Abstract
Developmental neurotoxicity (DNT) testing has seen enormous progress over the last two decades. Preceding even the publication of the animal-based OECD test guideline for DNT testing in 2007, a series of non-animal technology workshops and conferences that started in 2005 has shaped a community that has delivered a comprehensive battery of in vitro test methods (DNT IVB). Its data interpretation is now covered by a very recent OECD guidance (No. 377). Here, we overview the progress in the field, focusing on the evolution of testing strategies, the role of emerging technologies, and the impact of OECD test guidelines on DNT testing. In particular, this is an example of the targeted development of an animal-free testing approach for one of the most complex hazards of chemicals to human health. These developments started literally from a blank slate, with no proposed alternative methods available. Over two decades, cutting-edge science enabled the design of a testing approach that spares animals and enables throughput to address this challenging hazard. While it is evident that the field needs guidance and regulation, the massive economic impact of decreased human cognitive capacity caused by chemical exposure should be prioritized more highly. Beyond this, the claim to fame of DNT in vitro testing is the enormous scientific progress it has brought for understanding the human brain, its development, and how it can be perturbed.

Plain language summary
Developmental neurotoxicity (DNT) testing predicts the hazard of exposure to chemicals to human brain development. Comprehensive advanced non-animal testing strategies using cutting-edge technology can now replace animal-based approaches to assess this complex hazard. These strategies can assess large numbers of chemicals more accurately and efficiently than the animal-based approach. Recent OECD test guidance has formalized this battery of in vitro test methods for DNT, marking a pivotal achievement in the field. The shift towards non-animal testing reflects both a commitment to animal welfare and a growing recognition of the economic and public health impacts associated with impaired cognitive function caused by chemical exposures. These innovations ultimately contribute to safer chemical management and better protection of human health, especially during the vulnerable stages of brain development.
1 Introduction

Our 2014 article in this Food for Thought … series, “Developmental neurotoxicity – Challenges in the 21st century and in vitro opportunities” analyzed the state of the art and emerging opportunities of developmental neurotoxicity (DNT) testing (Smirnova et al., 2014). Much progress has been made since then. Here, we will relate the strategic development of the field from having no alternative to animal testing for DNT to the recent establishment of the Organisation for Economic Co-operation and Development (OECD) “Initial Recommendations on Evaluation of Data from the Developmental Neurotoxicity (DNT) In Vitro Testing Battery” (OECD, 2023) over 18 years.

DNT is a major public health concern in the context of autism spectrum disorder (ASD), attention deficit hyperactivity disorder (ADHD), and other neurodevelopmental delays (Tab. 1). The developing brain is more vulnerable to environmental perturbations than the mature brain due to pharmacokinetic factors, immature defense mechanisms, and the complex processes of neurodevelopment (e.g., proliferation, migration, and differentiation) (Rice and Barone, 2000; Rodier, 1995). These processes take place in strictly controlled timeframes over several years and create different windows of vulnerability. Interference with developmental processes by, e.g., chemical exposure, can contribute to neurodevelopmental disorders such as ASD, ADHD, and intellectual disabilities (Kuehn, 2010; Sagiv et al., 2010; Grandjean and Landrigan, 2004; Landrigan, 2010, Rossignol et al., 2014). However, very few substances have been identified as neurotoxicants due to the general lack of toxicity information on chemicals (Grandjean and Landrigan, 2006, 2014) (Fig. 1), and current animal testing strategies for DNT have several limitations (US EPA, 1998; OECD, 2007): They are costly ($1.4 million per substance), time consuming (about two years), and require large numbers of animals (> 1,000). In addition, there are justifiable scientific concerns as to the relevance of these studies for human health (see Section 2).

DNT testing is crucial in the context of public health due to the increasing incidence and public awareness of neurodevelopmental disorders such as ASD; the latest data from 2023 suggest that in the US 1 in 36 children of 8 years old is now diagnosed with ASD (Maenner et al., 2023) (Tab. 2). Similarly, from 2009 to 2017, there was a significant increase in the prevalence of ADHD among US children aged 3–17 years (increasing 8.5% to 9.5% according to the National Health Interview Survey, which is a nationally representative survey of the civilian noninstitutionalized population (Zablotsky et al., 2019)) and other conditions that affect cognitive and behavioral development in children. These disorders have significant societal and economic impacts, including lost productivity and the need for lifelong care, which can amount to millions of dollars per case.

Endocrine disruption, especially of the thyroid hormone system, is considered one of the contributors to DNT. Gaylord et al. (2020) analyzed National Health and Nutrition Examination Survey (NHANES) data from 2001-2016 to assess the neurodevelopmental disability burden and associated costs due to early life exposure to polybrominated diphenyl ethers (PBDEs), organophosphates, methylmercury, and lead in the United States. PBDE exposure in utero contributed the most to intellectual disability (ID) burden, resulting in 162 million intelligence quotient (IQ) points lost and 738,000 ID cases, followed by lead, organophosphates, and methylmercury. From 2001 to 2016, IQ loss attributable to PBDEs, methylmercury, and lead decreased or remained stagnant, while organophosphate-attributable IQ loss increased. The overall cost of ID cases decreased from $159 billion annually in 2001-2002 to about $50 billion annually in 2015-2016, largely due to reductions in PBDE and lead exposures. The authors noted that while these trends generally show benefits from chemical regulation, the increased use of potentially harmful substitutions for phased-out chemicals is concerning. Boyle et al. (2021) estimated that early childhood lead exposure in the US from 1999-2010 resulted in

---

Tab. 1: The many reasons why we need DNT testing

| 1. (Global) public health protection by prevention of developmental disorders – The developing brain is particularly vulnerable to toxicants in the context of autism spectrum disorders (ASD), attention deficit hyperactivity disorder (ADHD), and cognitive impairments. |
| 2. Regulatory compliance |
| 3. Guidance for pregnant women |
| 4. De-risking of pediatric clinical trials |
| 5. Economic impacts (a) – Loss of cognitive capacity (IQ score) costs society. |
| 6. Economic impacts (b) – One animal test costs at least $1.4 million. |
| 7. Animal welfare – More than 1000 animals are “used” for each chemical. |
| 8. Scientific advancement |
| 9. Environmental protection |
| 10. Legal and forensic applications – evidence in legal cases and contributing to justice and remediation efforts |
| 11. Consumer confidence |

---

1 https://www.cdc.gov/ncbddd/autism/data.html

Abbreviations: ADHD, attention deficit hyperactivity disorder; AI, artificial intelligence; AO, adverse outcome; AOP, adverse outcome pathway; ASD, autism spectrum disorder; DIV, days in vitro; DNT, developmental neurotoxicity; DNT IVB, DNT in vitro testing battery; DRF, dose-range finding study; EDC, endocrine disrupting chemical; HCS, high content screening; HTS, high-throughput system; IATA, integrated approaches to testing and assessment; ID, intellectual disability; iPSC, induced pluripotent stem cell; IQ, intelligence quotient; ITS, integrated testing strategy; IVIVE, in vitro-to-in vivo; KE, key event; KNDP, key neurodevelopmental process; MIE, molecular initiating event; ML, machine learning; MPS, microphysiological systems; NAM, new approach methodologies; OECD, Organisation for Economic Co-operation and Development; OI, organoid intelligence; PBDE, polybrominated diphenyl ethers; TG, test guideline
a loss of about $46.2 billion annually. Lead exposure of African American infants had the greatest economic impact, with an average loss of $47,116 per child, accounting for 20% ($110 billion) of the total losses despite comprising only 14% of the population. Notably, 74% of the total economic burden was attributed to blood lead levels below 5 μg/dL, highlighting the substantial consequences of even low-level lead exposure in early life.

Bellanger et al. (2015) estimated the neurodevelopmental disability burden and costs attributable to PBDE, organophosphate, and other EDC exposures in the European Union. An expert panel evaluated epidemiological and toxicological evidence contributing to IQ loss, intellectual disability, ASD, and ADHD. PBDE exposure was associated with a 70-100% probability of 873,000 lost IQ points and 3,290 intellectual disability cases, costing €9.59 billion annually. Organophosphate exposure had a 70-100% probability of 13 million lost IQ points and 59,300 intellectual disability cases, at €146 billion annually. EDC exposure had a 20-39% probability of causing 316 ASD cases at €199 million annually, and a 20-69% probability of causing 19,300-31,200 ADHD cases at €1.21-2.86 billion annually. The authors concluded that EDC exposure likely contributes substantially to neurobehavioral deficits and associated costs.

The rising rates of these disorders suggest that lifestyle factors and chemical exposures are likely contributors, making DNT testing an essential tool for identifying and mitigating potential environmental causes. The lack of comprehensive data on the neurotoxic effects of most chemicals, including those found in environmental pollutants, industrial chemicals, drugs, consumer products,
and food additives, underscores the urgent need for reliable DNT testing strategies. However, there is no legal requirement to provide DNT data for most chemicals. Thus, there is no data at all on 99% of all chemicals that humans are exposed to.

Epidemiology is equally complex. On this basis, an authoritative review of the field concluded that only 14 substances have been sufficiently characterized as human developmental neurotoxicants (Grandjean and Landrigan, 2014). We cannot trace ASD or other neurodevelopmental disorders back to just these substances, so there are certainly more, and this is what DNT testing is about. DNT testing can identify environmental and genetic interactions that may lead to these conditions and can thus inform regulations for protecting public health.

Our 2014 article (Smirnova et al., 2014) delved into the pressing issue of DNT testing. It underscored the limitations inherent in traditional animal-based studies. The article highlighted the increasing incidence and public awareness of neurodevelopmental disorders and the significant societal and economic impacts these conditions entail. It also pointed out the challenges faced by epidemiological studies in establishing causal relationships between environmental exposures and neurodevelopmental disorders, due to difficulties in study design, biosampling, and exposure metrics. The article explored the movement towards alternative DNT assessment methods, initiated through international workshops and conferences by expert groups and institutions. This effort aimed to identify promising alternative approaches, such as the use of adverse outcome pathways (AOPs) (Leist et al., 2017) and integrated testing strategies (ITS) (Hartung et al., 2013a; Rovida et al., 2015), to enhance the predictivity and relevance of DNT testing.

The article emphasized the need for innovative biotechnological and computational methods to overcome the limitations of current testing paradigms. Despite the progress made until then, it acknowledged the ongoing challenges in validating and integrating these new approaches into regulatory frameworks. The article called for continued innovation, collaboration, and strategic evaluation to improve the quality and efficiency of DNT testing, ultimately aiming for better protection of human health without relying on animal testing.

2 Limitations of DNT animal testing

The case of DNT is a revealing illustration of how difficult testing for complex hazards in animal models is. Some DNT concerns are covered in the two-generation study OECD Test Guideline (TG) 416, considered the most definitive study for developmental and reproductive toxicity (DART) in animals (OECD, 2001). Due to the potential impact of chemicals on human brain development, the development and adoption of OECD TG 426 (OECD, 2007) and the extended one-generation reproductive toxicity test guideline 443 (OECD, 2018) were significant milestones in the evolution of DNT testing. Both TG 426 and TG 443 have been subject to critical evaluation and refinement to ensure that they reflect the best available science for assessing DNT potential in human health risk assessment using animals. The guidelines emphasize the importance of considering the complex nature of neurodevelopment and the need for a variety of test methods to address different aspects of DNT (Makris et al., 2009; Arts et al., 2023).

OECD TG 426 provides a framework for conducting in vivo DNT studies. The guideline is designed to identify chemicals that may affect the nervous system during development, characterizing any chemical-induced alterations and estimating dose levels for regulatory uses. It includes specific endpoints to evaluate functional, behavioral, and morphological effects on the nervous system, with additional testing of offspring exposed in utero and during early lactation (Makris et al., 2009). However, it cannot realistically measure typical human adverse outcomes such as reduced language capacities, autism spectrum behavior or reduced IQ scores. Instead, it measures some endpoints that are not typical characteristics of human pathologies like altered grooming behavior, altered light-dark preferences, whisker reflexes, delayed eye opening, etc. These guideline tests were not formally validated in ring trials and are both costly and time-consuming. For a single study, approximately 1,200 rat pups are used, and the experimental part of the study lasts about three months. Data evaluation (e.g., pathology reading) takes about two years, with overall costs reaching around $1.4 million per substance. The sensitivity of the tests may not be sufficient to detect subtle neurotoxic effects that could be relevant to human health.

OECD TG 443, also known as the extended one-generation reproductive toxicity study, expands upon previous guidelines by including assessments of reproductive and developmental effects within a single study (Moore et al., 2016). It allows for the examination of multiple generations and includes optional modules to assess DNT and immunotoxicity. This guideline aims to provide a more comprehensive understanding of a substance’s potential reproductive and developmental hazards while using fewer animals than traditional two-generation studies (OECD TG 416). In actual practice, the reduced number of animals needed does not really hold (see below).

