t^4 Workshop Report*

Recommendation on Test Readiness Criteria for New Approach Methods in Toxicology: Exemplified for Developmental Neurotoxicity

Anna Bal-Price[1], *Helena T. Hogberg*[2], *Kevin M. Crofton*[3], *Mardas Daneshian*[4], *Rex E. FitzGerald*[5], *Ellen Fritsche*[6], *Tuula Heinonen*[7], *Susanne Hougaard Bennekou*[8], *Stefanie Klima*[9], *Aldert H. Piersma*[10], *Magdalini Sachana*[11], *Timothy J. Shafer*[3], *Andrea Terron*[12], *Florianne Monnet-Tschudi*[5,13], *Barbara Viviani*[14], *Tanja Waldmann*[9], *Remco H. S. Westerink*[15], *Martin F. Wilks*[5], *Hilda Witters*[16], *Marie-Gabrielle Zurich*[5,13] *and Marcel Leist*[4,9]

[1]European Commission, Joint Research Centre (EC JRC), Ispra, Italy; [2]Center for Alternatives to Animal Testing (CAAT), Johns Hopkins University, Baltimore, MD, USA; [3]National Center for Computational Toxicology, US EPA, RTP, NC, USA; [4]Center for Alternatives to Animal Testing, CAAT-Europe, University of Konstanz, Konstanz, Germany; [5]Swiss Centre for Human Applied Toxicology, SCAHT, University of Basle, Basle, Switzerland; [6]IUF – Leibniz Research Institute for Environmental Medicine & Heinrich-Heine-University, Düsseldorf, Germany; [7]Finnish Centre for Alternative Methods (FICAM), University of Tampere, Tampere, Finland; [8]Danish Environmental Protection Agency, Copenhagen, Denmark; [9]In vitro Toxicology and Biomedicine, Dept inaugurated by the Doerenkamp-Zbinden Foundation, University of Konstanz, Konstanz, Germany; [10]RIVM, National Institute for Public Health and the Environment, Bilthoven, and Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands; [11]Organisation for Economic Co-operation and Development (OECD), Paris, France; [12]European Food Safety Authority, Parma, Italy; [13]Department of Physiology, University of Lausanne, Lausanne, Switzerland; [14]Department of Pharmacological and Biomolecular Sciences, University of Milan, Milan, Italy; [15]Neurotoxicology Research Group, Institute for Risk Assessment Sciences (IRAS), Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands; [16]VITO, Flemish Institute for Technological Research, Unit Environmental Risk and Health, Mol, Belgium

Abstract

Multiple non-animal-based test methods have never been formally validated. In order to use such new approach methods (NAMs) in a regulatory context, criteria to define their readiness are necessary. The field of developmental neurotoxicity (DNT) testing is used to exemplify the application of readiness criteria. The costs and number of untested chemicals are overwhelming for *in vivo* DNT testing. Thus, there is a need for inexpensive, high-throughput NAMs to obtain initial information on potential hazards, and to allow prioritization for further testing. A background on the regulatory and scientific status of DNT testing is provided showing different types of test readiness levels, depending on the intended use of data from NAMs. Readiness criteria, compiled during a stakeholder workshop that united scientists from academia, industry and regulatory authorities, are presented. An important step beyond the listing of criteria was the suggestion of a preliminary scoring scheme. On this basis a (semi)-quantitative analysis process was assembled on test readiness of 17 NAMs with respect to various uses (e.g., prioritization/screening, risk assessment). The scoring results suggest that several assays are currently at high readiness levels. Therefore, suggestions are made on how DNT NAMs may be assembled into an integrated approach to testing and assessment (IATA). In parallel, the testing state in these assays was compiled for more than 1000 compounds. Finally, a vision is presented on how further NAM development may be guided by knowledge of signaling pathways necessary for brain development, DNT pathophysiology, and relevant adverse outcome pathways (AOP).

---

*A report of t^4 – the transatlantic think tank for toxicology, a collaboration of the toxicologically oriented chairs in Baltimore, Konstanz and Utrecht sponsored by the Doerenkamp Zbinden Foundation. The views expressed in this article are those of the contributing authors and do not necessarily reflect those of their institution of employment.

Disclaimer: The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency (US-EPA). Mention of trade names or commercial products does not constitute endorsement or recommendation for use. The opinions expressed and arguments employed herein are those of the authors and do not necessarily reflect the official views of the OECD or of the governments of its member countries.

Received December 8, 2017; Accepted January 29, 2018;
Epub February 21, 2018; © The Authors, 2018.

ALTEX 35(3), 306-352. doi:10.14573/altex.1712081

Correspondence: Tanja Waldmann, PhD, Universitätsstrasse 10,
Postbox M657, University of Konstanz, 78457 Konstanz, Germany
(tanja.waldmann@uni-konstanz.de)

1 Introduction

1.1 Objectives of the meeting and follow-up activities

A recent OECD/European Food Safety Authority (EFSA) workshop on the use of non-animal test methods for regulatory purposes in the area of developmental neurotoxicity (DNT) proposed to develop a standardized *in vitro* testing battery that could be used to generate data on the toxic effects of chemicals on the developing nervous system. It was recognized that there is an urgent need for a new alternative testing strategy that supports regulatory decisions with a focus on two specific aims: the first is to use existing alternative test methods to support screening and prioritization of chemicals for future testing, the second aim is to generate data that aid in guiding risk management decisions. The workshop concluded that the next task was to establish performance standards and develop a guidance document for an *in vitro* DNT testing battery (Fritsche et al., 2017a).

The International Stakeholder Network (ISTNET) on DNT testing is a collaborative effort of groups from academia, industry and regulatory bodies that aims to align the development of alternative (non-animal) testing methods with the needs of regulatory decision-making. A first meeting in Zurich in January 2014 explored the potential of applying the adverse outcome pathway (AOP) framework to promote test system development according to regulatory needs, and to assemble predictive integrated testing strategies (ITS) for DNT (Bal-Price et al., 2015a).

