altex.1712181
- Hartung, T. (2018b). Making big sense from big data. *Front Big Data* 1, 5. doi:10.3389/fdata.2018.00005
- Hartung, T. (2018c). Perspectives on in vitro to in vivo extrapolations. *J Appl In Vitro Toxicol* 4, 305-316. doi:10.1089/aivt.2016.0026
- Hartung, T., de Vries, R., Hoffmann, S. et al. (2019). Toward good in vitro reporting standards. *ALTEX* 36, 3-17. doi:10.14573/altex.1812191
- Hartung, T. (2021a). Pyrogen testing revisited on occasion of the 25th anniversary of the whole blood test. *ALTEX* 38, 3-19. doi:10.14573/altex.2101051
- Hartung, T. (2021b). Evidence integration in the era of information flooding – The advent of the comprehensive review. *Front Public Health* 9, 763828. doi:10.3389/fpubh.2021.763828
- Hartung, T. and Tsatsakis, A. M. (2021). The state of the scientific revolution in toxicology. *ALTEX* 38, 379-386. doi:10.14573/altex.2106101
- Hasiwa, N., Bailey, J., Clausen, P. et al. (2011). Critical evaluation of the use of dogs in biomedical research and testing in Europe. *ALTEX* 28, 326-340. doi:10.14573/altex.2011.4.326
- Hasiwa, N., Daneshian, M., Bruegger, P. et al. (2013). Evidence for the detection of non-endotoxin pyrogens by the whole blood monocyte activation test. *ALTEX* 30, 169-208. doi:10.14573/altex.2013.2.169
- Heinonen, T. (2015). Better science with human cell-based organ and tissue models. *Altern Lab Anim* 43, 29-38. doi:10.1177/026119291504300107
- Helman, G., Shah, I., Williams, A. J. et al. (2019). Generalized read-across (GenRA): A workflow implemented into the EPA CompTox chemicals dashboard. *ALTEX* 36, 462-465. doi:10.14573/altex.1811292
- Hermann, K., Pistollato, F. and Stephens, M. (2019). Beyond the 3Rs: Expanding the use of human-relevant replacement methods in biomedical research. *ALTEX* 36, 343-352. doi:10.14573/altex.1907031
- Herzler, M., Marx-Stoelting, P., Pirow, R. et al. (2021). The “EU chemicals strategy for sustainability” questions regulatory toxicology as we know it: Is it all rooted in sound scientific evidence? *Arch Toxicol* 95, 2589-2601. doi:10.1007/s00204-021-03091-3
- Hoffmann, S. and Hartung, T. (2006a). Designing validation studies more efficiently according to the modular approach: Retrospective analysis of the EPISKIN test for skin corrosion. *Altern Lab Anim* 34, 177-191. doi:10.1177/026119290603400209
- Hoffmann, S. and Hartung, T. (2006b). Towards an evidence-based toxicology. *Hum Exp Toxicol* 25, 497-513. doi:10.1191/0960327106het6480a
- Hoffmann, S., Edler, L., Gardner, I. et al. (2008). Points of reference in validation – The report and recommendations of ECVAM workshop. *Altern Lab Anim* 36, 343-352. doi:10.1177/026119290803600311
- Hoffmann, S., de Vries, R. B. M., Stephens, M. L. et al. (2017). A primer on systematic reviews in toxicology. *Arch Toxicol* 91, 2551-2575. doi:10.1007/s00204-017-1980-3
- Hoffmann, S., Margliani, B., Akgün-Ölmez, S. G. et al. (2021). A systematic review to compare chemical hazard predictions of the zebrafish embryo test with mammalian prenatal developmental toxicity. *Toxicol Sci* 183, 14-35. doi:10.1093/toxsci/kfab072
- Hoffmann, S., Whaley, P. and Tsaïoun, K. (2022a). How evidence-based methodologies can help identify and reduce uncertainty in chemical risk assessment. *ALTEX* 39, 175-182. doi:10.14573/altex.2201131
- Hoffmann, S., Aiassa, E., Angrish, M. et al. (2022b). Application of evidence-based methods to construct mechanism-driven chemical assessment frameworks. *ALTEX* 39, 499-518. doi:10.14573/altex.2202141
- Jacobs, M. N., Versteegen, R. J., Treasure, C. et al. (2019). Addressing potential ethical issues regarding the supply of human-derived products or reagents in in vitro OECD test guidelines. *ALTEX* 36, 163-176. doi:10.14573/altex.1901281
- Jahn, S. W., Plass, M. and Moïnfar, F. (2020). Digital pathology:



- Advantages, limitations and emerging perspectives. *J Clin Med* 18, 3697. doi:10.3390/jcm9113697