The scientific relevance of animal DNT studies to human health effects is questionable (Smirnova et al., 2014):

Interspecies differences: The interspecies differences between humans and the animals used in DNT testing such as rats or mice affect the extrapolation of data from animals to humans, as the mechanisms of neural development and the responses to toxic substances can vary significantly between species. The human brain is our most complex organ, and replicating its development in animal models is inherently challenging.

Genetic variability: Another limitation is the inability of animal models to reflect the inter-individual genetic and epigenetic differences found in human populations. These differences can influence susceptibility to neurotoxicants, and animal models based on inbred strains typically lack this variability.

Behavioral and histological interpretation: Interpreting behavioral effects in animals is difficult. Behavioral tests in animals may not accurately represent the complex human behaviors or cognitive functions that could be affected by neurotoxicants. Additionally, potential artifacts in morphometric neuropathological measures can arise, and criteria for observation measures are not always clear, leading to uncertainty in the evaluation of histological data.
In addition to the scientific concerns, ethical concerns regarding the use of animals in research are significant and have led to calls for the reduction, refinement, and replacement of animal testing. Moreover, the expense and time- and labor-intensity of animal-based DNT tests is prohibitive for routine chemical screening. This results in many potential DNT chemicals remaining unidentified due to the impracticality of testing them all using current animal-based guidelines.

Economic aspects (Meigs et al., 2018) also need to be considered: TG 443, which has a DNT module that can be triggered or be part of routine depending on regulatory requirements, was expected to reduce the costs of TG 416 studies (Schiffelers et al., 2015). However, an analysis commissioned by the European Chemicals Agency\(^2\) showed that the worldwide average price for TG 416 is €285,842 (European average is €318,295), while for TG 443 the average price for the basic study (without second generation and extra cohorts) is €414,273, and with second generation €469,778, i.e., 60% more than TG 416. With extra cohorts, the price increases to €507,444 for the DNT cohort or €440,414 for the immunotoxicity cohort, including both cohorts and the second generation costs €655,195. The expectation that TG 443 would save drastically on animal numbers has also been disproven (Knight et al., 2023; Rovida et al., 2023): While TG 416 in REACH registrations used on average 2,590 animals (3,098 with dose range finding studies (DRF)), TG 443 used 1,318 (1,826 with DRF) and with second generation 2,226 (2,734 with DRF), i.e. a 50% reduction but only for the basic study. Test duration of TG 443 compared to TG 416 is an advantage as long as the second generation is not triggered.

In summary, while animal-based DNT testing has provided critical data for individual regulatory decisions, the limitations have urged the development of alternative testing strategies that are more efficient, cost-effective, and human-relevant. However, the initial absence of alternatives to animal testing posed an enormous challenge.

3 The path of alternatives to in vivo DNT testing up to the DNT-4 meeting in 2014

Our first article on DNT in this series (Smirnova et al., 2014) was written in conjunction with the DNT-4 meeting in Philadelphia. The following summary of how we got to Philadelphia shall illustrate the unique strategic development to replace animal testing. It all started with Alan Goldberg, at the time Director of CAAT, approaching Thomas Hartung, at the time Head of the European Center for Validation of Alternative Methods (ECVAM), in 2004, saying, “Thomas, we have to do something about this.” This sparked an ever-growing team of collaborators working together in this endeavor. Alan Goldberg stated, “I chose DNT as the topic for TestSmart as there were no specific requirements for DNT, so we did not have to match against an in vivo approach.” The most important strategic decision was the engagement of regulatory

---

agencies: The US Environmental Protection Agency (US EPA) from the very beginning, and later the European Food Safety Authority (EFSA) and the OECD (Fig. 2).

The discussion started in Aspen around 2003 with Pamela Lein, Alan Goldberg, and Kevin Crofton talking about the need for DNT testing (Lein et al., 2005). Sandra Coecke, Anna Bal-Price, Chantra Eskes, and Helena Hogberg started working on this topic at ECVAM around the same time (Hogberg et al., 2010, 2011; Bal-Price et al., 2010). Over time, a variety of assays were developed and brought in a growing “family” of researchers including Bill Mundy, Ellen Fritsche, Marcel Leist, Tim Shafer, and many others. When Thomas Hartung, Erwin van Vliet, and Helena Hogberg moved to CAAT in 2009, they started up laboratories focusing on DNT and held DNT workshops. The series of workshops and conferences co-organized by CAAT, ECVAM, and US EPA is summarized in the following.

3.1 The first DNT workshop at ECVAM 2005
At ECVAM, the first workshop was organized as a collaboration between CAAT, ECVAM, and CEFIC3 (the chemical industry in Europe) on April 19-21, 2005, in Ispra, Italy. The article by Coecke et al. (2007) reports on the first workshop. The goal of the workshop was to identify and catalog in vitro methods that could predict and identify DNT hazards. Discussions focused on the science of DNT, including models that could capture critical mechanisms and processes, and policy and strategy for integrating alternative methods into regulatory frameworks. The report outlined recommendations and priorities for future work, emphasizing the need for high-throughput in vitro DNT models and their regulatory validation. It highlighted the importance of identifying in vitro tests that can effectively predict DNT hazards, the critical role of mechanistic understanding in developing these tests, and the need for a collaborative approach of scientists, regulators, and stakeholders to integrate these methods into risk assessment strategies.

Already in this paper on the 2005 workshop, we discussed that embryonic stem cells might be a possibility to humanize the assays. This was a year before Yamanaka developed induced pluripotent stem cells (iPSCs) and was then still an ethical dilemma. We also needed to determine what endpoints and aspects of neural development would need to be modeled in these cultures. The focus was placed on a number of key events, which we felt should be measured to model or assess DNT.

A further question was which type of cell model would be appropriate – how complex would it need to be, and how simple could it be to allow sufficient throughput? There were arguments for and against simple cell lines, complex microphysiological systems (MPS), less sentient organisms, and emerging computational methods. At ECVAM, and later at CAAT, we started to focus on 3D models, i.e., aggregates, which we felt was a good trade-off between complexity and throughput. Honegger et al. (1979) had published in Nature the model of reaggregating fetal rat brain cells, which we successfully adopted. It led us to the first paper describing the use of electrophysiological recording for DNT (van Vliet et al., 2007), and to the first paper on the assessment of DNT using metabolomics (van Vliet et al., 2008). More recently, it allowed us to address the DNT potential of flame retardants (Hogberg et al., 2021).

For many, the call for a concerted effort to address scientific and regulatory challenges sounded impossible at the time, but it started to break down the Herculean challenge into manageable steps and ultimately led to the OECD guidance (OECD, 2023).

3.2 The first international DNT conference DNT-1 in 2006
The first international DNT meeting, TestSmart DNT-1, took place in Reston, VA, in 2006. Its primary objective was to bring together stakeholders (e.g., industry, academia and regulators) to identify concerns relating to science and policy and discuss how to develop alternative testing methodologies (Lein et al., 2007). The resulting paper discussed the need for developing alternative methodologies to traditional animal-based DNT testing and emphasized the collaboration between CAAT, the U.S. EPA, and the National Toxicology Program towards creating TestSmart DNT – aimed at fostering alternative DNT models. These initiatives focused on identifying and prioritizing chemicals posing DNT risks, integrating alternative DNT methods into regulatory decision-making, and exploring possibilities for reducing, refining, or replacing animal use in DNT testing.

The progress from the 2005 workshop to the 2006 conference already reflected significant advances in the field of DNT testing: While Coecke et al. (2007) focused on the initial discussions and the conceptual framework for incorporating in vitro methods into regulatory practices, Lein et al. (2007) highlighted active collaborations and specific initiatives like TestSmart DNT, aimed at the practical implementation of these alternative methods. This transition marked a move from theoretical and strategic planning to actionable steps towards validation and regulatory acceptance and demonstrated an increased commitment to reducing animal testing through scientifically robust in vitro DNT testing methods.

3.3 The second international DNT conference DNT-2 in 2008
In 2008, TestSmart DNT-2 in Reston, VA, again, assessed the progress made in developing DNT alternatives, reassessed the priorities and recommendations set at DNT-1, and established ways to use in vitro data in decision-making. The meeting produced a document with recommendations on how to develop alternative DNT methods for screening and prioritization of chemicals (Crofton et al., 2011), which included:

1. Test methods should incorporate one or more endpoints that model key aspects of human neurodevelopment. The test method models the biological process, the test system employs the

3 https://cefic.org; Cefic, the European Chemical Industry Council, founded in 1972, is the voice of large, medium and small chemical companies across Europe, which provide 1.2 million jobs and accounts for approximately about 14% of world chemicals production.
appropriate cells/organisms, and the endpoints measure relevant developmental features.

2. The ability to correctly and accurately measure the intended DNT endpoint must be demonstrated by the use of a set of compounds termed “endpoint-specific controls” or “tool compounds”. The measured endpoint should reflect the intended neurodevelopmental process.

3. The dynamic range of the DNT endpoint should be characterized to determine the measurable extent of change from control values.

4. Concentration-response relationships should be characterized, ideally testing at least five concentrations over a wide range. This is critical for comparing the sensitivity of different methods.

5. Positive and negative control chemicals should be tested that are known to reliably affect or not affect the measured DNT endpoint by known mechanisms.

6. Initial training sets of chemicals should be tested, including those known to elicit or not elicit DNT responses based on in vitro data. This evaluates the methods’ ability to screen moderate numbers of chemicals.

7. Larger testing sets of chemicals should then be screened, including those known to cause or not cause DNT effects in vivo. This demonstrates the method’s ability to test larger chemical numbers.

8. The method’s sensitivity, specificity and predictivity in identifying chemicals’ DNT potential should be analyzed based on these reference chemical sets.

In summary, the recommendations focused on demonstrating the relevance, reliability, sensitivity, and throughput of alternative DNT methods using well-characterized endpoint assays and reference chemical testing sets as a framework for developing screening level alternative DNT tests. The goal was to establish the methods’ fitness for screening and prioritizing large numbers of chemicals. The DNT-2 recommendations aimed to facilitate a transition from conceptual discussions to practical assay development and use for regulatory purposes. These considerations and experience from the DNT field had an important impact on the development of broader concepts for in vitro test method development and for novel validation approaches of such assays (Leist et al., 2010, 2012; van Thriel et al., 2012).

### 3.4 The third international DNT conference DNT-3 in 2011 in Varese, Italy

DNT-3 was held in Varese, Italy, in 2011 and concluded substantial progress since DNT-2 in applying alternative DNT tests (Bal-Price, 2012). DNT-3 built upon the recommendations from Crofton et al. (2011) and showed progress in several key areas:

1. There was general consensus on the urgent need to develop alternative DNT testing strategies that are faster, more cost-efficient, and predictive of human outcomes. High-throughput in vitro DNT models are needed to test the large numbers of chemicals for which there is little to no DNT data.

2. Significant progress was reported in applying in vitro and non-mammalian test systems to DNT, including human stem/progenitor cell-based assays, though more work is needed to validate these alternative models against human DNT outcomes.

3. Generating data across multiple alternative models using a common set of test chemicals was identified as a critical need to facilitate comparisons and determine which models/endpoints are most predictive.

4. Establishing a reference set of positive and negative control chemicals for DNT was also deemed important for evaluating alternative models.

5. Cell-based assays covering key neurodevelopmental processes like proliferation, migration, differentiation, synaptogenesis, and network formation should be applied as functional DNT endpoints.

6. Computational modeling and bioinformatics approaches should be utilized to evaluate the predictive capacity of alternative DNT models/batteries.

7. Despite progress, in vitro models are not yet able to replicate the complexity of the developing nervous system. Their relevance to human outcomes must still be cautiously evaluated.

In summary, DNT-3 highlighted advancements in developing higher-throughput alternative models based on the DNT-2 framework while identifying crucial data gaps, such as the need for more cross-model comparisons using standardized reference chemicals. The emphasis shifted to practical application of DNT alternatives for screening/prioritization, beyond the initial proof-of-concept stage.

### 4 The first Food for Thought ... article on DNT

Building upon the seminal works discussed above, our 2014 paper (Smirnova et al., 2014) presented several key conceptual advances in the development of alternative approaches for DNT testing and provided a forward-looking perspective on the future of DNT testing.

First and foremost, we emphasized the pressing need for more efficient, cost-effective, and mechanistically informative methods to address the paucity of DNT data for the vast majority of chemicals in commerce. This built upon the discussions in the previous papers, which highlighted the limitations of current animal-based DNT testing paradigms, including their high cost, low throughput, and questionable relevance to human health outcomes. We argued that the development of cheaper, faster, and more predictive alternatives is not only a scientific imperative but also an ethical one, given the large numbers of animals used in regulatory DNT studies.