With the outcome of the OECD/EFSA workshop in mind, a second ISTNET Workshop took place in Konstanz in January 2017, focused on practical aspects of such pathway-based testing, and in particular on performance standards that should be applied to alternative DNT tests. The immediate objectives of the meeting and its follow-up activities were:

1. Define criteria for evaluation for readiness of a given test method.
2. Evaluate to what extent these criteria are fulfilled.

For the second objective, proof-of-principle examples are given here on how an evaluation may be performed; as information only historical, published information was used. Therefore, mid-term objectives were defined to continue this process:

A. Establish a standardized evaluation system for assay readiness.
B. Define a list of suitable test methods based on these criteria.
C. Establish criteria for a battery of tests for use in a DNT IATA based on readiness scores.
D. Build an IATA for initial chemical screening and prioritization.

The long-term goal is to define a battery of alternative tests based on developmental ontologies (in contrast to the mid-term goal of performance-based test definition). Such a battery would include the relevant tests for all biological pathways, processes and domains implicated in DNT.

1.2 Background on the use of existing *in vivo* test methods: Why are alternatives needed?

At present, there is no regulatory requirement for pesticides or other chemicals to be tested for DNT prior to registration. Instead, DNT testing can be triggered based on observed neurotoxic effects in repeat-dose *in vivo* animal testing, a known neurotoxic mode of action, or a structure-activity alert, in Europe for pesticides, biocides and chemicals and in the US for pesticides. In these triggered cases, DNT testing is performed as an *in vivo* higher-tier test as there are no regulatory accepted alternative methods for this purpose. There are two regulatory guidelines for DNT testing, both in rodents: OECD TG 426 (Developmental Neurotoxicity Study), which is an update of the 1998 US EPA DNT Guideline, and OECD TG 443 (Extended One-Generation. Reproductive Toxicity Study, DNT cohort). Both require neurobehavioral evaluation of cognitive, sensory and motor function, accompanied by histopathological and morphometric evaluation of the brain, but they do not provide detailed guidance on the use of specific behavioral tests, leaving flexibility in the study design and in the interpretation of the results obtained. Moreover, TG 426 and TG 443 present a number of challenges and limitations (Claudio et al., 2000; Crofton et al., 2004, 2011; Tsuji and Crofton, 2012; Smirnova et al., 2014), including:

1. They are time- and resource-consuming, low-throughput assays.
2. A large number of animals is required.
3. Differences in techniques and measures, especially for behavioral endpoints, can make it difficult to compare data between studies.
4. Implementation of the DNT guideline methods in contract laboratories has resulted in datasets with high variability and low reproducibility, even for positive controls.
5. Measured pathological and behavioral endpoints provide no mechanistic understanding of the underlying effects.
6. The currently required tests do not capture important complex endpoints of relevance for humans, for example higher cognitive functions.
7. The predictivity for protection of the human brain is based on a very limited number of chemicals, and rodent studies may not reflect some toxicodynamic processes in humans, leading in some cases to uncertainty about the relevance of animal outcomes for human DNT.

In reality, TG 426 and 443 are seldom conducted. Studies are currently available for only a relatively limited number of substances (about 120) (van Thriel et al., 2012; Kadereit et al., 2012; Crofton et al., 2012). Therefore, the urgent aim is to develop alternative test methods as part of a test strategy that

---

Abbreviations

ADME, absorption, distribution, metabolism and excretion; AKT, protein kinase B; AOP, adverse outcome pathway; BDNF, brain derived neurotrophic factor; BMP, bone morphogenic protein; BrdU, bromodeoxyuridine; CREB, cAMP responsive element binding protein; DA, defined approach; DNT, developmental toxicity; EGF, epidermal growth factor; ERK, extracellular signal-regulated kinase; ESC, embryonic stem cells; GFAP, glial fibrillary acidic protein; GW, gestation week; HCIA, high content image analysis; HBRV, health-based reference value; IATA, integrated approach to testing and assessment; ITS, integrated testing strategy; KE, key event; LTP, long-term potentiation; MIE, molecular initiating event; NAM, new approach method; NPC, neural progenitor cell; PI3K, phosphatidylinositol-3 kinase; PSC, pluripotent stem cells; QSAR, quantitative structure activity relationship; RG, radial glia; RTK, receptor tyrosine kinase; TG, test guideline.

at least can identify DNT alerts and guide prioritization at a lower-tier level.

A recent review, focused on pesticide active substances, was presented at the DNT OECD/EFSA workshop in Brussels (Fritsche et al., 2017a) by the German Federal Institute for Risk Assessment (BfR). To date, DNT studies have been conducted on only 35 of the 485 pesticide active substances currently approved in the EU. Of these 35, 19 displayed positive *in vivo* evidence of DNT. It should be noted that a large proportion of these 485 pesticide active substances were classified as adult neurotoxicants (Grandjean and Landrigan, 2006). It is unknown whether a similarly high rate (> 50%) of positive DNT results would be seen for other classes of chemicals that are not enriched in neurotoxicants. Moreover, the DNT testing led to health-based guideline reference values for only 2 of these 19 positive compounds.

An alternative analysis of DNT studies by the USEPA in 2010 demonstrated that of 72 DNT studies, 15 were used to determine the point of departure for one or more risk assessment scenarios, and an additional 13 were determined to have the potential for use as a point of departure for future risk assessments (Raffaele et al., 2010). These assessments are limited to a small number of chemicals that in no way represents the known chemical space of environmental chemicals (Richard et al., 2016). Thus, to clarify the need for DNT testing for regulatory purposes, experimental evidence on the potential for DNT hazard for many more chemicals is required. However, for this purpose the tests need to be more time- and cost-effective.