- Jones, P. A. and Baylin, S. B. (2007). The epigenomics of cancer. *Cell* 128, 683-692. doi:10.1016/j.cell.2007.01.029
- Jonker, D. M., Kenna, L. A., Leishman, D. et al. (2005). A pharmacokinetic-pharmacodynamic model for the quantitative prediction of dofetilide clinical QT prolongation from human ether-a-go-go-related gene current inhibition data. *Clin Pharmacol Ther* 77, 572-582. doi:10.1016/j.clpt.2005.02.004
- Juberg, D. R., Borghoff, S. J., Becker, R. A. et al. (2014). Lessons learned, challenges, and opportunities: The U.S. endocrine disruptor screening program. *ALTEX* 31, 63-78. doi:10.14573/altex.1309171
- Kang, I., Smirnova, L., Kuhn, J. H. et al. (2021). COVID-19 – Prime time for microphysiological systems, as illustrated for the brain. *ALTEX* 38, 535-549. doi:10.14573/altex.2110131
- Kerecman Myers, D., Goldberg, A. M., Poth, A. et al. (2017). From in vivo to in vitro: The medical device testing paradigm shift. *ALTEX* 34, 479-500. doi:10.14573/altex.1608081
- Kim, J. A., Han, E., Eun, C. J. et al. (2012). Real-time concurrent monitoring of apoptosis, cytosolic calcium, and mitochondria permeability transition for hypermulticolor high-content screening of drug-induced mitochondrial dysfunction-mediated hepatotoxicity. *Toxicol Lett* 214, 175-181. doi:10.1016/j.toxlet.2012.08.027
- Kirkland, D., Pfuhler, S., Tweats, D. et al. (2007). How to reduce false positive results when undertaking in vitro genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of an ECVAM workshop. *Mutat Res* 628, 31-55. doi:10.1016/j.mrgentox.2006.11.008
- Kleensang, A., Maertens, A., Rosenberg, M. et al. (2014). Pathways of toxicity. *ALTEX* 31, 53-61. doi:10.14573/altex.1309261
- Kleinstreuer, N. C., Hoffmann, S., Alépée, N. et al. (2018). Non-animal methods to predict skin sensitization (II): An assessment of defined approaches*. *Crit Rev Toxicol* 48, 359-374. doi:10.1080/10408444.2018.1429386
- Knight, J., Rovida, C., Kreiling, R. et al. (2021). Continuing animal tests on cosmetic ingredients for REACH in the EU. *ALTEX* 38, 653-668. doi:10.14573/altex.2104221
- Kolle, S. N., Landsiedel, R. and Natsch, A. (2020). Replacing the refinement for skin sensitization testing: Considerations to the implementation of adverse outcome pathway (AOP)-based defined approaches (DA) in OECD guidelines. *Regul Toxicol Pharmacol* 115, 104713. doi:10.1016/j.yrtph.2020.104713
- Kostadinova, R., Boess, F., Applegate, D. et al. (2013). A long-term three dimensional liver co-culture system for improved prediction of clinically relevant drug-induced hepatotoxicity. *Toxicol Appl Pharmacol* 268, 1-16. doi:10.1016/j.taap.2013.01.012
- Krebs, A., Waldmann, T., Wilks, M. F. 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. doi:10.14573/altex.1909271
- Krewski, D., Acosta, D. Jr., Andersen, M. et al. (2010). Toxicity testing in the 21st century: A vision and a strategy. *J Toxicol Environ Health B Crit Rev* 13, 51-138. doi:10.1080/10937404.2010.483176
- Krewski, D., Andersen, M. E., Tyshenko, M. G. et al. (2020). Toxicity testing in the 21st century: Progress in the past decade and future perspectives. *Arch Toxicol* 94, 1-58. doi:10.1007/s00204-019-02613-4
- Krewski, D., Saunders-Hastings, P., Baan, R. A. et al. (2022). Development of an evidence-based risk assessment framework. *ALTEX* 39, 667-693. doi:10.14573/altex.2004041
- Kuegler, P. B., Zimmer, B., Waldmann, T. et al. (2010). Markers of murine embryonic and neural stem cells, neurons and astrocytes: Reference points for developmental neurotoxicity testing. *ALTEX* 27, 17-42. doi:10.14573/altex.2010.1.16
- Lakhani, C. M., Tierney, B. T., Manrai, A. K. et al. (2019). Repurposing large health insurance claims data to estimate genetic and environmental contributions in 560 phenotypes. *Nat Genet* 51, 327-334. doi:10.1038/s41588-018-0313-7
- LeCluyse, E. L., Witek, R. P., Andersen, M. E. et al. (2012). Organotypic liver culture models: Meeting current challenges in toxicity testing. *Crit Rev Toxicol* 42, 501-548. doi:10.3109/10408444.2012.682115