Secondly, the paper discussed the concept of AOPs and their central role in guiding the development of ITS for DNT. AOPs provide a structured framework for linking molecular initiating events (MIEs) to adverse outcomes (AO) at the organism or population level, via a series of key events (KEs) at the cellular, tissue, and organ levels. By mapping the landscape of known or putative AOPs relevant to DNT, researchers can identify critical pathways and processes that should be targeted by alternative test methods. This advances the mechanistic approach advocated by Crofton et al. (2011) and provides a rational basis for designing
and interpreting in vitro and in silico assays in terms of their relevance to in vivo outcomes.

Thirdly, we highlighted the transformative impact of recent technological advances on the development of alternative DNT models. These included the use of human stem cell-derived neural models, which offer a more physiologically relevant and species-specific platform for studying neurodevelopmental processes and their perturbation by chemicals. Advances in 3D cell culture techniques, such as brain organoids, now allowed the recreation of complex tissue-level interactions in vitro, bridging the gap between traditional 2D models and intact organisms. High-throughput screening (HTS) technologies, such as automated imaging and multi-electrode arrays, now enabled the rapid and quantitative assessment of key neurodevelopmental events, such as neurite outgrowth, synaptogenesis, and network formation. These technological breakthroughs, which were still in their infancy when the earlier papers were published, had greatly expanded the toolkit available for alternative DNT testing.

Building on these advances, we proposed a conceptual framework for linking in vitro test systems to an endophenotype of disturbed functional or structural connectivity in the brain. It was postulated and assumed that all DNT effects eventually resulted from the disturbance of at least one key neurodevelopmental process (KNDP), such as cell proliferation, differentiation, migration, and myelination. This pragmatic approach acknowledged that our current understanding of the complete chain of events linking MIEs to AOs in the developing brain was (and is) still limited. It also anticipated currently discussed test strategies that focus more on determining the highest non-toxic (no toxicity endophenotype) concentration/dose than on exactly defining the type of AO (e.g., attention deficit vs language disturbance). By focusing on a core set of well-characterized neurodevelopmental processes known to be sensitive to chemical perturbation and critical for proper brain development, researchers could develop a battery of fit-for-purpose assays to cover the main pathways of DNT. This framework provided a roadmap for assay development and validation in the absence of complete AOPs.

Another important issue tackled in 2014 was the challenge of defining adversity in the context of alternative DNT models. While apical endpoints, such as changes in behavior or cognitive function, are typically used to define adverse effects in regulatory animal studies, these complex outcomes would be difficult to recapitulate in vitro. We discussed the need to establish clear criteria for distinguishing between adaptive and adverse responses at the cellular and molecular level, taking into account factors such as the magnitude, duration, and reversibility of the effects. This discussion was very much driven by our concomitant discussions on pathways of toxicity (Kleensang et al., 2014) versus pathways of defense (Hartung et al., 2012) in the context of the resilience of biological systems (Smirnova et al., 2015b). We also highlighted the importance of considering the biological relevance of the observed changes, rather than relying solely on statistical significance. This nuanced discussion of adversity built upon the earlier works (Blaauboer et al., 2012) and underscored the need for a careful and context-dependent interpretation of alternative DNT assay results.

Finally, we emphasized several key factors that should be considered when designing and interpreting alternative DNT models. These include potential interspecies differences in neurodevelopmental processes and chemical susceptibility, which underscore the need for human-relevant models, and the importance of considering indirect mechanisms of DNT, such as alterations in thyroid hormone signaling or placental function, which may not be captured by models focused solely on direct effects on neural cells. Additionally, we highlighted the critical role of exposure timing, given the dynamic nature of neurodevelopment and the existence of critical windows of vulnerability. These considerations, while not entirely new, were given greater prominence in the paper, reflecting a more nuanced and biologically informed approach to alternative DNT testing.

In conclusion, the 2014 paper summarized a significant conceptual advance in the field of alternative DNT testing, building upon the groundwork laid by earlier publications. By integrating the latest scientific understanding of neurodevelopment with technological advances in in vitro and in silico modeling, we provided a comprehensive and forward-looking perspective on the challenges and opportunities in this field. We articulated a vision for a new paradigm of predictive toxicology, based on a mechanistic understanding of the pathways underlying DNT and their perturbation by chemicals. Driven by the limitations of current alternative models, we offered a roadmap for future research and development, emphasizing the need for a multi-disciplinary and collaborative approach. As such, the paper served as a valuable synthesis of the state-of-the-art in 2014 and a guidepost for the future direction of alternative DNT testing.

5 DNT-4 – Toward AOPs and fit-for-purpose assays for DNT

DNT-4 was held in Philadelphia, PA, in 2014. The meeting brought together diverse stakeholders (academia, industry, and regulatory bodies) from around the globe (Asia, Canada, Europe, and US), who discussed the next steps required to move in vitro DNT tests forward. Based on experiences from previous DNT meetings, the meeting format included scientific presentations followed by smaller breakout groups in which specific topics were discussed. The following summarizes the discussions at DNT-4. It does not necessarily reflect all opinions of the participants, and the concluding statements are our own reflections of the meeting. As similar topics were discussed in several breakout groups, the report has been organized based on the initial meeting program.

5.1 New concepts and test strategies

The first session explained the AOP framework and gave examples on how it could be used for DNT assessment. The keynote presentation by Dr Kevin Crofton (US EPA, presented by Dr William Mundy, US EPA) described the critical importance of the AOP framework to link DNT research to regulatory needs. Dr Anna Bal-Price, European Commission, further explained the concept of AOPs applied to DNT evaluation. Dr Ellen Fritsche, IUF – Leibniz Research Institute of Environmental Medicine, and Dr
Pamela Lein, University of California Davis, gave practical examples of how in vitro and epidemiological data can inform AOPs. The AOP framework is a tool to combine existing knowledge concerning the linkage between a MIE and an AO at the individual or population level (Ankley et al., 2010). The AOP covers the whole pathway, including chemical properties, MIE, cellular response, organ response, organism response, and finally the population response. One test or study alone will not be able to capture this whole pathway. However, combining data from several DNT tests, including different models (e.g., cells and non-mammalian species) and endpoints for different developmental processes (e.g., proliferation, differentiation, and myelination) with existing in vivo data and data from epidemiological studies can help to reduce the uncertainty and give a better toxicity prediction. It should be emphasized that AOPs are chemical agnostic, meaning that any chemical that perturbs the MIE with sufficient potency and duration will have an effect on the following chain of KEs identified in the AOP. The advantage with this approach is that not all elements in the AOP need to be identified before the concept can be useful (Bal-Price and Meek, 2017). By assembling all existing information, data gaps can be identified to focus efforts on generating useful data and methods. DNT AOPs will provide confidence that cellular changes can lead to DNT effects in humans and will increase the certainty of in vitro tests for risk assessors and regulators to make decisions.

Currently, only a few AOPs for DNT have been developed according to the OECD guidance document (Bal-Price et al., 2015a), but more AOPs can be built once more data is obtained. The AOP framework faces special challenges for neurotoxicity and DNT assessment. Firstly, there is a lack of basic knowledge on the pathophysiology of neurological diseases and the underlying MIEs and KEs. Moreover, there is little understanding of compensatory processes in the nervous system, such as homeostasis and resilience. The complexity of the development of the central nervous system will likely construct a network of AOPs instead of the suggested linear AOPs. However, while it takes time to build AOPs, the incompleteness of an AOP should not impede its use.

It is likely that gene-environment interactions play a crucial role in the case of many neurodevelopmental diseases. While the increase in ASD prevalence is partly due to increased diagnostic criteria, it cannot be fully attributed to diagnostic substitution (Hertz-Picciotto and Delwiche, 2009) and genetic causes; an increasing number of studies suggest that environmental factors contribute to this increase. Several genes involved in ASD and other neurodevelopmental disorders have been shown to be susceptible to environmental perturbation (Pessah and Lein, 2008). In vitro models are preferable to evaluate these gene and environmental interactions and how they might affect neurodevelopment at the cellular level. The adverse effect on the cellular level can then be linked to adverse effects in patients with developmental disorders to build confidence in the AOP. For example, genes involved in synapse formation and elimination have been shown to be disturbed by environmental exposure, and epidemiological studies have shown that many children with ASD have increased connectivity in local circuits of the cortex (Keown et al., 2013). Such an approach to associate cellular effects with disease outcome could be informative for regulators and demonstrate the use of in vitro data for risk assessment.

A breakout group “Science of DNT Models and AOPs” elaborated further on this topic. The previous DNT meetings had already identified several promising models for DNT studies, e.g., cell lines, primary cultures, 3D aggregating cultures, zebrafish, C. elegans, and drosophila (Coecke et al., 2007; Lein et al., 2007). At DNT-4 this breakout group deliberated further advantages and limitations of these models with a focus on what window of neurodevelopment they cover and whether and how genetic variation can influence DNT endpoints in vitro. It was concluded that different alternative models are suitable for various stages of neurodevelopment. Suggested as the most fitting cell models for earlier stages were precursor cell lines and stem cells (e.g., embryonic and iPSCs) together with non-mammalian species such as C. elegans and zebrafish. Some of these models, like iPSCs and the non-mammalian species, can also cover the later stages of development while primary cultures are mainly relevant for the later periods. It was further discussed that different species have divergent developmental time frames, both in vitro and in vivo, e.g., rodent cells develop faster than human cells, and zebrafish have shorter developing time than mice. Shorter development can be an advantage as it shortens the experimental time. However, it is unclear how species might differ in sensitivity to environmental perturbations. Several cell-based studies have reported species differences in response to various compounds; however, there were concerns among the participants that these effects could be artifacts. Especially in the early stage of development the group felt that differences could be large, while the later stages could be more easily compared among the models. It was remarked that these challenges do not only concern differences between species but differences between cells from different donors. Such differences may depend on when the cells are taken and on the genetic background. It is well known that genetics play a role in the sensitivity of individuals to exposure to chemicals and drugs, and it is therefore to be expected that genetics will influence DNT endpoints in vitro. As the human population is heterogenetic and includes vulnerable subgroups, this can be seen as an advantage. Genetic variations have long been explored in simpler organisms such as C. elegans and zebrafish (Nishimura et al., 2015). However, with the recent use of iPSCs, interindividual differences are increasingly studied also in human cells (Pamies et al., 2017).

It was specified that human cells would be preferred in DNT studies, as this would avoid species differences. However, there is a tremendous dearth of relevant data on human exposure and DNT effects, while in vivo rodent data is available. Therefore, correlating data from rodent cells with rodent in vivo data can act as a bridge to extrapolate data on human cells to the human situation (Maass et al., 2023; Algharabily et al., 2023). Nevertheless, models need to be characterized based on maturation as a function of

---

days in vitro (DIV) and associated with the developmental stage in vivo. But is 7 DIV comparable to 7 days in vivo? It was discussed that measurements of specific markers could contribute to identifying the developmental stage of the model, e.g., specific miRNA, mRNAs and/or proteins.

The conclusion of this breakout group was that alternative models need to be well characterized to be useful in any toxicity study, including DNT. Often models lack information about metabolic capacity, understanding of different cell types, and association to the in vivo situation. However, many models have shown great competence and are useful for measurement of several developmental processes such as differentiation, proliferation, migration, and survival. The in vivo (rodent) predictability for a subset of chemicals seems to be good, however, many more chemicals need to be tested to fully understand the real potential (see below for deeper discussion). To better understand the relevance to the human situation, the DNT community could benefit by interacting with the clinical community to correlate toxicity to, e.g., exposure scenarios, existing biomarkers, and pathological mechanisms.

5.2 Mechanistic and omics tools, as well as functional endpoints, to increase assay information content
In this session, different approaches to evaluate DNT using omics tools were presented. Dr Jennifer Freeman, Purdue University, gave some examples using the zebrafish, and Dr Lena Smirnova described the use of 3D in vitro brain models. Dr Milou Dingemans, Utrecht University, informed how to integrate multiple data from the European Commission-funded project DENAMIC5.

The benefits of using non-mammalian species, such as zebrafish and C. elegans, for DNT assessment include that they are inexpensive, fast, amenable to genetic modifications, and allow whole organism and behavior studies. Zebrafish can be combined with transcriptomic and epigenomic approaches to identify DNT mechanisms for chemicals (Lee and Freeman, 2014). The DENAMIC project includes in vitro, zebrafish, in vivo models, and human cohorts, and applies transcriptomics, proteomics, neurotransmitter profiles, and miRNA to identify biomarkers for neurodevelopmental disorders. The focus is on mixtures, and the data is integrated with the aim to understand mechanisms of the interactions of low doses and neurodevelopmental disorders.

The benefit of using 3D in vitro models include increased cell-cell interactions and better representation of the complexity of the in vivo brain. Different single or multi-cellular human and rodent model types were evaluated for their advantages and disadvantages. The models are combined with transcriptomics, metabolomics, and miRNA profiles to identify pathways of toxicity after chemical exposure that may lead to adverse effects on the developing brain (Smirnova et al., 2015a).

The use of omics approaches will generate high-content information to better understand the mechanisms of chemicals as in the vision for toxicology in the 21st century (Krewski et al., 2020). However, the generation of massive data might lead to new challenges and was discussed in the breakout group on techniques.