The sensitivity of the currently used *in vivo* DNT test has been questioned (Claudio et al., 2000; Vorhees and Makris, 2015). Some of the issues may be due to toxicodynamics, others may be explained by different toxicokinetics among species (metabolic activity or placental transfer in animals compared to humans as exemplified earlier (reviewed in Aschner et al., 2017)). The issue of sensitivity is, for example, evident regarding the predictivity value of the rat DNT assay for the evaluation of chemicals acting on the hypothalamic-thyroid axis. Despite the human evidence linking developmental hypothyroxinemia with changes in brain development in children (Haddow et al., 1999; Henrichs et al., 2010), several DNT studies investigating rodent offspring from hypothyroid/hypothyroxinemic dams have shown that adverse behavioral outcomes were not always present (York et al., 2005). Although multiple explanations may clarify this issue and should be taken into account (e.g., severity of the effect in the dams, limited milk transfer of the compound, neurobehavioral assessment methods not suited for the detection of subtle effects in the brain, presence of compensatory mechanisms), it is evident that design, conduct and interpretation of *in vivo* DNT studies are complicated. Species differences of developing brain cells in response to thyroid hormones have recently been reported also on the level of pharmacodynamics (Dach et al., 2017).

Due to these issues, the US EPA Office of Pesticide Programs (OPP) suggested to include alternative approaches in the testing paradigm to improve DNT hazard identification in the context of analyzing DNT *in vivo* studies for 72 pesticide active substances (Raffaele et al., 2010).

Another reason regulatory bodies and authorities support the development of alternative medium- to high-throughput assays is the need for testing large numbers of chemicals for their DNT potential (Crofton et al., 2012; EFSA, 2013; Bal-Price et al., 2015a; Fritsche et al., 2017a).

1.3 Making alternative methodologies for DNT testing acceptable for regulatory purposes

Reliability and human relevance are the two critical requirements that have to be addressed for regulatory acceptance of alternative test methods. The OECD Adverse Outcome Pathway (AOP) framework (OECD, 2013; Ankley et al., 2010; Bal-Price et al., 2015b; Leist et al., 2017; Terron et al., 2018) is useful in defining the human relevance of data from individual test systems as it takes all available data, including human epidemiology and human *in vitro* data, into consideration. Moreover, it allows development of quantitation and threshold models on the basis of quantified key events (KE) in an established AOP.

The assessment of the readiness and reliability of alternative DNT methods for regulatory purposes is currently lagging behind the extremely rapid development of new technologies (e.g., induced pluripotent stem cells, 3D cell co-cultures and organoids, high-content omics measurements, bioinformatics tools, etc.) (Leist et al., 2008a, 2014; Marx et al., 2016; Rovida et al., 2015; Smirnova et al., 2016). This is unfortunate, since more guidance on how to ensure reliability of the available and new *in vitro* DNT assays would help researchers in designing, conducting, and reporting studies. It would also encourage regulators to take NAMs into account.

Therefore, the major focus of this workshop report is to provide a set of readiness criteria that potentially could be acceptable to both regulators and test developers. Moreover, examples are given on how a readiness evaluation of existing *in vitro* DNT assays could be applied to various regulatory applications. Preliminary scoring by workshop participants of over a dozen methods demonstrates that the field of DNT-NAM is ready to support some regulatory decisions. The readiness criteria will also be helpful to harmonize development of new *in vitro* tests and to ensure their reliability and relevance.

In addition to data reliability and relevance evaluation, both researchers and regulators will need guidance on data integration from a battery of alternative DNT assays (Behl et al., 2015) in the form of ITS and defined approaches (DA) (OECD, 2016c). This enables a tiered approach, spanning the spectrum from hazard identification/characterization as an input to quantitative risk assessment, aiding the application of human health-related decisions based on data coming from alternative approaches.

Outstanding regulatory challenges for accepting alternative DNT test data are similar for most alternative methods and include uncertainty due to genetic background, cell type and topography, life-stage, and exposure temporality in dose-response modeling (Hartung et al., 2017a,b). Some of these issues are addressed in the AOP framework (Bal-Price and Meek, 2017; Leist et al., 2017; Terron et al., 2018), which will thus help in their resolution.

Current hazard identification processes based on *in vitro* tests accepted by regulatory agencies rely on molecular and cellular KEs within AOPs. Here, the most prominent example is the application of a testing battery based on KEs identified in the AOP