- Leist, M., Kadereit, S. and Schildknecht, S. (2008a). Food for thought ... on the real success of 3R approaches. *ALTEX* 25, 17-32. doi:10.14573/altex.2008.1.17
- Leist, M., Bremer, S., Brundin, P. et al. (2008b). The biological and ethical basis of the use of human embryonic stem cells for in vitro test systems or cell therapy. *ALTEX* 25, 163-190. doi:10.14573/altex.2008.3.163
- Leist, M., Efremova, L. and Karreman, C. (2010). Food for thought ... considerations and guidelines for basic test method descriptions in toxicology. *ALTEX* 27, 309-317. doi:10.14573/altex.2010.4.309
- Leist, M., Lidbury, B. A., Yang, C. et al. (2012). Novel technologies and an overall strategy to allow hazard assessment and risk prediction of chemicals, cosmetics, and drugs with animal-free methods. *ALTEX* 29, 373-388. doi:10.14573/altex.2012.4.373
- 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
- Leist, M., Ghallab, A., Graepel, R. et al. (2017). Adverse outcome pathways: Opportunities, limitations and open questions. *Arch Toxicol* 31, 221-229. doi:10.1007/s00204-017-2045-3
- Lieber, C. S. (1980). Interaction of ethanol with drug toxicity. *Am J Gastroenterol* 74, 313-320.
- Linkov, I., Massey, O., Keisler, J. et al. (2015). From “weight of evidence” to quantitative data integration using multicriteria decision analysis and Bayesian methods. *ALTEX* 32, 3-8. doi:10.14573/altex.1412231
- Lippa, K. A., Aristizabal-Henao, J. J., Beger, R. D. et al. (2022). Reference materials for MS-based untargeted metabolomics and lipidomics: A review by the metabolomics quality assurance and quality control consortium (mQACC). *Metabolomics* 18, 24. doi:10.1007/s11306-021-01848-6
- Locke, P. A. and Myers, D. B. (2011). A replacement-first approach to toxicity testing is necessary to successfully reauthorize TSCA. *ALTEX* 28, 266-272. doi:10.14573/altex.2011.4.266
- Luechtefeld, T. and Hartung, T. (2017). Computational approaches to chemical hazard assessment. *ALTEX* 34, 459-478. doi:10.14573/altex.1710141
- 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. doi:10.1093/toxsci/kfy152
- Madia, F., Corvi, R., Worth, A. et al. (2020). Making better use of toxicity studies for human health by extrapolating across endpoints. *ALTEX* 37, 519-531. doi:10.14573/altex.2005061
- Maertens, A., Anastas, N., Spencer, P. J. et al. (2014). Green toxicology. *ALTEX* 31, 243-249. doi:10.14573/altex.1406181
- Maertens, A., Luechtefeld, T., Kleensang, A. et al. (2015). MPTP's pathway of toxicity indicates central role of transcription factor SP1. *Arch Toxicol* 89, 743-755. doi:10.1007/s00204-015-1509-6
- Maertens, A., Bouhifd, M., Zhao, L. et al. (2017). Metabolomic network analysis of estrogen-stimulated MCF-7 cells: A comparison of over-representation analysis, quantitative enrichment analysis and pathway analysis versus metabolite network analysis. *Arch Toxicol* 91, 217-230. doi:10.1007/s00204-016-1695-x
- 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
- Maertens, A., Golden, E., Luechtefeld, T. H. et al. (2022). Probabilistic risk assessment – The keystone for the future of toxicology. *ALTEX* 39, 3-29. doi:10.14573/altex.2201081
- Manikkam, M., Haque, M. M., Guerrero-Bosagna, C. et al. (2014). Pesticide methoxychlor promotes the epigenetic transgenerational inheritance of adult-onset disease through the female germline. *PLoS One* 9, e102091. doi:10.1371/journal.pone.0102091
- 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
- Marx, U., Akabane, T., Andersson, T. et al. (2020). Biology-inspired microphysiological systems to advance patient benefit and animal welfare in drug development. *ALTEX* 37, 365-394. doi:10.14573/altex.2001241
- Mazzucato, M. (2013). *The Entrepreneurial State: Debunking Public vs. Private Sector Myths*. 1st edition. Anthem Press. doi:10.4337/roke.2015.01.10
- McNicol, D. (2016). *A Primer of Signal Detection Theory*. United Kingdom: Taylor & Francis Group.