The keynote by Marcel Leist described the use of human ESCs combined with transcriptomics endpoints. It was demonstrated that statistical prediction models could identify signature transcriptome changes for subgroups of DNT toxicants. Short toxicant exposures may be more useful to identify modes-of-action, while longer exposures of differentiating stem cell models lead to epigenetic changes that affect the differentiation track of the cells and have permanent effects that remain even after drug washout. The conclusion was that transcriptome patterns, recorded at non-cytotoxic chemical exposures, allow chemical classification and, therefore, hazard prediction.

5.3 How to accelerate testing for DNT
Speakers from US government described current screening approaches for DNT at NIEHS by Dr Mamta Behl and at US EPA by Dr William Mundy and Dr Timothy Shafer. As the number of chemicals that lack DNT data is too high to test with traditional in vivo approaches, the National Toxicology Program at NIEHS is working on strategies to screen and prioritize compounds for in vivo DNT testing. A list of 80 compounds was developed and tested by collaborators using various cell systems such as primary, ESC, iPSs, and cell lines containing both neurons and astrocytes, and on non-neuronal cells, e.g., cardiomyocytes. Alternative models such as nematode and zebrafish were also included. The prioritization is based on various data analysis models.

The high-throughput testing program includes ToxCast (700 assays and over 2000 chemicals tested) and Tox21 (in partnership with NTP, FDA and NCATS, screening more than 10,000 chemicals). Moreover, National Health and Environmental Effects Research Laboratory/ToxCast are developing assays for more complex endpoints including DNT. EPA assays for DNT were developed in line with the AOP framework with high-throughput molecular assays, e.g., ion channels, receptors, enzymes; for MIIs, proliferation, differentiation, neurite outgrowth, synaptogenesis, migration, and apoptosis; multi electrode array for KEs, and a zebrafish behavior assay for AO. A list of positive and negative reference chemicals has been developed.

5.4 The use of AOPs and alternative assays for safety assessment
For the alternative assays to be useful, they need to be applied in a regulatory framework. However, different regulatory bodies have different requirements and needs. In this session Dr William Slikker from NCTR, FDA, and Dr Anna Lowit from Health Effects Division, US EPA, gave two regulatory agencies’ perspectives on the use of alternative methods for DNT assessment.

In the case of drug regulation at FDA, it was described that the challenge with only using mechanistic data is that drugs are not entirely pharmacologically identical. In this case, the data may suggest similar risk, but the “safety margin” may not be the same. However, mechanistic data can still be helpful and together with in vivo experiments provide a better understanding of the DNT effects.

5 https://cordis.europa.eu/project/id/282957
The regulation at US EPA is more flexible, and the use of *in vitro* data when there is knowledge from an AOP can be used to support read-across for similar compounds. Furthermore, *in vitro* data can be used in a weight of evidence evaluation to determine data needs or to review a waiver justification. The US EPA supports the 3Rs in their regulatory program. In the case of DNT, US EPA granted 8 waivers and requested 1 study between December 8, 2011, and April 11, 2014. Several *in vitro* study design considerations to satisfy regulatory use were discussed including cell types, presence/absence of serum, parent chemical vs. metabolite, and test concentrations. Furthermore, there are different fit for purpose criteria depending on the intended use, e.g., screening and prioritization vs replacement of *in vivo* guideline or single assay vs test battery.

Dr Thomas Hartung discussed the use of ITS (Hartung et al., 2013a) for DNT assessment. An ITS is an algorithm to combine (different) test result(s) and, possibly, non-test information (existing data, *in silico* extrapolations from existing data or modeling) to give a combined test result. They often will have interim decision points at which further building blocks may be considered. Several aspects should be considered, for example flexibility in combining components of the ITS, the applicability domain, and efficiency in terms of cost, time, and technical difficulties.

### 5.5 Frontiers in DNT testing

The last session of the meeting described cutting-edge technologies for DNT presented by stakeholders from government and academia. Dr Andrea Seiler from the German Federal Institute for Risk Assessment (BfR) described a joint project funded by the German Ministry for Research and Education (BMBF) with the goal to develop standardized predictive cell-based *in vitro* assays for DNT testing. Dr Nisha Sipes from US EPA used literature mining to identify MIEs in an AOP framework for cleft palate. Such informatic data can be applied to develop multicellular virtual-tissue models, e.g., neural tube closure in the virtual embryo, that are based on cell-level models driven by biological networks and rules that can be used for predictive toxicology. Extensive characterization of human iPSCs cultured in 3D was presented by Dr David Parnies from Johns Hopkins University. Human iPSCs allow the study of patients with different genetic backgrounds and can explore whether certain genes convey sensitivity to chemical exposure. Dr Luc Stoppini from the University of Applied Sciences Geneva investigated the functionality of 3D ESCs and iPSCs using various multielectrode arrays. Effects of exposure to chemicals were further evaluated with receptor gene expression studies and metabolomics for neurotransmitter and blood brain barrier. Finally, Dr Keith Cheng from Penn State College of Medicine discussed a novel imaging technology in whole zebrafish. The high-throughput technique can detect changes in any cell type at cell resolution (voxel dimensions of ~1 micron) and makes it possible to identify and characterize virtually every organ, tissue, and cell type contributing to soft tissue architecture, including specific structures such as nerve tracts and vessels.

#### 5.6 Conclusions from DNT-4

The meeting ended with a panel and plenary discussion of the steering committee moderated by Dr Alan Goldberg. Several challenges and limitations with current approaches were identified. The relatively new AOP framework is still not mature enough to use for regulatory decisions, but it will be important to understand how to link cellular events to AOs. The major challenges with the AOPs are to understand the relationships between the events and to get quantitative data. To achieve this, different models and technologies are needed. Some of the new human cell-based models (e.g., iPSCs) and technologies (e.g., omics and miRNA) might be useful tools, particularly when linked to epidemiological studies. However, the models need to be very well characterized.

There is a need to accelerate chemical testing in the assays that have been developed, make sense of the data from high-content assays, and link cellular effects to AOs. Prediction and interpretation cannot be determined without screening more compounds. The major limitation of current DNT tests is still the same as at DNT-2 and -3: lack of data generation.

For this, reference chemicals are needed, which is challenging as there is limited DNT *in vivo* data for most chemicals. After identification and testing of reference chemicals there should be a selection of tests to build a battery and eventually an ITS. To collect the generated and existing data in a common database will be essential. Furthermore, a strong policy program that can influence funding organizations (both in EU and US) to support assay development and chemical screening is needed. Establishment of a DNT secretariat including different stakeholders from various parts of the world was suggested as a path forward.

There seems to be a change in the mindset of regulators, with a first step to make use of *in vitro* tests combined with revised *in vivo* testing for risk assessment. However, it should be kept in mind that different regulatory bodies have different needs. It will be crucial to have them involved during test development. To demonstrate that *in vitro* tests are useful we need to generate more data, validate the tests, and link the generated data to AOs in humans.

### 6 Current methodological advances as a basis for a new approach for DNT

The biotech revolution has clearly fueled DNT testing: the use of stem cells, the development of organotypic cultures, and high-content methods have significantly contributed to the development of alternative testing methods for DNT by providing more human-relevant and mechanistically informative systems that can reduce reliance on animal testing (Groot et al., 2013).

#### 6.1 Stem cells

The availability of human stem cells, particularly human iPSCs, has revolutionized the field of DNT testing by providing new methodologies that are more relevant to human biology. iPSCs can be differentiated into various cell types of the central and peripheral nervous systems, enabling the modeling of different stages of brain development and the assessment of chemical toxicity during these stages (Yamada et al., 2019; Kobolak et al., 2020).
Gene-environment studies using patient cells: iPSCs facilitate the study of gene-environment interactions by using cells from patients with developmental disorders (Ilieva et al., 2018; Russo et al., 2019; Wegscheid et al., 2021; Villa et al., 2021; Santos et al., 2023; Kilpatrick et al., 2023). This approach provides insight into whether certain genetic backgrounds confer increased sensitivity to environmental stressors. This kind of research is crucial for understanding the complex interplay between genetic predispositions and environmental factors in the development of neurological disorders.

Gene engineering with risk genes: The advent of gene-editing technologies like CRISPR/Cas9 has allowed to introduce or correct mutations in iPSCs. This enables the study of specific risk genes in a controlled environment and the observation of their effects on neurodevelopment: For instance, our earlier study demonstrated a potential synergy between a mutation in the high-risk autism gene CHD8 and exposure to the organophosphate pesticide chlorpyrifos in an iPSC-derived 3D brain model (Modafferi et al., 2021). By engineering iPSCs with known risk genes for neurological conditions, scientists can dissect the pathways through which these genes contribute to disease and identify potential therapeutic targets.

“Living biopsy” and disease modeling: iPSCs can be considered a “living biopsy” of a patient’s condition, as they capture the patient’s genetic makeup and can be differentiated into disease-relevant cell types, providing a platform for studying disease mechanisms and potential treatments (Smirnova and Hartung, 2024).

Personalized medicine and toxicology: The use of iPSCs in DNT testing paves the way for personalized medicine and toxicology. By generating iPSCs from individual patients, it is possible to create personalized models of disease and predict individual responses to drugs and environmental toxicants (Fritsche et al., 2018). This approach could lead to more tailored and effective treatments with fewer side effects, as well as safer and more targeted drug development.

Challenges and future directions: Despite these advances, challenges remain for the use of human stem cells for DNT testing. These include the need for improved methods to differentiate iPSCs into fully mature and functional neurons and glia, the development of standardized protocols for toxicity testing, and the integration of these new methods into regulatory frameworks (Fritsche et al., 2018; Yamada et al., 2019).

As the technology progresses, it is expected that human stem cell-based models will become increasingly important tools for assessing the safety of chemicals and drugs, ultimately reducing the reliance on animal testing and improving human health outcomes.

6.2 Organotypic cultures and microphysiological systems
Organotypic cultures are complex in vitro models that maintain or reconstruct the architecture and multi-cellular complexity of the original tissue. We see this as a revolutionary change (Hartung and Tsatsakis, 2021) in achieving relevant human cell cultures. These cultures can, in the case of DNT, mimic key aspects of human neurodevelopment and can be used to study the basic biological processes that are fundamental to understanding DNT, such as differentiation, proliferation, migration, and neurite growth. By testing the disturbance of these biological activities by chemicals, we can identify potential neurotoxicants. The bioengineering of MPS is an enormously important task.

We have organized three large workshops with opinion leaders from all over the world (Marx et al., 2016, 2020, and in preparation). The second one led to a Science paper on human MPS for drug development (Roth et al., 2021). Out of this, also, developed a series of MPS World Summits and the international MPS society. In 2022 in New Orleans, out of 655 registrants, 65 came from the FDA, showing the enormous resonance of this topic with the agency. 430 people met there in person, which was tripled to 1300 participants in Berlin 2023, and similar numbers are expected for Seattle in June 2024.

Our own work led us tohumanize brain organoids (Hogberg et al., 2013), published shortly after the first iPSC-derived model (Lancaster et al., 2013) of the first brain organoids. In 2016, at AAAS, we finally showed how we can mass-produce them (Pamies et al., 2017). Since then, we have been using this approach for a number of disease models. These models are spontaneously electrophysiologically active, and they include most of the brain cells we know except for microglia, which can be added. This includes all types of neurons we looked for, as well as astrocytes, oligodendrocytes, even in reasonable proportions – the standard model has about 40% glial cells, and we have just developed protocols to increase them to physiological levels of 50% (Morales Pantoja et al., 2023b). One of the key features is myelination, with about 30% of the axons being myelinated (Pamies et al., 2017; Chesnut et al., 2021; Romero et al., 2023). One can see the beautiful structures of oligodendrocytes wrapping themselves around the axons, which is a quite unique feature, because very few human models to date show myelination. We used this model for DNT, assessing pesticides such as rotenone (Pamies et al., 2018) and chlorpyrifos (Modafferi et al., 2021), but also the antidepressant paroxetine (Zhong et al., 2020), which has been under debate for about 20 years as possibly contributing to neurodevelopmental disorders. The discussion was not closed, but we could show that clinically relevant concentrations disturb brain development in our experimental system.

One of the key ideas is that these brain organoids show all of the mechanisms we have identified before as critical for DNT. Is this, perhaps, a possibility to replace the extensive battery of tests with something that combines all of these assays? Because we observe aspects of neurodifferentiation, myelination, neurite outgrowth, synaptogenesis, glial migration, and gliosis, and the neural network through electrophysiology in a single model, it appears possible to multiplex the different assays. The brain organoids undergo at least a critical phase of development, reflecting currently about five months of embryo development, possibly even more if we only drive them towards this. So, we can probably observe a lot of DNT-relevant mechanisms in a single set-up. This may be more promising than developing a lot of different assays, which no lab can have all at the same time for the...
same substance. This was the starting point for a project funded by the EPA, announced when the EPA in 2019 decided to move out of animal testing by 2035. The project aims to develop a mini “brainbow” of fluorescent reporter genes engineered into these brain organoids so that, noninvasively, all of these mechanisms can be studied at the same time (Romero et al., 2023). We have developed shell electrodes that embrace the brain organoids, allowing 3D electrophysiology around them (Huang et al., 2022).