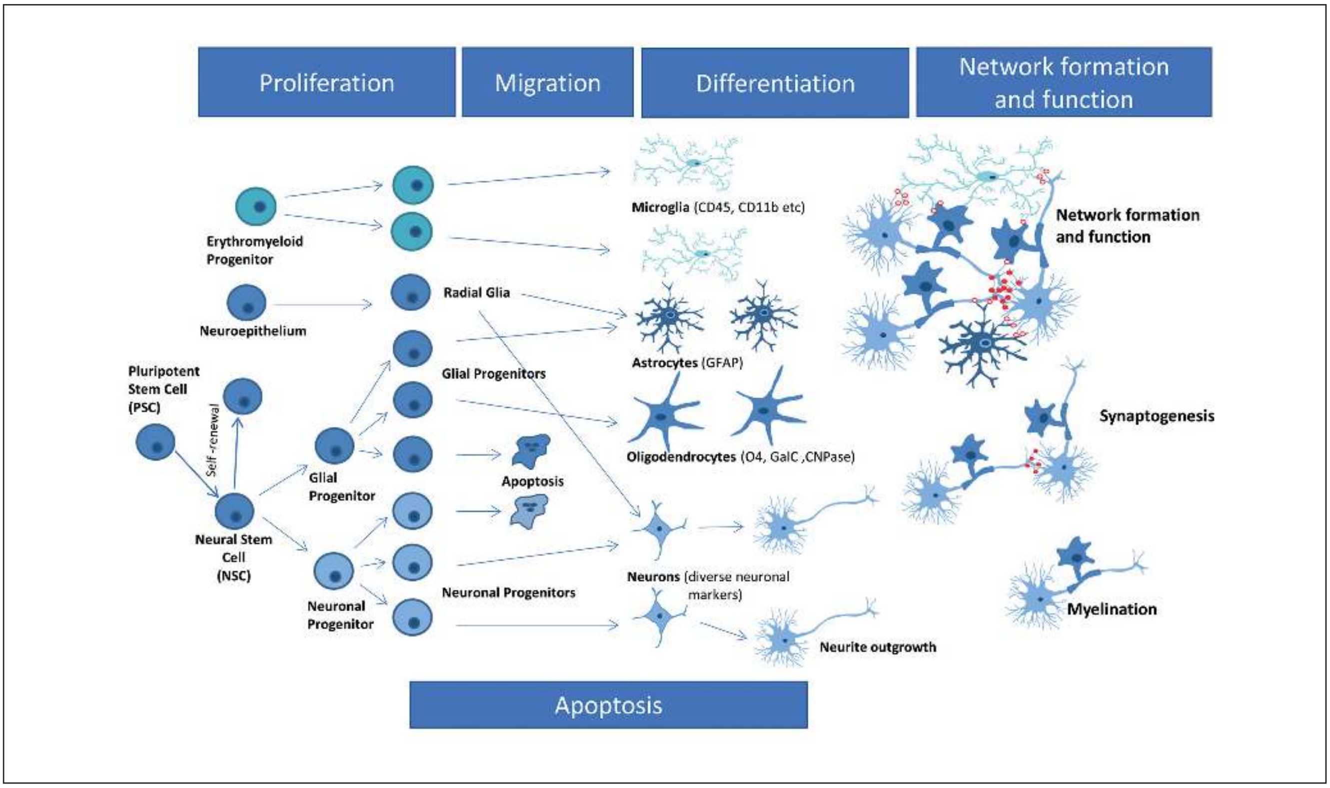


Fig. 1: Fundamental neurodevelopmental processes relevant for DNT
Several neurodevelopmental processes are essential for nervous system development. These processes, known from *in vivo* studies, can be relatively faithfully modelled *in vitro*. It is assumed that DNT toxicants exert their toxicity by disturbing at least one of these processes. Therefore, disturbances of the processes depicted here in blue boxes are KEs of AOPs relevant for DNT. The figure gives a short overview of nervous system development from simple precursors (left side) to complex functional tissue (with cell-cell interactions) on the right side. For a DNT test battery all these biological processes should be covered by one or more test methods.

for skin sensitization (OECD 2016b; Delrue et al., 2016; Adeleye et al., 2015; Urbisch et al., 2015). Transferring this concept to DNT, where currently only a few relevant AOPs are available, and where many more pathways might underlie toxicity for the developing brain (Bal-Price et al., 2015b, 2017; Bal-Price and Meek, 2017), a similar procedure is not yet feasible. Therefore, *in vitro* assays anchored to key cellular neurodevelopmental processes should guide the development of an alternative DNT testing battery (Fritsche et al., 2017a; Aschner et al., 2017; Schmidt et al., 2017; Bal-Price et al., 2010, 2012; Crofton et al., 2011).

Since 2005, an international community used the CAAT Tox-Smart DNT meetings as a basis to propose alternative approaches for DNT evaluation (Lein et al., 2005; Coecke et al., 2007; Crofton et al., 2011; Bal-Price et al., 2012, 2015a; Smirnova, 2014; Leist et al., 2012). The above-mentioned processes-based alternative DNT testing strategy is a result of this ongoing exchange between basic researchers and regulatory scientists. Such cellular KEs are intermediate to late KEs in an AOP, and examples from existing DNT AOPs include, e.g., "impaired neuronal differentiation" (Bal-Price et al., 2015b; Bal-Price and Meek, 2017), "decreased synaptogenesis" or "decreased neuronal network function" (Bal-Price et al., 2015b; Sachana et al., 2016), see also AOP-Wiki[1]. However, as the number of available DNT AOPs is small, basic clinical as well as toxicological sciences may inform us on relevant and measurable neurodevelopmental KEs, as summarized in Fritsche et al. (2015) and Fritsche (2017b).

Examples from the toxicological side include methylmercury-induced inhibition of neural cell migration (Bal-Price et al., 2015b; Moors et al., 2007), arsenic-induced inhibition of neural progenitor cell (NPC) proliferation (Chattopadhyay et al., 2002), valproic acid-induced inhibition of neural crest cell migration (Zimmer et al., 2012), or neuronal differentiation (Foti et al., 2013; Balmer et al., 2012, 2014; Waldmann et al., 2014, 2017). For these examples, the compounds' modes of action (MoAs) are being elucidated (Bal-Price et al., 2015b).

Knowledge from clinical research on neurodevelopmental disorders with genetic alterations as basis for disease are also helpful in determining human-relevant, cell-based endpoints. Here, for example, diverse receptor tyrosine kinase (RTK) mutations, leading to activation of protein kinase B (AKT, PKB), can cause

---

[1] https://aopwiki.org/

a variety of morphological disturbances in humans that are based on deregulation of brain cell proliferation and apoptosis (reviewed in Hevner, 2015). Also, aberrant expression of the brain derived neurotrophic factor (BDNF) and its dependent molecules, extracellular signal-regulated kinase (ERK) and cAMP responsive element binding protein (CREB), have been linked to numerous psychiatric disorders, including autism spectrum disorders, mood disorders and schizophrenia. Cellular functions controlled by these pathways are numerous, including brain cell proliferation, dendritogenesis, and synaptogenesis (reviewed in Ehrlich and Josselyn, 2016). These are only examples; a more detailed compilation of relevant neurodevelopmental pathways and cellular functions can be found in Fritsche et al. (2017b).

Modelling these key neurodevelopmental processes, from cell division up to neuronal network formation (Fig. 1), in a NAM testing battery will yield information on relative sensitivities of the processes to chemicals. For a small subset of endpoints, the principle of detecting the most sensitive process, and extrapolating from its disturbance *in vitro* to an *in vivo* hazard, has been exemplified in Baumann et al. (2016). Thus, information from batteries of tests run in parallel will not only serve as readouts for DNT hazard but will also inform future assay development and design of AOPs. While focusing on all these positive aspects, it will be important to bear in mind that fundamental issues of *in vitro* assays need to be kept in mind: for instance, the metabolic capacities that may differ from the *in vivo* situation, the interaction of different cell types that may largely affect their response pattern (Gantner et al., 1996), and issues of biological barriers (Leist et al., 2014; Kadereit et al., 2012; Aschner et al., 2017).