- Meigs, L., Smirnova, L., Rovida, C. et al. (2018). Animal testing and its alternatives – The most important omics is economics. *ALTEX* 35, 275-305. doi:10.14573/altex.1807041
- Miller, G. W. and Jones, D. P. (2014). Nature of nurture: Refining the definition of the exposome. *Toxicol Sci* 137, 1-2. doi:10.1093/toxsci/kft251
- Miller, G. W. (2020). *The Exposome – A New Paradigm for the Environment and Health*. 2nd edition. Academic Press. doi:10.1016/C2017-0-00630-4
- Modafferi, S., Zhong, X., Kleensang, A. et al. (2021). Gene-environment interactions in developmental neurotoxicity: A case study of synergy between chlorpyrifos and CHD8 knockout in human BrainSpheres. *Environ Health Perspect* 129, 77001. doi:10.1289/EHP8580
- Mondou, M., Maguire, S., Pain, G. et al. (2021). Envisioning an international validation process for new approach methodologies in chemical hazard and risk assessment. *Environ Adv* 4, 100061. doi:10.1016/j.envadv.2021.100061
- Muratov, E. N., Bajorath, J., Sheridan, R. P. et al. (2020). QSAR without borders. *Chem Soc Rev* 49, 3525-3564. doi:10.1039/d0cs00098a
- Nantasenamat, C. (2020). Best practices for constructing reproducible QSAR models. In K. Roy (ed.), *Ecotoxicological QSARs. Methods in Pharmacology and Toxicology*. New York, NY, USA: Humana. doi:10.1007/978-1-0716-0150-1_3
- Nelson, P., Urs, N. V. and Kasichyanula, T. R. (2022). Progress in natural language processing and language understanding. In M. V. Albert, L. Lin, M. J. Spector et al. (eds), *Bridging Human Intelligence and Artificial Intelligence. Educational Communications and Technology: Issues and Innovations*. Cham, Switzerland: Springer. doi:10.1007/978-3-030-84729-6_6
- Nielsen, J. B., Graff, C., Rasmussen, P. V. et al. (2014). Risk prediction of cardiovascular death based on the QTc interval: Evaluating age and gender differences in a large primary care population. *Eur Heart J* 35, 1335-1344. doi:10.1093/eurheartj/ehu081
- NRC –National Research Council (2007). *Toxicity Testing in the 21st Century: A Vision and a Strategy*. New York, USA: The National Academies Press.
- O'Brien, P. J. (2014). High-content analysis in toxicology: Screening substances for human toxicity potential, elucidating subcellular mechanisms and in vivo use as translational safety biomarkers. *Basic Clin Pharmacol Toxicol* 115, 4-17. doi:10.1111/bcpt.12227
- 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 and 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
- 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. doi:10.14573/altex.2007091
- Pamies, D., Leist, M., Coecke, S. et al. (2022). Guidance document on good cell and tissue culture practice 2.0 (GCCP 2.0). *ALTEX* 39, 30-70. doi:10.14573/altex.2111011
- Paparella, M., Daneshian, M., Hornek-Gausterer, R. et al. (2013). Uncertainty of testing methods – What do we (want to) know? *ALTEX* 30, 131-144. doi:10.14573/altex.2013.2.131
- Parks Saldutti, L., Beyer, B. K., Breslin, W. et al. (2013). In vitro testicular toxicity models: Opportunities for advancement via biomedical engineering techniques. *ALTEX* 30, 353-377. doi:10.14573/altex.2013.3.353
- Partosch, F., Mielke, H., Stahlmann, R. et al. (2015). Internal threshold of toxicological concern values: Enabling route-to-route extrapolation. *Arch Toxicol* 89, 941-948. doi:10.1007/s00204-014-1287-6
- Patlewicz, G., Ball, N., Becker, R. A. et al. (2014). Read-across approaches – Misconceptions, promises and challenges ahead. *ALTEX* 31, 387-396. doi:10.14573/altex.1410071



- Pendse, S. N., Maertens, A., Rosenberg, M. et al. (2017). Information-dependent enrichment analysis reveals time-dependent transcriptional regulation of the estrogen pathway of toxicity. *Arch Toxicol* 91, 1749-1762. doi:10.1007/s00204-016-1824-6