Another line of work is to use brain organoids to study gene-environment interactions (Butera et al., 2023; Suciu et al., 2023a). We cannot explain the enormous increase in ASD (Tab. 2) based only on genomics and suspect there is an environmental component. Our hypothesis is that a susceptible genome meeting chemical exposure at a vulnerable time contributes to the development of ASD. In Modafferi et al. (2021), we showed such a pair for the first time: CHD8, a known risk gene for autism, and chlorpyrifos, a substance with some liabilities at least at high concentrations and doses, synergize. This result has now led to the creation of an NIH Autism Center of Excellence, where we are following in total 175,000 children in 18 partner centers to look for such gene-environment interactions, and we are verifying these pairs experimentally in brain organoid systems as a second ongoing part of the project.

Challenges and future directions: Though more complex models have the potential to enhance DNT testing, there are some challenges to consider before their use in regulatory applications (Hogberg and Smirnova, 2022). The major limitations are lack of standardization of protocols and lower reproducibility. Moreover, the throughput of such models is still low compared to the simple methods previously developed. The culture time is often extended and can be costly, making such models, as of today, more useful as a follow-up method for a small set of prioritized compounds identified in an initial screening approach.

Taken together, organotypic cultures and MPS as complementary or orthogonal assays for the current screening battery of assays for DNT can support the translation of in vitro mechanistic effects to in vivo DNT outcomes. Once these models can demonstrate that they follow the readiness criteria for regulatory application (Bal-Price et al., 2018a) it will be time to refine the current DNT in vitro test battery (DNT IVB).

Enhanced efficiency and throughput: HCS methods enable the rapid and efficient processing of moderate to large numbers of chemicals, essential for the development of high-throughput DNT testing strategies. This capability is crucial given the vast number of chemicals in the environment that have not been tested for neurotoxicity due to the time-consuming and resource-intensive nature of in vivo DNT testing.

Multiparametric analysis: One of the key strengths of HCS is its ability to simultaneously measure multiple parameters within the same assay, providing a more comprehensive understanding of neurotoxic effects. This multiplexing capability allows the assessment of various aspects of neuronal health, including cell viability, neurite outgrowth, synaptic function, and cellular signaling pathways, within a single experiment (Li and Xia, 2019).

Improved sensitivity and specificity: The automated nature of HCS, combined with advanced image analysis algorithms, can enhance the sensitivity and specificity of DNT testing. High-content methods can detect subtle changes in neuronal morphology and function that may be indicative of neurotoxicity, even at low doses of chemicals (Schmuck et al., 2017; Persson and Hogberg, 2016). This sensitivity is critical for identifying potential neurotoxicants that may not produce overt toxicity but could still have significant impacts on neural development and function.

Application to complex models: HCS methods are compatible with complex in vitro models, including 3D organoid cultures and human iPSC-derived neuronal models (Fritsche et al., 2018). These advanced models more closely mimic human neural development and disease, enhancing the relevance of DNT testing to human health. High-content methods allow the detailed analysis of these complex models, providing insights into the mechanisms of neurotoxicity and the potential for developmental disorders (Schmuck et al., 2017).

Challenges and future directions: Despite the advantages of HCS in DNT testing, challenges remain. These include the need for standardized protocols and validation of HCS methods for regulatory acceptance. Additionally, the complexity of data generated by high-content methods requires sophisticated bioinformatics tools for analysis and interpretation (Li and Xia, 2019). The integration of HCS with emerging technologies, such as machine learning (ML) and artificial intelligence (AI), holds promise for addressing these challenges. These technologies can enhance the analysis of complex datasets, improve the predictive power of DNT testing, and facilitate the identification of novel neurotoxicants (Schmuck et al., 2017).

In conclusion, HCS methods have revolutionized DNT testing by providing a more efficient, sensitive, and comprehensive approach to evaluating the neurotoxic potential of chemicals. As these methods continue to evolve and integrate with advanced computational tools, they will play an increasingly important role in protecting human health from the adverse effects of environmental chemicals.

6 https://www.epa.gov/newsreleases/epa-awards-nearly-850000-johns-hopkins-university-advance-research-alternative-methods
7 https://www.epa.gov/newsreleases/administrator-wheeler-signs-memo-reduce-animal-testing-awards-425-million-advance
6.4 Key characteristics of developmental neurotoxicants

The concept of “key characteristics”, properties of chemicals and other agents that confer potential hazard, was first developed for carcinogens and was based on properties of known human carcinogens as classified by the International Agency for Research on Cancer (IARC). These key characteristics of carcinogens were applied in the evaluation of diverse carcinogens and are now used as the basis for the evaluation of mechanistic data at IARC. Recently, the key characteristics of male and female reproductive toxicants (Arzuaga et al., 2019; Luderer et al., 2019) and of endocrine disrupting chemicals (La Merrill et al., 2020; Cediel-Ulloa et al., 2022) were described, and those for other toxicant areas are in development. A (developmental) neurotoxicity working group met at UC Davis on September 17-18, 2019, to develop key criteria for DNT. Pamela Lein, UC Davis, and Thomas Hartung, Johns Hopkins, co-chaired the meeting which was hosted by Martyn Smith (UC Berkeley) and Lauren Zeise (OEHHHA, CalEPA). The group examined the literature and developed 12 key characteristics, considering neurotoxicity both during development and in later life. They are currently refining the key characteristics in a series of follow-up teleconference calls and preparing a manuscript for publication.

6.5 Advances in artificial intelligence supporting DNT testing

AI has significantly transformed DNT testing, offering innovative approaches to understanding and predicting the neurotoxic effects of chemicals and drugs. This transformation is evident in several key areas, including the enhancement of predictive models, the integration with complex biological systems, and the improvement of data analysis and interpretation.

Enhanced predictive models: AI, particularly ML, has been instrumental in developing predictive models for DNT testing in the context of the ongoing ONTOX project (Vinken et al., 2021, see below). These models can analyze vast datasets, identifying patterns and relationships that may not be apparent using traditional analysis methods. For instance, a study demonstrated the use of ML to predict in vitro neurotoxicity induced by nanoparticles, highlighting the potential of non-testing approaches in hazard assessment (Furxhi and Murphy, 2020). By leveraging features such as exposure dose, duration, and cell type, AI models can provide a more nuanced understanding of neurotoxicity risks.

Integration with complex biological systems: AI’s ability to handle complex, high-dimensional data makes it particularly suited to integrating with advanced biological systems used in DNT testing, such as human iPSC-derived neural constructs (Schwartz et al., 2015). These systems can model human neurodevelopment more accurately than animal models, but they generate large amounts of data that can be challenging to analyze. AI can process and interpret this data, identifying key indicators of neurotoxicity and enhancing the relevance of DNT testing to human health. We have discussed earlier the opportunities of modeling MPS by computational approaches (Smirnova et al., 2018).

Improvement of data analysis and interpretation: HCS, which generate large volumes of data on cellular responses to chemicals, have become a cornerstone of modern DNT testing (Shafer, 2019). AI algorithms can analyze these data efficiently, extracting meaningful insights on neurotoxic effects. For example, AI has been used to map drug-induced neuropathy through in situ motor protein tracking, combining imaging data with ML for a more accurate assessment of neurotoxicity (Yi et al., 2021).

Challenges and future directions: Despite these advances, challenges remain in the application of AI to DNT testing. These include the need for large, high-quality datasets for training AI models, the interpretation of AI-generated predictions, and the integration of AI approaches into regulatory frameworks (Fritsche et al., 2017). Addressing these challenges will require continued collaboration between toxicologists, data scientists, and regulatory bodies.

AI has the potential to revolutionize DNT testing by enhancing the predictive accuracy of neurotoxicity assessments, enabling the integration of complex biological data, and improving the efficiency of data analysis. As AI technologies continue to evolve, they will play an increasingly important role in identifying neurotoxic risks and protecting human health.

The EU ONTOX project (Vinken et al., 2021), funded by Horizon 2020, is at the forefront of leveraging AI for DNT testing (beside liver and kidney toxicity) and broader chemical risk assessments without the use of animal testing. By focusing on the development of non-animal new approach methodologies (NAMs) and probabilistic risk assessment, ONTOX aims to align with 21st-century toxicity testing principles. A key aspect of the project involves addressing the acceptance and validation of AI in risk assessment, as highlighted during the first ONTOX Stakeholder Network Meeting held in March 2023 (Diemar et al., 2024). This meeting brought together various stakeholders, including regulatory authorities, companies, academia, and non-governmental organizations, to discuss the challenges and opportunities associated with implementing AI-driven NAMs and probabilistic risk assessment. The discussions underscored the need for capacity building, sustainability, and regulatory acceptance of AI technologies in the context of ensuring consumer safety and advancing chemical risk assessment methodologies. The ONTOX project’s efforts to integrate AI into DNT testing and chemical risk assessment represent a significant step towards reducing reliance on animal testing while enhancing the accuracy and efficiency of toxicity evaluations (Diemar et al., 2024).

In summary, these advanced biotechnological tools have enabled the creation of more relevant and efficient in vitro models for DNT testing. They have the potential to provide mechanistic insights into how chemicals affect neurodevelopment, which is
crucial for the development of ITS that can ultimately reduce the need for animal testing.

7 Conceptual advances as a basis for a new approach for DNT

7.1 The rise of integrated testing strategies (ITS) aka integrated approaches to testing and assessment (IATA)

The consensus report on the future of animal-free systemic toxicity testing, which emerged from expert workshops convened by CAAT-Europe, outlined a general strategy for animal-free test approaches (Basketter et al., 2012; Leist et al., 2014). This strategy was informed by the US National Research Council’s vision for toxicity testing in the 21st century (NRC, 2007), which advocated for a shift towards more human-relevant, non-animal methods.

The early discussions and recommendations for animal-free systemic toxicity testing laid the groundwork for the development of ITS. These strategies were envisioned to provide a more effective approximation of regulatory information needs than standalone assays. ITS are designed to integrate various information sources, including in vitro assays, in silico models, and human biomonitoring data, to predict the toxicity of substances. This integration is facilitated by computational tools such as Bayesian networks and ML. Bayesian networks are probabilistic models that can combine data from different sources and account for uncertainties, providing a structured approach to integrating evidence and making predictions. ML, on the other hand, can analyze large datasets to identify features that are predictive of toxicity outcomes, improving the accuracy and efficiency of toxicity predictions.

The promise of ITS lies in their ability to approximate the information that would otherwise be obtained from animal testing, but in a more human-relevant and ethical manner. By leveraging emerging tools for data integration, ITS can provide a more comprehensive assessment of potential toxicants, considering various factors such as toxicokinetics, hazard testing, and the mapping of information along AOPs (Leist et al., 2014).

The development of ITS represents a significant advance in the field of toxicology, aiming to reduce the reliance on animal testing while still meeting the regulatory requirements for safety assessment. As these strategies continue to evolve, they are expected to become an integral part of the regulatory landscape, providing a more efficient and ethically responsible approach to toxicity testing.

7.2 The role of AOPs

The concept of AOPs (Willet, 2018) has become a cornerstone in the rational design of ITS for DNT testing. This framework is instrumental in identifying specific targets and mechanisms that are critical in the development of neurotoxic effects, thereby facilitating the development of testing strategies that are both more mechanistically informed and more predictive of human health outcomes.

Importance of AOPs in ITS design for DNT testing: AOPs offer a structured approach to understanding the complex mechanisms underlying DNT, enabling researchers to pinpoint specific biological processes that can be targeted for testing. By mapping out the sequence of events that leads to adverse neurodevelopmental outcomes, AOPs help in identifying relevant biomarkers and endpoints that can be incorporated into ITS (Hernández-Jerez et al., 2021; Willet, 2018). This mechanistic understanding is crucial for developing assays that are not only sensitive to specific neurotoxic effects but also relevant to human health.

Facilitating mechanistically informed testing strategies: The identification of KEs and targets within DNT AOPs has significantly contributed to the development of more mechanistically informed testing strategies. For instance, the EFSA Panel on Plant Protection Products and their Residues developed AOP-informed IATA case studies for the DNT hazard identification of pesticides like deltamethrin and flufenacet (Hernández-Jerez et al., 2021). By focusing on specific KEs identified in the AOPs, such as alterations in neural proliferation, differentiation, and synaptogenesis, these strategies can more accurately predict the potential for chemicals to cause DNT.