2 General guidance of quality and performance standards

2.1 OECD guidance on test descriptions and readiness

The rationale for alternative DNT testing is given by the consensus between academic, industry and regulatory scientists that chemicals with the potential to trigger DNT should be properly identified and that the current testing paradigm, based on *in vivo* studies, does not satisfy this need (Fritsche et al., 2017a). For moving alternative DNT tests into action, scientists should focus on defining and applying test specifications (Leist et al., 2010, 2012) and validation paradigms to evaluate their readiness and draw a roadmap for their application in a regulatory context.

The meaning of the term readiness varies strongly between different interest groups (Fig. 2). For instance, an academic researcher uses a cellular model system to investigate pathways of cellular functions and needs a reliable model that mimics human effects. However, this is only the starting point for the work of the test system developer.

Regulatory acceptance of individual tests will be facilitated by adherence to international regulatory consensus guidance. For instance, the OECD Guidance Document No. 211 (GD 211) provides a template for assay annotations of non-guideline *in vitro* methods (OECD, 2014b). GD 211 harmonizes the manner in which non-guideline *in vitro* methods are described, and thereby facilitates assessment (by the regulator) of the reliability and relevance of the produced data. The US EPA Office of Pesticide Programs recommends following this guidance to describe non-guideline *in vitro* methods for acute toxicity (EPA, 2016). According to this guidance (OECD, 2014b), the method description should include purpose and scope of the assay, method components including protocol and reference chemicals, the stage of development of the assay, the quality/acceptance criteria, data interpretation and prediction model(s), and performance metrics including sensitivity and predictivity (i.e., proportion of false negatives for positive controls, and of false positives for negative controls).

A further important document is the guidance document on *Good In vitro Method Practices* (GIVIMP) for the development and implementation of *in vitro* methods for regulatory use in human safety assessment (expected release: 2018). This guidance (draft version OECD, 2017a) will be of central importance in regulatory acceptance of the proposed DNT alternative methods. It describes the scientific, technical and quality practices needed at all stages between *in vitro* method development to implementation for regulatory use. These include roles and responsibilities (of developers, component providers and users), quality considerations, facilities, apparatus, material and reagents, test systems, test and reference/control items, standard operating procedures, method performance, and reporting of the results. The GIVIMP document has been written for various users, including GLP test facilities but also research laboratories developing new *in vitro* methods for regulatory purposes. In the latter case, full compliance with GIVIMP may not be realistic, but compliance with as many as possible of the “good practices” will facilitate the acceptance and routine use of the *in vitro* method in a regulatory environment.

It is understandable that the completeness of the information recommended in the OECD guidance will vary, because the level of development of the DNT alternative methods is different, and this in turn impacts the use of the methods for different regulatory applications. However, in all cases, the suggested framework aims to cover some information on 1) a test method definition (including purpose, scientific principle, metabolic competence, quality control criteria, technical limitations and strengths); 2) test method performance (robustness, reference chemicals, performance measures/predictive capacity); 3) data interpretation; 4) potential applications; and 5) supporting information available in the existing databases (e.g., DataBase on ALternative Methods DB-ALM of EURL-ECVAM[2]).

In this context, the consideration of “applicability domains” takes an important and often underestimated role. The test method must be considered like a tool. And like all tools, it has a proper domain of application (e.g., scissors to cut paper), borderline domains of application that require case-by-case evaluations (e.g., use of scissors to punch holes or to open a bottle),

---

[2] https://ecvam-dbalm.jrc.ec.europa.eu/

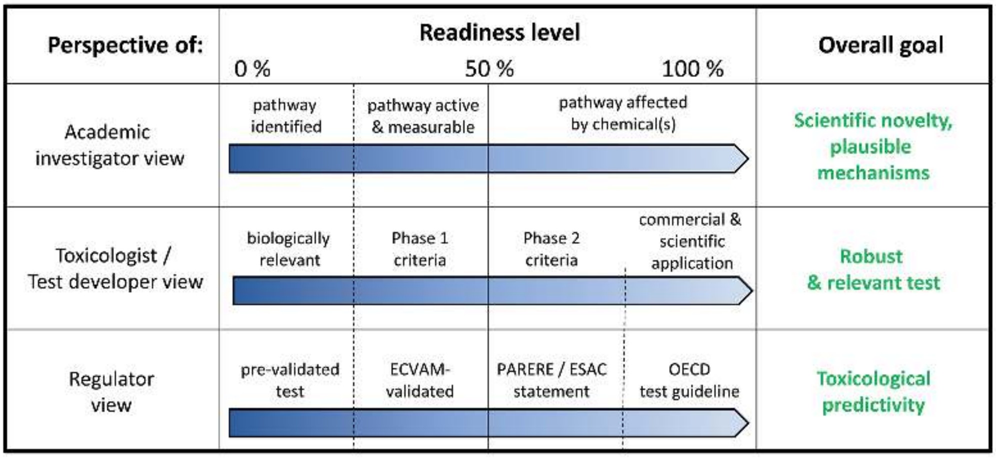


Fig. 2: Different perspectives of DNT alternative methods' readiness evaluation
In the discussion on "test readiness" it is important to note that different fields and stakeholders have their own perspective. Three of these perspectives are outlined. For each of them, examples for increasing grades of readiness and final goals are given. These perspectives are interdependent to some degree: (i) a test that is 100% ready for an academic investigator in basic science can form the starting point for a toxicological test developer; (ii) a test that is considered ready by the test developer may be at the start of regulatory readiness, e.g., with respect to formal validation; and (iii) a test that is at the highest regulatory readiness level (OECD TG) may provide a starting point for academic researchers who want to unravel key mechanisms and pathways that are essential and that biologically explain the test read outs.

and applications that are physically possible, but usually lead to non-satisfactory results (e.g., use of scissors to open a can or to turn screws). For DNT test methods, several dimensions of "applicability domains" are important. The three most important ones are:
(i) the type of chemicals to be tested;
(ii) the type of mechanisms explored;
(iii) the type of (regulatory) questions addressed.
Thus, a given method may be more ready for certain applications and less ready for others!