- Persson, M., Loye, A. F., Jacquet, M. et al. (2014). High-content analysis/screening for predictive toxicology: Application to hepatotoxicity and genotoxicity. *Basic Clin Pharmacol Toxicol* 115, 18-23. doi:10.1111/bcpt.12200
- Piir, G., Kahn, I., García-Sosa, A. T. et al. (2018). Best practices for QSAR model reporting: Physical and chemical properties, ecotoxicity, environmental fate, human health, and toxicokinetics endpoints. *Environ Health Perspect* 126, 126001. doi:10.1289/EHP3264
- Price, E. J., Vitale, C. M., Miller, G. W. et al. (2022). Merging the exposome into an integrated framework for “omics” sciences. *iScience* 25, 103976. doi:10.1016/j.isci.2022.103976
- Prüss-Ustün, A., Vickers, C., Haeffliger, P. et al. (2011). Knowns and unknowns on burden of disease due to chemicals: A systematic review. *Environ Health* 10, 9. doi:10.1186/1476-069X-10-9
- Ramirez, T., Daneshian, M., Kamp, H. et al. (2013). Metabolomics in toxicology and preclinical research. *ALTEX* 30, 209-225. doi:10.14573/altex.2013.2.209
- Rappaport, S. M. and Smith, M. T. (2010). Environment and disease risks. *Science* 330, 460-461. doi:10.1126/science.1192603
- Rossini, G. P. and Hartung, T. (2012). Food for thought ... towards tailored assays for cell-based approaches to toxicity testing. *ALTEX* 29, 359-372. doi:10.14573/altex.2012.4.359
- Roth, A. and MPS-WS Berlin (2019). Human microphysiological systems for drug development. *Science* 373, 1304-1306. doi:10.1126/science.abc3734
- Rothen-Rutishauser, B., Blank, F., Muhlfeld, C. et al. (2008). In vitro models of the human epithelial airway barrier to study the toxic potential of particulate matter. *Expert Opin Drug Metab Toxicol* 4, 1075-1089. doi:10.1517/17425255.4.8.1075
- Rovida, C. and Hartung, T. (2009). Re-evaluation of animal numbers and costs for in vivo tests to accomplish REACH legislation requirements for chemicals. *ALTEX* 26, 187-208. doi:10.14573/altex.2009.3.187
- Rovida, C. (2010). Food for thought ... why no new in vitro tests will be done for REACH by registrants. *ALTEX* 27, 175-183. doi:10.14573/altex.2010.3.175
- Rovida, C., Longo, F. and Rabbit, R. R. (2011). How are reproductive toxicity and developmental toxicity addressed in REACH dossiers? *ALTEX* 28, 273-294. doi:10.14573/altex.2011.4.273
- Rovida, C., Alépée, N., Api, A. M. et al. (2015a). Integrated testing strategies (ITS) for safety assessment. *ALTEX* 32, 171-181. doi:10.14573/altex.1411011
- Rovida, C., Asakura, S., Daneshian, M. et al. (2015b). Toxicity testing in the 21st century beyond environmental chemicals. *ALTEX* 32, 171-181. doi:10.14573/altex.1506201
- Rovida, C., Barton-Maclaren, T., Benfenati, E. et al. (2020). Internationalisation of read-across as a validated new approach 2 method (NAM) for regulatory toxicology. *ALTEX* 37, 579-606. doi:10.14573/altex.1912181
- Samuel, G. O., Hoffmann, S., Wright, R. et al. (2016). Guidance on assessing the methodological and reporting quality of toxicologically relevant studies: A scoping review. *Environ Int* 92-93, 630-646. doi:10.1016/j.envint.2016.03.010
- Sayette, M. A. (2017). The effects of alcohol on emotion in social drinkers. *Behav Res Ther* 88, 76-89. doi:10.1016/j.brat.2016.06.005
- Schmidt, C. W. (2015). Diversity outbred: A new generation of mouse model. *Environ Health Perspect* 123, A64-67. doi:10.1289/ehp.123-A64
- Silbergeld, E. K., Contreras, E. Q., Hartung, T. et al. (2011). Nanotoxicology: “The end of the beginning” – Signs on the roadmap to a strategy for assuring the safe application and use of nanomaterials. *ALTEX* 28, 236-241. doi:10.14573/altex.2011.3.236
- Sillé, F. C. M., Karakitsios, S., Kleensang, A. et al. (2020). The exposome – A new approach for risk assessment. *ALTEX* 37, 3-23. doi:10.14573/altex.2001051