Enhancing predictivity with emerging tools for data integration: The integration of emerging tools for data analysis, such as Bayesian networks and ML, has further enhanced the predictivity of ITS for DNT testing. Bayesian networks, for example, allow for the probabilistic quantification of the weight of evidence across different data sources, including in vitro assays and in silico models, within the AOP framework. ML algorithms can analyze complex datasets from high-throughput screening assays to identify patterns and predict outcomes based on identified KEs (Zhang et al., 2023). These computational tools enable the integration of diverse data types, improving the ability of ITS to approximate regulatory information needs effectively.

Conclusion: The development and application of AOPs in the design of ITS for DNT testing represents a significant advance in the field of toxicology. The currently available AOPs for DNT, however, are relatively few: Data retrieved from AOP-Wiki11 on 24 Mar 2024 list eight AOPs (#6, #12, #13, #17, #31, #54, #499, #500), of which five are endorsed by OECD’s WPHA/WNT12. It will take a community effort to expand this, to combine it to a network (Pistollato et al., 2020; Sachana et al., 2021b; Spînu et al., 2022; Pitzer et al., 2023), and to make it the basis of a testing strategy. By providing a structured framework to understand the mechanistic basis of neurotoxicity, AOPs can facilitate the development of testing strategies that are not only more informed by the underlying biology but also more predictive of human health outcomes. The integration of advanced computational tools for data analysis further enhances the capacity of ITS to provide com-

11 https://aopwiki.org
8 The development of new approaches for DNT as an example for the strategic development of alternatives to animal testing

The 2014 conference represented a turning point toward regulatory engagement. Here, we wish to highlight the involvement of regulatory agencies (EPA, EFSA, OECD) and the development of the guidance documents that emerged from these collaborations. In parallel to agencies embracing the technological developments, the concepts for tackling DNT developed further with this discussion, moving more toward implementation and regulatory use (Bal-Price et al., 2018b).

8.1 Conceptual workshops further advancing DNT testing

The International STakeholder NETwork (ISTNET), which took place in 2014 in Zurich, Switzerland, helped to define how to bring all the technical developments into a regulatory context. The workshop was sponsored by EPAA (The European Partnership for Alternatives to Animal Testing), CAAT, and SCAHT (Swiss Centre for Applied Human Toxicology).

The ISTNET meeting (Croflon et al., 2014; Bal-Price et al., 2015b) focused on the concept of AOPs as a promising tool to promote the development of test systems aligned with regulatory needs. AOPs were considered crucial for assembling predictive ITS for DNT. A stepwise approach to AOP-based DNT testing was outlined, starting with incomplete AOPs for compound grouping and focusing on KEs of neurodevelopment. The next steps included applying the AOP concept to regulatory DNT testing, using AOP interactions for economic development of screening assays, and transitioning from qualitative descriptions to quantitative network modeling.

The report also highlighted the importance of communication and discussions between stakeholders – regulators, industry, and academia – to define a regulatory need-driven roadmap for an ITS for DNT. It acknowledged the limitations of current animal-based test methods, advocating for in vitro/in silico modeling approaches to provide value-added data for regulatory purposes. These approaches were expected to reduce animal testing, lower costs, and increase testing efficiency by using HTS to estimate environmental hazards to human health.

The ISTNET meeting concluded that alternative approaches such as in vitro test methods, quantitative structure-activity relationships (QSARs), read-across, and the application of the concepts of KNPD, AOP, and toxicity endophenotypes could meet regulatory requirements for DNT testing. The main focus of the meeting was increasing the use of alternative data sources in DNT risk assessment and risk management decisions.

In summary, the article by Bal-Price et al. (2015b) advanced the concepts of DNT testing by promoting the use of AOPs to guide the development of ITSs, advocating for the integration of in vitro and in silico methods into regulatory frameworks, and emphasizing the need for a collaborative approach among stakeholders to address the challenges in DNT testing.

The workshop played a pivotal role in integrating EFSA and the OECD into the development of animal-free systemic toxicity testing strategies, including DNT testing. These discussions laid the groundwork for future activities aimed at integrating AOP-based ITS into regulatory decision-making processes for chemical safety assessment. The contributions of EFSA and OECD to these activities have been significant, providing scientific expertise and regulatory perspectives that are essential for the acceptance and implementation of animal-free testing strategies. Their activities have helped to advance the field of DNT testing by promoting the development of more predictive, efficient, and ethically responsible testing strategies that better reflect human health outcomes: EFSA and the OECD have been actively involved in advancing the field of DNT testing, such as the development of non-animal test methods and IATA for DNT. The OECD has coordinated international efforts to enhance DNT testing, acknowledging the limited historical use of in vivo DNT test guidelines. Workshops over the past decade have led to a consensus among stakeholders on the need for a DNT testing battery based on in vitro endpoints and alternative species assays. The OECD initiated specific activities, including collating available DNT in vitro methods, forming a DNT testing battery, generating a reference set of chemicals for testing, and developing an OECD guidance document including IATA case studies (Sachana et al., 2019). These case studies aimed to assess the applicability of the DNT IVB (Blum et al., 2023) in the regulatory risk assessment of pesticides. The approach included systematic literature reviews, expert knowledge elicitation, and Bayesian network analysis to integrate evidence within the AOP framework.

The OECD/EFSA workshop on DNT in October 2016 (Fritsche et al., 2017) made significant progress towards advancing the use of alternative, non-animal test methods for regulatory purposes. Key points of consensus reached by the diverse group of international stakeholders included:

1. There is an urgent need for a DNT testing strategy using in vitro methods and alternative models to begin screening and prioritizing the large number of untested chemicals for their potential effects on the developing nervous system.
2. A battery of currently available in vitro DNT assays, based on critical neurodevelopmental processes, is ready for use now for screening and prioritization purposes. Further development and standardization of the testing battery can enable its use for hazard assessment and to support risk management decisions in the future.
3. A roadmap should be established to define procedures and milestones for implementing this new approach to DNT testing. Priorities include chemical testing to build confidence in the testing battery, establishing performance standards, and developing an OECD guidance document on an integrated approach to DNT testing and assessment.

The OECD project was summarized by Sachana et al. (2019, 2021a). Key steps include: 1) Generating a reference set of chemicals to test a battery of in vitro DNT assays spanning key neurode-
velopmental processes; 2) selecting the battery of in vitro DNT assays based on readiness criteria; 3) testing the reference chemicals in the battery to build confidence in the alternative approaches; 4) developing IATA case studies using DNT in vitro battery data for different regulatory applications; and 5) drafting an OECD guidance document on the use of alternative DNT testing methods within an IATA framework.

The review by Schmidt et al. (2016) presents a comprehensive overview of the progress and technical possibilities in in vitro neurotoxicity and DNT screening, emphasizing the urgent need for alternative testing methods due to challenges in extrapolating animal data to humans and the limited capacity of animal testing to cover all substances requiring evaluation. The paper discusses various cellular platforms used in neurotoxicity testing, ranging from animal and human cell lines to advanced human hiPSCs and organ-on-a-chip models, highlighting their respective advantages and limitations. It covers common endpoints of neurotoxicity testing, including cell viability, neurite outgrowth, synaptic function, and electrophysiological properties, among others. Furthermore, it explores analytical methods such as high-content imaging and electrophysiological screens to assess these endpoints. The review underscores the importance of integrating multiple cellular models and endpoints to capture the complexity of neurotoxic effects and advocates for the development of ITS and multi-omics approaches to improve prediction models for assessing chemical hazard potential. This work contributes significantly to the field by outlining the current state, challenges, and future directions for in vitro neurotoxicity and DNT screening, promoting the shift towards more human-relevant, efficient, and ethical testing methods.

The workshop and resulting paper by Aschner et al. (2017) addressed the critical gap in information regarding the DNT hazard posed by industrial and environmental chemicals. Recognizing that future testing approaches will likely involve a battery of alternative and complementary tests, the paper focused on the first generation of alternative DNT tests that target fundamental neurodevelopmental processes, such as neuronal differentiation, precursor cell migration, or neuronal network formation. These processes are crucial as they capture toxicants with diverse targets and modes of action and can be linked to toxicity endophenotypes, which are alterations in neural connectivity leading to neurofunctional deficits in humans. The workshop that led to the review defined criteria for selecting positive/negative controls, prepared recommendations for their use, and initiated the setup of a directory of reference chemicals. Over 50 endpoint-specific control compounds were identified for initial technical optimization of tests, and an additional set of 33 chemicals, considered direct DNT toxicants, was proposed for further test development.

We next assessed test readiness in a subsequent workshop and developed a list of reference compounds: Bal-Price et al. (2018a) presents a major conceptual advance in providing a framework to evaluate the readiness of NAMs for DNT testing for regulatory purposes. A set of readiness criteria and a scoring system to assess the readiness of individual DNT in vitro assays as well as the overall DNT in vitro battery (DNT IVB) for various regulatory applications were put forward. This allows a quantitative assessment of the status of assay development and what further work is needed to increase regulatory confidence in these alternative methods. The approach outlined provides a role model for evaluating alternative methods in other fields of toxicology. Key elements include: 1) defining readiness criteria based on the regulatory need and context of use, 2) quantitative scoring of assay readiness based on multiple defined criteria, 3) evaluating the battery of assays as a whole in terms of biological coverage, predictive performance, and overall readiness for different regulatory uses, and 4) using case studies to demonstrate how the alternative methods can be applied in an IATA. This framework allows a transparent and objective evaluation of NAMs to facilitate their regulatory acceptance and use. While more work is needed to fully validate the DNT IVB (Juberg et al., 2023), the strategy presented was an important step forward in advancing the use of alternative methods for safety testing.

The US EPA held a public meeting of the FIFRA Scientific Advisory Panel on September 15-17, 2020, to review the use of NAMs to derive extrapolation factors and evaluate DNT for human health risk assessment13. Key points included14.

1. The Panel reviewed several in vitro NAM assays developed by EPA and European researchers to evaluate important neurodevelopmental processes that may be disrupted by chemical exposure. While finding strengths in the assays, the Panel noted limitations such as the lack of important cell types, absence of functional/mechanistic assessments, and difficulties extrapolating to in vivo effects.

2. The Panel considered the use of in vitro-to-in vivo extrapolation (IVIVE) approaches to compare NAM assay results to doses causing acetylcholinesterase inhibition in animals for organophosphate pesticides. They found the IVIVE approach reasonable but raised concerns about model assumptions and performance.

3. The Panel reviewed analyses deriving interspecies and intraspecies extrapolation factors from in vitro acetylcholinesterase inhibition data in rat and human tissues. Limitations were noted in the analysis methods, sample representativeness, and sample sizes to characterize human variability.

4. Overall, the Panel saw value in the NAM approaches while providing numerous recommendations to address limitations and uncertainties in using the data for human health risk assessment.

This evaluation represented an important starting point for the regulatory implementation of in vitro DNT approaches for the most important use case of agrochemicals. The specific concerns14 represented an important agenda for future developments:

- **a. The absence of hormonal factors (sex hormones, thyroid, stress hormones)**
- **b. The influence of neurotransmitter signaling**

---

c. The influence of chemical-induced systemic changes (e.g., inflammation, oxygen levels and distribution)
d. The influence of maternal factors (maternal infection, hormonal, organ system dysfunction, placenta integrity)

In addition, it was stated that the in vitro assays:
a. Will be limited in their ability to detect adaptive or compensatory processes.
b. Do not account for critical cell-cell interactions required during neurodevelopment.
c. Have difficulty distinguishing between neuroactive and neurotoxic compounds.
d. Do not reflect human genetic diversity when using human cell lines from one human.”

8.2 The OECD guidelines and the guidance document
An OECD expert group for in vitro DNT testing was created in 2017. The outcome of this was the proposal to develop, in essence, an IATA. The OECD calls it an IATA – we recently called it invitrois – in vitro, in vivo, and in silico combined (Caloni et al., 2022).

An important contribution was a report commissioned by EFSA: Crofton and Mundy (2021) provide guidance on the interpretation and use of data generated from assays currently included in the DNT IVB for regulatory purposes. Their report has the following key points:

1. The DNT IVB includes a battery of in vitro assays that cover key neurodevelopmental processes such as proliferation, migration, differentiation, neurite outgrowth, and neural network formation.
2. These assays are considered key events at the cellular level that are plausibly related to modes of action of developmental neurotoxicants in vivo.
3. When interpreting data from individual assays, it’s important to evaluate the biological relevance of the test system, assay quality and reproducibility, testing with a training set of chemicals, data analysis methods, and use of a decision model to classify chemicals.
4. To evaluate the DNT IVB as a whole, factors to consider include the predictive power compared to in vivo data, consistency of results across the battery, comparison of potency to other in vitro endpoints, and mapping to AOPs.
5. Use of DNT IVB data should be guided by the consistency of the in vitro data, biological plausibility, incorporation of in vitro to in vivo extrapolation models, and consideration of uncertainties in the context of the regulatory need.
6. While not a full replacement for animal studies, DNT IVB data is already being used to inform screening, prioritization, and weight of evidence approaches in different regulatory applications. Further work is needed to develop standardized reference chemical lists, data analysis pipelines, and tiered testing strategies.