2.2 Principles for evaluation of the readiness of test strategies based on multiple test methods

A systematic approach to building a test battery should first determine the readiness of individual alternative DNT methods. A general set of readiness criteria has been proposed by OECD (2014b), and these have been clustered in four categories (Tab. 1). Such guidance has been considered here in compiling specific readiness criteria for DNT test methods, and in devising a preliminary scoring system to obtain indications on the readiness status of various published tests (see chapters below). Currently, none of the proposed DNT alternative methods are stand-alone methods, thus a battery of the assays that capture essential information across neurodevelopmental processes and developmental timing is considered important for a comprehensive hazard assessment. Here, we discuss briefly the evaluation of ITS.

The evaluation of ITS could be based on the principles developed for the reporting of DAs to testing and assessment based on multiple information sources (OECD, 2016b). A DA can be built in various ways and may take the form of a sequential testing strategy (STS) or an ITS. The fixed data interpretation procedure is then used to interpret data generated with a defined set of alternative methods that can either be used on its own or together with other methods and existing information within an IATA (OECD, 2016a). In this case, the template for data reporting of individual information sources used in a DA that was published in an OECD guidance (OECD, 2016b) will ensure a transparent and accurate documentation of the methods used within a DA. Within such a DA, information has to be documented properly to ensure transparency of the methods used. The description should include the chemical and/or biological mechanism addressed by the methods and provide some indications of the plausible linkage of the modelled mechanisms or neurodevelopmental processes to the apical endpoint being predicted. Known scientific confidence and limitations of methods should also be reported, including a comparison to existing similar non-testing or testing methods.

Principle 1 aims to ensure clarity in the endpoint addressed, by defining it. From this perspective, a relationship between the combination of the alternative test methods' endpoint(s) and the biological phenomena of interest should be explored. The limitations (e.g., inability to determine DNT effects secondary to systemic effects like hormonal imbalance) are to be clearly identified. The scientific validation of the testing strategy should be based on a mechanistic ground with the assumption that a derangement of fundamental processes in neurodevelopment will lead to an adverse effect.

Principle 2 aims to ensure clarity in the purpose for which the combination of the alternative methods is proposed. Considering that a test method should fit for a specific purpose, the problem formulation should be defined at the beginning of the process. This would not only include the regulatory purpose, i.e., screening and prioritization vs. single chemical hazard identification, but it would also specify the target performance values (predictive capacity required).

Principle 3 intends to provide transparency on the rationale used for applying DAs. The rationale may be based on an existing AOP or network of AOPs or other mechanistic information relevant to the endpoint. In the case of DNT, due to the limited number of available AOPs, mechanistic information derived from studies exploring disturbance of brain development processes by well-established DNT compounds can form the rationale for constructing a DNT testing strategy that relies on alternative methods (Fritsche, 2017a,b).

Principle 4 deals with data generated by the different information sources and how it is used within the DA to derive a prediction/assessment and aims to provide transparency on this aspect. The description should ideally include a schematic representation (e.g., flowchart or decision tree) to illustrate the procedure. The approach followed to provide prediction needs to be documented and understandable by the regulators.

Principle 5 allows the capture of the sources of uncertainty in predictions. Of particular interest would be to define if the proposed DNT testing strategy is reliable only for positive predictions or only for specific pathways or mechanisms of action. Additionally, the level of confidence (reliability of prediction) associated with the application of the testing strategy to different chemicals is needed. It is relevant to include as many chemicals as feasible as the determination of the applicability domain is expected to be correlated with the number and diversity of chemicals tested. More importantly, this principle aims to capture the variability of the data produced by the alternative methods as well as the variability of the output data (i.e., from the DA) associated with the reference data (e.g., animal or human DNT data) used as benchmark data. In other words, the prediction of a DA aims to capture the variability and uncertainty of the alternative approach and the reliability of the gold standard data by applying appropriate statistical concepts and qualitative approaches.

The application of these criteria and principles helps to establish the overall relevance of the alternative methods and of the testing strategy.

3 Evaluation of *in vitro* DNT assays against defined readiness criteria

3.1 Compilation of readiness criteria

The development and application of *in vitro* test methods is driven by various stakeholders: basic academic researchers, test developers in industry and public institutions, and regulatory decision makers. As shown and discussed in Figure 2, these three

Tab. 1: Example for ranking parameters for *in vitro* methods to detect chemicals that disturb the thyroid hormone axis
Ranking parameters were established by OECD for thyroid-disrupting chemicals to determine the readiness of tests for validation (OECD, 2014a). The criteria in *Category 1* are considered of highest priority. Each criterion within this category is considered to have equal weight, and all are essential to demonstrate the readiness of the assay. For instance, the assessment of the biological plausibility is considered very important in defining readiness of the method for validation. However, criteria in this category are hard to quantify. Moreover, many DNT tests cover multiple mechanisms and processes with varying levels of plausibility and data on their *in vivo* relationship. Thus, the practical value of such criteria for DNT methods needs to be considered case-by-case. The criteria for *Category 2* are better defined and quantifiable. They relate to the evaluation of reliability and efficacy of the method. Sufficient positive and negative compounds should be included to assess specificity and sensitivity, and focus should be given to the robustness of the assay. Regarding *Category 3*, the criteria are also relevant to assay performance evaluation. However, the particular performance issues described under this category are considered to be of less significance during initial phases of test development and evaluation. *Category 4* contains criteria for the methods that are considered good to meet in order to gain broad acceptance.