- Skinner, M. K. (2014). Endocrine disruptor induction of epigenetic transgenerational inheritance of disease. *Mol Cell Endocrinol* 398, 4-12. doi:10.1016/j.mce.2014.07.019
- Smirnova, L., Hogberg, H. T., Leist, M. et al. (2014). Developmental neurotoxicity – Challenges in the 21st century and in vitro opportunities. *ALTEX* 31, 129-156. doi:10.14573/altex.1403271
- Smirnova, L., Harris, G., Leist, M. et al. (2015). Cellular resilience. *ALTEX* 32, 247-260. doi:10.14573/altex.1509271
- 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
- Spielmann, H. (2019). Progress in eliminating one-year dog studies for the safety assessment of pesticides. In H. Kojima, T. Seidle, and H. Spielmann (eds), *Alternatives to Animal Testing*. Singapore: Springer. doi:10.1007/978-981-13-2447-5_6
- Steger-Hartmann T., Kreuchwig, A., Vaas, L. et al. (2020). Introducing the concept of virtual control groups into preclinical toxicology animal testing. *ALTEX* 37, 343-349. doi:10.14573/altex.2001311
- Stephens, M. L., Andersen, M., Becker, R. A. et al. (2013). Evidence-based toxicology for the 21st century: Opportunities and challenges. *ALTEX* 30, 74-104. doi:10.14573/altex.2013.1.074
- Stephens, M. L., Betts, K., Beck, N. B. et al. (2016). The emergence of systematic review in toxicology. *Toxicol Sci* 152, 10-16. doi:10.1093/toxsci/kfw059
- Stephens, M., Akgün-Ölmez Gül, S., Hoffmann, S. et al. (2019). Adaptation of the systematic review framework to the assessment of toxicological test methods: Early lessons learned using the zebrafish embryotoxicity test as the test case. *Toxicol Sci* 171, 56-68. doi:10.1093/toxsci/kfz128
- Stucki, A. O., Barton-Maclaren, T. S., Bhuller, Y. et al. (2022). Use of new approach methodologies (NAMs) to meet regulatory requirements for the assessment of industrial chemicals and pesticides for effects on human health. *Front Toxicol* 4, 964553. doi:10.3389/ftox.2022.964553
- Taylor, K., Stengel, W., Casalegno, C. et al. (2014). Experiences of the REACH testing proposals system to reduce animal testing. *ALTEX* 31, 107-128. doi:10.14573/altex.1311151
- Thomas, R. S., Paules, R. S., Simeonov, A. et al. (2018). The US Federal Tox21 program: A strategic and operational plan for

- continued leadership. *ALTEX* 35, 163-168. doi:10.14573/altex.1803011
- Threadgill, D. W. and Churchill, G. A. (2012). Ten years of the collaborative cross. *Genetics* 190, 291-294. doi:10.1534/genetics.111.138032
- Ting, A. H., McGarvey, K. M. and Baylin, S. B. (2006). The cancer epigenome – Components and functional correlates. *Genes Dev* 20, 3215-3231. doi:10.1101/gad.1464906
- Tollefsen, K. E., Scholz, S., Cronin, M. T. et al. (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
- Tolosa, L., Pinto, S. and Donato, M. T. (2012). Development of a multiparametric cell-based protocol to screen and classify the hepatotoxicity potential of drugs. *Toxicol Sci* 127, 187-198. doi:10.1093/toxsci/kfs083
- Tolosa, L., Gomez-Lechon, M. J. and Donato, M. T. (2015). High-content screening technology for studying drug-induced hepatotoxicity in cell models. *Arch Toxicol* 89, 1007-1022. doi:10.1007/s00204-015-1503-z
- Tralau, T. and Luch, A. (2015). Moving from rats to cellular omics in regulatory toxicology: Great challenge toward sustainability or “up-shit-creek without a paddle”? *Arch Toxicol* 89, 819-821. doi:10.1007/s00204-015-1511-z
- Tralau, T., Oelgeschläger, M., Gürtler, R. et al. (2015). Regulatory toxicology in the twenty-first century: Challenges, perspectives and possible solutions. *Arch Toxicol* 89, 823-850. doi:10.1007/s00204-015-1510-0