8.3 From screening hits to potential DNT toxicants
Screening takes an increasingly larger role in DNT, driven, e.g., by large screening and prioritization programs of the US EPA or the NTP/NIEHS, and by the availability of more, better, and more robust high-throughput assays in cells, organoids, and simple organisms. Screening is a scientific discipline developed in drug discovery, and a lot of experience has been collected that was incorporated into a screening culture and best practices within the discipline of pharmacology. In toxicology, the technology has been adopted and adapted, but the culture is lagging behind. Often, the distinction between a screen hit and a potential toxicant is not clear and needs further clarification for efficient use of this approach for DNT testing. The pharmacological terminology is as follows: screens produce “preliminary positives” in the respective assay. After a confirmation under more controlled conditions, they can be called “confirmed positives”. They need to be filtered, e.g., for technical artifacts, or secondary effects due to cytotoxicity. In many pharmacological screens this eliminates 90% of the screen, and only few “true positives” remain. These survivors are usually further tested in secondary assays. If they survive this step, then one can talk of potentially interesting assay hits and one can formulate the hypothesis that these are potential toxicants (see Fig. 3).

This hypothesis then needs to be further evaluated to call compounds “relevant toxicological hits”. The evaluation is a process that either disproves the hypothesis or gathers evidence that increases the plausibility of the hypothesis to be true. Important factors in this are similarity to other well-known toxicants or a known mechanism that feeds into an AOP, in addition to toxicokinetic behavior likely to lead to target occupancy at relevant exposures. In the absence of a strong structural similarity argument or mechanistic plausibility, the relevance is often hard to evaluate and requires extremely relevant and predictive DNT assays. Exemplary publications of this process are based on a neural crest migration assay and polychlorinated biphenyls as preliminary hits (Nyyfeler et al., 2018) or a neurite outgrowth assay (Krug et al., 2013b) and berberine as hit compound (Suciu et al., 2023b). A very recent example is a screen of 1800 ToxCast compounds for toxicity to developing oligodendrocytes. The extensive hit characterization and toxicological follow-up was necessary to profile, e.g., quaternary amine-based amphiphilic compounds as a new class of potential DNT toxicants (Cohn et al., 2024).

9 Challenges and future directions
One of the primary challenges is the need for further validation of alternative DNT methods. While a large number of in vitro and in silico models have been developed and shown to be promising predictors of DNT, most of these methods have not yet undergone the rigorous validation process required for regulatory acceptance. Traditional validation involves demonstrating the reproducibility, predictivity, and relevance of a test method for its intended purpose, typically through a series of inter-laboratory studies using a standardized protocol and a set of diverse reference chemicals. This process can be time-consuming and resource-intensive, requiring collaboration among multiple stakeholders, including test method developers, validation bodies, and regulatory agencies. This concept of validation is outdated, as it was developed for simple endpoints and for methods that by themselves (i.e., one-to-one) predict such endpoints. For methods that are part of a test battery, the principles of relevance and predictivity require a
account the specific challenges and opportunities associated with DNT testing, such as the complexity of neurodevelopmental processes, the importance of considering species differences (Baumann et al., 2016) and exposure timing, and the need for a battery of complementary assays that cover different modes of action.

The recently updated document by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), "Validation, Qualification, and Regulatory Acceptance of New Approach Methodologies" describes a more flexible, fit-for-purpose validation strategy with emphasis on integrating results from multiple NAMs rather than one-to-one replacement. DNT is described as one endpoint example where such a strategy will be re-definition, and the use of reference sets of chemicals is only of limited use. For DNT, the overall number of reference chemicals is too small for a classical relevance approach with a statistical prediction model (Mundy et al., 2015). More importantly, also the few reference chemicals with known DNT effects in humans or animals are mostly not well characterized for their mechanism and their mode of action. They can therefore hardly be used for mechanistic assays or KNDP assays.

To facilitate the validation of alternative DNT methods, there is a need for a clear and consensus-based framework that outlines the key steps and criteria for establishing the scientific and regulatory validity of these methods. This framework should take into account the specific challenges and opportunities associated with DNT testing, such as the complexity of neurodevelopmental processes, the importance of considering species differences (Baumann et al., 2016) and exposure timing, and the need for a battery of complementary assays that cover different modes of action. The recently updated document by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), "Validation, Qualification, and Regulatory Acceptance of New Approach Methodologies" describes a more flexible, fit-for-purpose validation strategy with emphasis on integrating results from multiple NAMs rather than one-to-one replacement. DNT is described as one endpoint example where such a strategy will be

Fig. 3: How to move from screening results to toxicological hits

Screens, if run by high-throughput technology, often generate “preliminary positives” that need to be further qualified and characterized. A first step involves repetition of the test, possibly with a more stringent prediction model, more stringent laboratory procedure controls (e.g., according to the OECD Guidance on Good In Vitro Method Practices (GIVIMP)) and with new, well-controlled compound stock solutions. The true positives (TP) are then obtained after elimination of technical artifacts (e.g., fluorescent compounds, or compounds binding unspecifically to proteins), control for cytotoxicity, and cheminformatics filtering for known pan-assay interfering substances (PAINS). Certainty on the bioactivity of such compounds is then obtained classically in secondary and tertiary assays for similar target pathways or structures. These are ideally orthogonal in the sense that they use other test systems and/or readout technologies.

In a second phase, such “convincing” assay hits are evaluated for their toxicological relevance. Strategy I tests the relevance by establishing the similarity of the assay hit to known, well-characterized toxicants. The definition of similarity relies on (i) structural similarity, (ii) similar ADME properties, and (iii) a similar bioactivity profile/mode-of-action (MoA). Strategy II tests the relevance based on the expected MoA or the activation of a reliable adverse outcome pathway (AOP). Essential issues are (i) whether the concentrations triggering an AOP molecular initiation event (MIE) or key event (KE) are realistically reached, whether it is likely that the triggered AOP also continues to the adverse outcome (AO). A more comprehensive investigation would use quantitative in vitro-to-in vivo (IVIVE) extrapolation models and would consider whether key event relationships (KER) affected by modifiers or counter-regulations.

important and hopefully speed up the process. Efforts to implement such a framework are ongoing, as exemplified by the work of the OECD’s DNT Expert Group and the collaborative workshops organized by CAAT, ICCVAM, ECVAM, and other stakeholders. Possibly, the concept needs to be even more radically revised for some assays, i.e., not relying at all on classical reference chemicals but rather on mechanistic validation concepts (Hartung et al., 2013b).

Another challenge is the integration of new technologies, such as HTS, omics approaches, and computational modeling, into regulatory DNT testing strategies. While these technologies offer the potential for more rapid, efficient, and mechanistically informative testing, their application in a regulatory context raises several issues related to data quality, interpretability, and relevance. For example, HTS assays may generate large amounts of data on the effects of chemicals on specific molecular or cellular endpoints, but the relationship between these effects and AOs in the developing brain may not always be clear. Similarly, omics approaches, such as transcriptomics (Krug et al., 2013a; Rempel et al., 2015) and metabolomics (Modafferi et al., 2021), can provide a wealth of information on the biological pathways and processes perturbed by chemical exposure, but translating this information into actionable regulatory decisions may require additional steps, such as the development of quantitative AOPs and the establishment of thresholds for adverse effects.

To address these challenges, there is a need for ongoing dialogue and collaboration between the developers of new technologies, the regulatory agencies responsible for implementing DNT testing requirements, and the broader scientific community. This dialogue should aim to identify the key data needs and decision-making criteria for regulatory DNT testing, and to develop a roadmap for the progressive integration of alternative methods and new technologies into regulatory frameworks. This may involve the development of tiered testing strategies that combine in vitro, in silico, and in vivo approaches in a way that balances the need for efficiency and predictivity with the requirements for regulatory certainty and human health protection.

Looking ahead, there are also significant opportunities for improving the predictivity and efficiency of DNT testing through advances in data integration and ML. One promising approach is the development of ITS that combine multiple types of data, such as in vitro assay results, physicochemical properties, and in silico predictions, to provide a more comprehensive and reliable assessment of DNT potential. By leveraging the strengths of different data streams and taking into account the uncertainties and limitations associated with each, ITS can potentially provide a more accurate and efficient means of prioritizing chemicals for further testing and informing regulatory decisions.

The development of ITS for DNT testing can be facilitated by advances in ML and AI (Kleinstreuer and Hartung, 2024), which offer powerful tools for integrating and analyzing large and complex datasets. ML algorithms, such as random forests, support vector machines, and deep neural networks, can be trained on existing DNT data to identify patterns and relationships that may not be apparent from traditional statistical analyses. These models can then be used to predict the DNT potential of new chemicals based on their structural and biological similarity to known toxicants, or to identify the most informative combinations of assays and endpoints for a given regulatory context.

As more data become available from alternative DNT testing methods and other sources, such as high-throughput screening programs and academic research, there will be increasing opportunities to refine and optimize ITS using ML approaches. This could involve the continuous updating of predictive models as new data are generated, the identification of novel biomarkers and endpoints that are more predictive of AOs, and the development of more sophisticated decision-making frameworks that take into account multiple lines of evidence and sources of uncertainty.

To fully realize these opportunities, however, there is a need for greater standardization and harmonization of DNT testing methods and data reporting practices. This includes the development of common ontologies and data formats for describing neurodevelopmental processes and endpoints, the establishment of minimum reporting standards for in vitro and in silico DNT studies, and the creation of centralized and publicly accessible databases for storing and sharing DNT data. Efforts to promote such standardization and harmonization are already underway, as exemplified by the work of the OECD’s Extended Advisory Group on Molecular Screening and Toxicogenomics (EAGMST)16 and the collaborative projects funded by the European Union’s Horizon 2020 research and innovation program.

In conclusion, while significant progress has been made in developing alternative approaches for DNT testing, there remain important challenges to be addressed in terms of method validation, regulatory acceptance, and data integration. At the same time, there are exciting opportunities for leveraging advances in ML and AI to improve the predictivity and efficiency of DNT testing, and to support the development of more effective and scientifically sound regulatory strategies for protecting human health. By working together to address these challenges and opportunities, the scientific and regulatory communities can help to ensure that the promise of alternative DNT testing is fully realized, and that the safety of chemicals is assessed in a way that is both rigorous and efficient and that reflects the best available science.

9.1 Implementation
Implementing advanced testing strategies for DNT in regulatory and research contexts presents several challenges. While ITS offer a promising approach to approximate regulatory information needs more effectively, their adoption is hindered by various factors. One of the main challenges is the complexity of DNT itself, with a multitude of KEs and targets that need to be covered by the building blocks of an ITS. The design of ITS is ideally based on established mechanisms of health effects, such as AOPs, but the small number of studies leading to regulatory decisions in DNT limits this approach.

16 https://www.oecd.org/chemicalsafety/testing/omics.htm
Moreover, the integration of emerging tools for data integration, including Bayesian networks, ML tools, and sensitivity analysis, is essential for the continuous optimization of ITS. However, the practical application of these tools in a regulatory context is still in its infancy, and there is a need for further development and validation of these methods. Additionally, the current regulatory frameworks may not be fully equipped to accommodate the novel approaches that ITS and AOPs represent, requiring updates and adaptations to existing guidelines and practices.

The challenges extend to the quality assurance of non-animal tests, where Good Cell Culture Practices (GCCP) (Pamies et al., 2022) and the concept of “mechanistic validation” (Hartung et al., 2013b) are critical to ensure the reliability of the results. However, establishing these quality standards and gaining widespread acceptance can be a slow and complex process. Furthermore, the interpretation of DNT study results requires a substantial amount of expertise, and there is considerable flexibility in the study design, which introduces potential sources of variability.

In summary, while the strategic development of pathway-based approaches to DNT testing is underway, the implementation of these advanced strategies faces challenges related to the complexity of DNT, the need for further development and validation of data integration tools, the adaptation of regulatory frameworks, and the establishment of quality standards for non-animal tests.

9.2 Vision for the future
The potential of advanced technologies in DNT testing extends far beyond merely replacing animal models; it opens new avenues for understanding the intricacies of neurodevelopmental disorders and toxicology as envisioned in a human exposome project (Hartung, 2023b). The integration of organotypic cultures, stem cells, high-content methods, and computational modeling offers a more nuanced and human-relevant approach to studying the effects of environmental chemicals on neurodevelopment. These technologies enable researchers to explore the cellular and molecular mechanisms underlying DNT, providing insights into the etiology of disorders such as ASD and ADHD. Furthermore, the development of AOPs and ITS based on these technologies facilitates a more systematic and predictive assessment of chemical risks. By harnessing the power of 21st-century technologies, the field of toxicology is poised to make significant strides in protecting public health, reducing reliance on animal testing, and uncovering new knowledge about the complex processes that govern neurodevelopment and its disruption by environmental factors (Smirnova et al., 2014).