CATEGORY 1 **Initial high priority considerations**	**CATEGORY 2** **Method performance considerations**
Biological plausibility Extrapolation to humans or broadly applicable across vertebrates/phyla Availability of resources Reference chemicals	Within-laboratory reproducibility Between-laboratory reproducibility Assay variability Accuracy Assay specificity/assay sensitivity
CATEGORY 3 **Technical capability**	**CATEGORY 4** **Other practical considerations**
Dynamic range/concentration test range Detection/adjustment of confounding factor and/or incorrect/inconclusive measurements and/or other bias Response characterization	Technological transferability/proprietary elements Transparency of the method Documentation of development and utility of the method

groups may have different points of view regarding the readiness of a test method. Moreover, readiness depends on the application of a method, on the field of toxicology in question (here DNT), and on the quality of animal experiments in the given field (Hartung and Leist, 2008). To take all this into account, a 2-step consensus process has been organized to establish a practical set of readiness criteria. They were first suggested and discussed during a workshop with different stakeholders and then assembled for this report by a working group. A third step (described below) involved testing of the applicability of the criteria for actual scoring.

The criteria were clustered into 13 groups, e.g., concerning the test system, the prediction model, or the applicability for screening. For each of the criteria, a short heading was defined (e.g., critical components of the cell system). Then, the criterion was described in more detail. To do this, specifying or guiding questions often have been defined that need to be answered to provide information on the respective criterion. For instance, for the "critical components of the cell system", this is "Have critical components and handling steps been identified and have they been clearly and explicitly described? Are examples for normal performance and morphology given? Are there examples for alerts?" Finally, examples of the type of information required are given. In the chosen example: "E.g., cell density on a specific day of differentiation could be a critical step; wrong, strange morphology of cells could be an alert". In this way, a compromise was reached between length (and clarity) of the document, and the information needed to perform a readiness evaluation (Tab. 2).

Our criteria list is meant to provide an easy-to-use tool for test developers and users in order to provide a quick and fast overview for them to judge how far the method is developed and what important points need to be addressed. Moreover, the semi-quantitative or quantitative scoring may help regulators to identify the strengths and weaknesses of a given test method. This could help them to decide to what extent the data generated by a given test method could be used. Notably, the tool may also be useful to identify and exclude data from non-ready methods from regulatory use or to prevent scientifically unsound data from creating anxiety in the general public.

Tab. 2: Performance criteria to define the readiness of test methods for hazard evaluation
This set of criteria was developed with the needs of toxicological test developers in mind. It should help them to prepare their assay for priority screening as well as for incorporation in an ITS. In the first column, the criteria are listed in their short form, the second column gives a definition or short description of each criterion (with some supporting and guiding questions), and the third column provides examples or further explanations for each of the criteria. The fourth column gives the maximum score that can be reached. There are 13 main categories of criteria, each with different numbers of sub-items. Within each main category, the sub-item can be scored for a readiness evaluation, and the sum of these scores results in the score for the main category. The fourth column indicates the maximum score that can be given for each category. The main criteria can be assigned to three different phases of test method development (Phase I in pale blue, phase II in medium blue, phase III in dark blue, Fig. 3). The topics printed in italics (e.g., 1j, 3a, 3c, 4d, 5d, 5e, 5i) may not apply to each test method. If they do not apply, the score is automatically set to 1 for these sub-items.

Abbreviations: AOP, adverse outcome pathway; BMCL, benchmark concentration lower bound; CRO, contract research organization; EC, endpoint-specific controls; KE, key event; MIE, molecular initiating event; NC, negative controls; PC, positive controls; S/N, signal noise ratio; SOP, standard operation procedure; STR, short tandem repeat; UC, unspecific controls.

Criteria	Description	Examples / Why is it important	Max. score
1 Test system		*Note: here scoring not for "test method"*	**10**
1a What is modelled	Is there a clear rationale given for what target organ/tissue relevant for human poisoning/pathology the test system should reflect?	Here: Question is not of relevance but whether there is documentation and a rationale at all.	1
1b Relevance	Is the chosen test system known to be a key component in pathogenesis, or why is it thought to reflect a key component, mechanism or tissue?	Here: Is the tissue/organ modelled important for regulatory toxicology or biomedical research purposes? Is evidence given for the relevance of the model by morphological comparison, gene expression or functional criteria? Are all/sufficient cell types included in the model?	1
1c System uncertainties and human correlate (HC)	(i) Is there a discussion on where the test system differs from the mimicked human tissue, and which gaps of analogy need to be considered? (ii) Do toxicant-altered genes (or other biomarkers) correspond to changes in mimicked human tissue (after poisoning or in relevant pathologies)?	(i) E.g., a differentiated cell or a cell line (such as HepG2) does not necessarily reflect all features of the corresponding *in vivo* tissue/conditions. (ii) This is an additional measure to increase confidence in the test – not mandatory, but helpful.	1

Criteria	Description	Examples / Why is it important	Max. score
1d Definition of cells	Is the test system sufficiently characterized (source, multiple positive and negative markers for cell identity, number, quality, composition, differentiation state, viability, usual morphology, basic function, basic reaction to stimuli, STR analysis)?	This is especially important for cells that have to be produced regularly, e.g., by differentiation or primary cell isolation.	1
1e Cell composition	For multi-component systems: information on all cellular subpopulations. What is the percentage of contaminating cells or in co-cultures what is the percentage of all subpopulations.	This is important for the test endpoints as it could be that only one cell type may be affected by a toxicant. For primary cells: Have cells from different sources (suppliers) been tested (e.g., hepatocytes from different suppliers may differ in purity and quality)? For routine use it would be beneficial to have pre-set acceptance criteria for each cell type.	1
1f Cellular environment	Information on structuring components of the test system: coating, scaffolds, matrix description, medium (supplements), microfluidic effects, supportive cells, dimensions and positioning/handling of 3D constructs, etc.	This means a very detailed description of the culture conditions, including temporal and spatial aspects. Cell differentiation and response (quality, quantity, kinetics) may depend on multiple external factors and on the 3D arrangement.	1
1g Biological consistency	(i) Has the variation of the test system been assessed, have influencing factors been identified? (ii) Have acceptance criteria and performance standards for the test system been defined (different from the test!)?	(i) E.g., do medium supplements have an influence on the outcome of the cells, such as batch effects of FCS or serum replacement additives? (ii) E.g., a range of marker expression levels, of biological function (proliferation, protein production, etc.), of structural features (cell number, organoid size, etc.). For cell lines: What is the optimum passage number for cells? For routine use it would be beneficial to have pre-set acceptance criteria for the whole model/test system.	1
1h Critical components	Have critical components and handling steps been identified and described? Are examples for normal performance and morphology given? Are there examples for alerts?	E.g., cell density on a specific day of differentiation could be a critical step; wrong, strange morphology of cells could be an alert. For routine use it would be beneficial to have pre-set acceptance criteria.	
1i Cell stability	Stability proven over multiple doublings; genetic stability shown; pluripotency/ multipotency (for stem cells) shown, cell identity shown.	For stem cells, stability needs to be shown over many passages (≥ 10). For primary cells, stability and identity of supply needs to be shown. Stability of function (e.g., xenobiotic metabolism) needs to be shown.	1
1j *Transgenic cells*	*Transgene characterized (source, sequence, regulation); insertion characterized; stability of function shown and quantified; cell identity and function related to wildtype; clonality documented.*		*1*
2 Exposure scheme			**3**
2a Description	Complete, detailed, unambiguous	Medium changes, re-additions, coating, treatment period and timing, incubation conditions (temperature, humidity, gassing, etc.)	1
2b Unique identity	Tests with multiple variants need to define very transparently, which variant the data comes from.	E.g., from which cell type/clone; which time; which plate format; which medium additives, etc.	1