- Trask, O. J., Jr., Moore, A. and LeCluyse, E. L. (2014). A micropatterned hepatocyte coculture model for assessment of liver toxicity using high-content imaging analysis. *Assay Drug Dev Technol* 12, 16-27. doi:10.1089/adt.2013.525
- Tropsha, A. (2010). Best practices for QSAR model development, validation, and exploitation. *Mol Inform* 29, 476-488. doi:10.1002/minf.201000061
- 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
- van der Zalm, A. J., Barroso, J., Browne, P. et al. (2022). A framework for establishing scientific confidence in new approach methodologies. *Arch Toxicol* 96, 2865-2879. doi:10.1007/s00204-022-03365-4
- 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
- Vinken, M., Benfenati, E., Busquet, F. et al. (2021). Safer chemicals using less animals: Kick-off of the European ONTOX project. *Toxicology* 458, 152846. doi:10.1016/j.tox.2021.152846
- Virchow, R. and Chance, F. (1860). *Cellular Pathology, as Based Upon Physiological and Pathological Histology. Twenty Lectures Delivered in the Pathological Institute of Berlin During the Months of February, March and April, 1858*. New York, USA: R. M. De Witt. doi:10.5962/bhl.title.110759
- von Aulock, S., Busquet, F., Locke, P. et al. (2022). Engagement of scientists with the public and policymakers to promote alternative methods. *ALTEX* 39, 543-559. doi:10.14573/altex.2209261
- Whaley, P., Blaauboer, B. J., Brozek, J. et al. (2021). Improving the quality of toxicology and environmental health systematic reviews: What journal editors can do. *ALTEX* 38, 513-522. doi:10.14573/altex.2106111
- Wilcox, N. and Goldberg, A. (2011). Food for thought ... on validation. A puzzle or a mystery: An approach founded on new science. *ALTEX* 28, 3-8. doi:10.14573/altex.2011.1.003
- Wild, C. P. (2019). The global cancer burden: Necessity is the mother of prevention. *Nat Rev Cancer* 19, 123-124. doi:10.1038/s41568-019-0110-3
- 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, 338-355. doi:10.1021/tx4004038
- Zhang, P., Carlsten, C., Chaleckis, R. et al. (2021). Defining the scope of exposome studies and research needs from a multidisciplinary perspective. *Environ Sci Technol Lett* 12, 839-852. doi:10.1021/acs.estlett.1c00648
- Zhang, Y., Post, W.S., Dalal, D. et al. (2011). QT-interval duration and mortality rate: Results from the third national health and nutrition examination survey. *Arch Intern Med* 171, 1727-1733. doi:10.1001/archinternmed.2011.433
- Zhu, H., Zhang, J., Kim, M. T. et al. (2014). Big data in chemical toxicity research: The use of high-throughput screening assays to identify potential toxicants. *Chem Res Toxicol* 27, 1643-1651. doi:10.1021/tx500145h
- Zhu, H., Bouhifd, M., Kleinstreuer, N. et al. (2016). Supporting read-across using biological data. *ALTEX* 33, 167-182. doi:10.14573/altex.1601252
- Zurlo, J., Bayne, K. A., Brown, D. C. et al. (2011). Critical evaluation of the use of dogs in biomedical research and testing in Europe. *ALTEX* 28, 355-359. doi:10.14573/altex.2011.4.355
- Zurlo, J. and Hutchinson, E. (2014). Refinement. *ALTEX* 31, 4-10. doi:10.14573/altex.1312191

Conflict of interest

Thomas Hartung is consultant for computational toxicology for Underwriters Laboratories (UL) and receives shares of their respective sales. He also holds stock options in and consults ToxTrack LLC. He is named inventor on a patent by Johns Hopkins University on the production of brain organoids, which is licensed to AxoSim, New Orleans, LA, USA, and receives royalty shares and consults AxoSim.

Data availability

No original data was created for this manuscript.

Acknowledgement

This work has received funding from the European Union's Horizon 2020 research and innovation program under grant agreements No. 681002 (EU-ToxRisk) and No. 963845 (ONTOX).