We have tremendous problems to find an exposure hypothesis for all these diseases. We also have tremendous problems to find a genetic hypothesis for all of them. The explanation for this may be very simple: gene-environment interaction. One of the biggest limitations of the current test battery is the lack of consideration of population variability as the cells are likely from typically developed donors. Therefore, it would be interesting to include assays with cells from different genetic backgrounds, e.g., from autistic children, to increase the possibility to assess gene interactions with these exposures.

AI, including the use of large language models and predictive toxicology, promises to enhance DNT testing strategies. In ON-TOX, one of the three targets is DNT, which brings a lot of very important elements into the discussion. The first is exploiting large language models in order to extract all of the available information. One of the most important goals is for the data from our DNT community to be made available for extraction. The data need to be brought together, annotated with metadata, so that it is possible to actually feed them into these models. The second part that ONTOX is addressing is essentially using the foundational model of read-across-based structure-activity relationships to enter these additional endpoints. This shall not be done in isolation like a QSAR, where DNT data would be used to predict DNT, but shall include everything we know about the connections among these chemicals. For example, simple addition of some data on androgen receptor activity led to 98% accurate predictions of substances (unpublished); the addition of cancer data from REACH registrations or reproductive toxicity data from REACH for about 1000 substances was 80% correctly predicted. So, there is a lot of promise in integrating DNT data into large foundational models for predictive toxicology. And finally, AI is helping us to understand what is contributing to the correct information. What are the important information sources — the evidence streams? This will help to prioritize assays that can tell us what the value of the in vivo information is and that can distinguish between valuable and less valuable information (Hartung, 2023a). This allows a type of sensitivity analysis, and can even be used to generate economic information, where you can ask whether it is worthwhile to do a certain assay and whether its results will really change the overall assessment because we know whether this assay is of high or low information value. We cannot overestimate at this moment, what integration of knowledge by AI will give us. We may end up accepting it as a copilot and receive its guidance to drive the further development of the DNT battery.

The concept of organoid intelligence (OI) and its potential to advance our understanding of neurodevelopment and toxicology is worth some consideration in this context. We have introduced the idea of simple forms of memory and learning as the ultimate functionality of brain organoids (Smirnova et al., 2023a,b). This combines technologies, i.e., the organoids and AI; essentially we are trying to talk to a brain organoid through AI. This is possible because organoids are electrophysiologically active and receptive. OI, as a type of intelligence-in-a-dish, promises to study toxicants and possible repairs by working on brain organoids from patient material. We are forming a community around this (Morales Pantoya et al., 2023a; Hartung et al., 2023), which includes important ethical discussions (Hartung et al., 2024).

10 Conclusions and outlook
Summarizing the impact of the strategic development of alternatives to animal testing for DNT, this field has benefited tremendously from the collaborative efforts of researchers, regulators, and industry stakeholders. The past two decades have seen a remarkable advance in our scientific understanding of neurodevelopmental processes and the mechanisms by which chemicals can interfere with these processes to cause AOs. This knowledge
has been leveraged to develop a wide range of alternative testing methods, including in vitro cell culture models, non-mammalian animal models, and computational approaches that offer the potential for a more rapid, efficient, and mechanistically informative assessment of DNT potential.

The development of these alternative methods has been driven by a recognition of the limitations of traditional animal-based testing paradigms, which are resource-intensive, time-consuming, and may not always be predictive of human health outcomes. By focusing on the key events and pathways that underlie neurodevelopment and its perturbation by chemicals, alternative methods have the potential to provide a more targeted and relevant assessment of DNT potential, while also reducing the reliance on animal testing. This shift towards a more mechanistically based and human-relevant testing paradigm represents a major advance in the field of DNT testing and reflects a broader trend in toxicology towards the use of alternative methods that are in line with the principles of the 3Rs.

The impact of this strategic development extends beyond scientific and ethical dimensions. By providing more rapid and efficient means of assessing DNT potential, alternative methods have the potential to significantly accelerate the pace of chemical safety assessment and support more effective regulatory decision-making. This is particularly important given the large number of chemicals in commerce that have not been tested for DNT potential and the growing recognition of the importance of protecting the developing brain from environmental chemical exposures. By enabling the prioritization of chemicals for further testing and providing a more comprehensive and reliable assessment of DNT potential, alternative methods can help to ensure that regulatory decisions are based on the best available science and are protective of human health.

The development of alternative methods for DNT testing has been a collaborative effort, involving contributions from researchers in academia, government, and industry, as well as from regulatory agencies and other stakeholders. This collaboration has been essential for ensuring that the methods developed are not only scientifically valid and relevant, but also practical and implementable within a regulatory context. The establishment of networks and forums for the exchange of knowledge and ideas, such as the DNT Workgroup and the series of DNT conferences organized by CAAT and its partners, has been crucial in fostering this collaboration and driving the field forward.

Looking back over the past 18 years, the field of DNT testing has made remarkable progress. In 2006, when the first DNT conference was held, the field was heavily reliant on animal-based testing methods, with few alternatives available. The discussions at that conference focused on the need for more rapid and efficient testing methods, and the potential for in vitro and non-mammalian models to fill this need. Over the subsequent years, a range of alternative methods was developed and evaluated, culminating in the publication of a comprehensive review by Bal-Price et al. (2012) that identified a number of promising in vitro and alternative animal models for DNT testing.

The period from 2014 to the present has seen a rapid acceleration in the development and application of alternative methods for DNT testing, driven in part by advances in stem cell biology, genome editing, high-throughput screening, and computational modeling. The establishment of the OECD DNT Expert Group in 2017 and the publication of the OECD’s “Initial Recommendations on Evaluation of Data from the Developmental Neurotoxicity (DNT) In Vitro Testing Battery” (OECD, 2023) marked important milestones in the regulatory acceptance of alternative methods.

Despite these advances, however, much work remains to be done to fully realize the potential of alternative methods for DNT testing. Further research is needed to optimize and validate these methods, to demonstrate their reliability and relevance for regulatory decision-making, and to integrate them into testing strategies that provide a comprehensive and efficient assessment of DNT potential. Continued collaboration among researchers, regulators, and industry stakeholders will be essential for addressing these challenges and advancing the field towards a more predictive and ethical testing paradigm.

In conclusion, the strategic development of alternatives to animal testing for DNT over the past 18 years represents a major advancement in the field of toxicology, one that has been driven by scientific innovation, regulatory need, and a commitment to more human-relevant and ethical testing approaches. While significant challenges remain, the progress made to date is a testament to the dedication and collaboration of the many stakeholders involved in this effort. By continuing to work together to refine and validate alternative methods for DNT testing, we can move towards a future in which the safety of chemicals is assessed in a way that is both scientifically rigorous and ethically sound, and that truly protects the health of our most vulnerable populations. What we have seen over the last two decades is a scientific revolution. For the first time, an example of a strategic development by a community led to the development of a testing strategy that is completely novel in its approach. It is a testing strategy composed of different components, it is mechanism-based, and it has undergone an international consensus process. These are all criteria we should apply to more areas, especially to animal experimentation. The combination of these disruptive technologies will lead us to far more than just replacing animal tests. We can do things we did not dream of when we started out. The challenge now is implementation!

References


Baumann, J., Gassmann, K., Masjosthusmann, S. et al. (2016). Comparative human and rat neurospheres reveal species differences in chemical effects on neurodevelopmental key events. *Arch Toxicol* 90, 1415-1427. doi:10.1007/s00204-015-1568-8


Kilpatrick, S., Irwin, C. and Singh, K. K. (2023). Human pluripotent stem cell (hPSC) and organoid models of autism: Opportunities and limitations. Transl Psychiatry 13, 217. doi:10.1038/s41398-023-02510-6


Lee, J. and Freeman, J. L. (2014). Zebrafish as a model for developmental neurotoxicity assessment: The application of the zebrafish in defining the effects of arsenic, methylmercury, or lead on early neurodevelopment. Toxics 2, 464-495. doi:10.3390/toxics20300464


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


Morales Pantoja, I. E., Smirnova, L., Muorri, A. R. et al. (2023a). First organoid intelligence (OI) workshop to form an OI community. Front Artif Intell 6, 1116870. doi:10.3389/frai.2023.1116870


Smirnova, L., Caffo, B. S., Gracias, D. H. et al. (2023a). Organoid intelligence (OI): The new frontier in biocomputing and...

Smirnova, L., Morales Pantoja, I. E. and Hartung, T. (2023b). Organo
doid Intelligence (OI) – The ultimate functionality of a brain microphysiological system. *ALTEX* 40, 191-203. doi:10.14573/altex.2303261

Smirnova, L. and Hartung, T. (2024). The promise and potential of
brain organoids. *Adv Healthcare Mater*, e2302745. Online ahead of
print. doi:10.1002/adhm.202302745


Suciu, I., Pamics, D., Peruzzo, R. et al. (2023a). GxE interactions as
a basis for toxicological uncertainty. *Arch Toxicol* 97, 2035-2049.
doi:10.1007/s00204-023-03500-9

Suciu, I., Delp, J., Gutbier, S. et al. (2023b). Definition of the neu
rotoxicity-associated metabolic signature triggered by berber
ine and other respiratory chain inhibitors. *Antioxidants* 13, 49. 
doi:10.3390/antiox13010049


van Thriel, C., Westerink, R. H., Beste, C. et al. (2012). Translating
neurobehavioural endpoints of developmental neurotoxicity tests into in vitro assays and readouts. *Neurotoxicology* 33, 911-924. 

van Vliet, E., Stoppini, L., Balestrino, M. et al. (2007). Electrophy-
siological recording of re-aggregating brain cell cultures on mul-
ti-electrode arrays to detect acute neurotoxic effects. *Neurotoxi-
cology* 28, 1136-1146. doi:10.1016/j.neuro.2007.06.004

van Vliet, E., Morath, S., Linge, J. et al. (2008). A novel in vitro meta-
tabolomics approach for neurotoxicity testing, proof of principle for methyl mercury chloride and caffeine. *Neurotoxicology* 29,
1-12. doi:10.1016/j.neuro.2007.09.007

van Vliet, E., Daneshian, M., Beilmann, M. et al. (2014). Current
approaches and future role of high content imaging in safety sci-

Villa, C., Combi, R., Conconi, D. et al. (2021). Patient-derived in-
duced pluripotent stem cells (iPSCs) and cerebral organoids for
drug screening and development in autism spectrum dis-
doi:10.3399/pharmaceutics1302080

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

microdeletion syndrome establishes CRLF3 as a critical regu-
celrep.2021.109315

Willett, C. E. (2018). The use of adverse outcome pathways (AOPs)
to support chemical safety decisions within the context of inte-
grated approaches to testing and assessment (IATA). In H. Ko-
jima, T. Seidle and H. Spielmann (eds), *Alternatives to Animal
2447-5_11

Yamada, S., Hirano, Y., Kurosawa, O. et al. (2019). Evaluation of
developmental neurotoxicity using neural differentiation poten-
ty in human iPSC cells. *Proc Ann Meet Jpn Pharmacol Soc* 92, 
https://www.jstage.jst.go.jp/article/jpsslupp/92/0/92_3-127/_/
pdf

jacs.1c07312

Zablotsky, B., Black, L. I., Maenner, M. J. et al. (2019). Preval-

Zhang, L., Li, M., Zhang, D. et al. (2023). Developmental neuro-
toxicity (DNT) QSAR combination prediction model establish-
ment and structural characteristics interpretation. *Toxicol Res* 13, 
tf1d16. doi:10.1003/toxres/tf1d116

Zhong, X., Harris, G., Smirnova, L. et al. (2020). Antidepressant
paroxetine exerts developmental neurotoxicity in an iPSC-

**Conflict of interest**

T. H. and H. T. H. are named inventors on a patent by Johns
Hopkins University on the production of mini-brains (also
called BrainSpheres), which is licensed to A xoSim, New Orleans,
LA, USA. T. H. and L. S. are consultants for A xoSim, New
Orleans, and T. H. is also a consultant for the American Type
Culture Collection (ATCC) and was until recently for AstraZen-
eca on advanced cell culture methods.

**Data availability**

No novel data was produced for this manuscript.

**Acknowledgements**

The authors acknowledge funding from the US Environmental Protection Agency (EPA) EPA-STAR grant (R83950501), the
BMBF (161I0243B, 016LW0146K), the Combining advances in Genomics and Environmental science to accelerate Action-
able Research in ASD (GEARS) Network (R01ES034554); Ladd-Acosta, Volk), and the Land Baden-Württemberg (NAM-
ACCEPt). Funding was also received from EFSA, the European
Union’s Horizon 2020 research and innovation program under
grant agreements No. 964537 (RISK-HUNT3R), No. 964518
ToxFree), No. 963845 (ONTOX), No. 101057014 (PARC), and a
Johns Hopkins Discovery award and a Johns Hopkins SURPASS
award.