Criteria	Description	Examples / Why is it important	Max. score
2c Graphical scheme	Complete sequence of events including endpoint assessment	Supports clarity and data assignment to test variants	1
3 Documentation / SOP			**5**
3a *Availability*	*Method description for test system, test procedure, analytical endpoints and prediction model; public availability of SOP (data bank or test developer upon request)*	*Normal scientific publications are usually not sufficient, except for specific methods papers. For transferability of the test method it is beneficial to have SOPs or other documents covering each component of the test method and the whole testing process*	*1*
3b Stage of development	Version history; updated		1
3c *For CRO tests*	*Are full performance standards and corresponding data delivered by the CRO along with test data (in case SOP details are not disclosed)*	*Non-disclosure of SOP is acceptable if full performance/readiness criteria are given.*	*1*
3d Test components	Documented and available (receipt, storage, handling and disposal documents); quality criteria and checking procedure established	E.g., for media, plates, coating it should be defined what is acceptable/non-acceptable and how this is controlled. Test chemical identity and purity (certificate of analysis) and safety data sheets for chemicals	1
3e Stocks	Procedure for preparation, storage and quality control of stocks established		1
4 Main endpoint(s)		*Mainly referring to specific/functional endpoints*	**4**
4a Biological relevance	Is there a rationale given why the test endpoint is relevant to adverse outcomes?	Helps to interpret the results.	1
4b Toxicological relevance	Are toxicants (≥ 3) known to affect the endpoint	Helps to interpret the results.	1
4c Analytical methods	Methods defined, rationale given; positive controls and acceptability criteria	Positive controls for analytical method may differ from controls for test/endpoint	1
4d Multiple endpoints	Are all endpoints and their relation to one another (priority, preference) defined?	E.g., neurite outgrowth, cytotoxicity	1
5 Cytotoxicity		*Here: if cytotoxicity is not main endpoint*	**5**
5a Cytotoxicity within test	Cytotoxicity is preferentially determined within same test compartment as the major endpoint; second choice is under same conditions in parallel	Control of cytotoxicity in a different format (e.g., other types of plates; other time point, is very problematic). Measuring cytotoxicity under the same test conditions as the main end point helps to interpret the mechanism related to the adverse effects for the main end point (specific or cytotoxicity-driven mechanism)	1
5b Subpopulation effects	Are subpopulations detected by measure for cytotoxicity or proliferation? Are minor changes detected? Has sensitivity been shown?	Usually at least three types of assay required (measurement of viability, measurement of cell death, single cell analysis)	0.5
5c Specificity (compared to cytotoxicity)	A measure needs to be established to distinguish a specific/functional endpoint from cytotoxicity	E.g., neurite outgrowth, migration inhibition in non-cytotoxic concentration ranges	0.5
5d *Timing within test*	*For repeated/prolonged dosing, early death and compensatory growth need to be considered.*	*Testing for cytotoxicity only at the end of the test may give false negative data if cells die early and this is no longer detectable at later time points because of compensatory proliferation.*	*0.5*

Criteria	Description	Examples / Why is it important	Max. score
5e *Timing after test*	*For very short endpoints, e.g., electrophysiology measured 30 min after toxicant exposure, delayed measure of cytotoxicity is necessary*	*Cells cannot die in a very short time even if compound triggers lethal changes. Data for 24 h exposure should be given.*	*0.5*
5f Curve fitting	Sufficient non-toxic data points (baseline); at least 40% toxicity / change to allow fitting		0.5
5g Non-cytotoxicity	Absence of ‘cytotoxicity’ does not mean non-cytotoxicity (question of power): Has data variation been considered; is a measure of uncertainty given for non-cytotoxicity (e.g., BMCL calculation)?		0.5
5h Bench mark response	Has a rationale been given for setting a threshold value for cytotoxicity (statistical or biological significance)	E.g., statistical: 3x standard deviation; biological: 90% viability; see also: http://invitrotox.uni-konstanz.de/	0.5
5i *Apoptosis/ proliferation*	*If natural feature of the test system: measure for normal rate required*		0.5
6 Test method controls			**4**
6a Positive controls (PC)	≥ 3 toxicants required for test definition; preferentially with different mechanisms; preferentially human-relevant toxicants; indicate variation of PC within and across assays.	Used to define acceptability criteria, S/N ratio or z’-value of screen	1
6b Negative controls (NC)	≥ 5 negative controls are required to define specificity at ±20% level; concentration of negatives needs to be defined and rationalized	Ways to define negatives: (i) e.g., compound only acting when metabolized, (ii) acting on another organ, (iii) known to be safe for pregnant women, (iv) is selective for another assay, (v) pairs/matches of a specific positive control (e.g., inactive metabolite)	1
6c Unspecific controls (UC)	A type of negative control for functional assays: not inactive, but only cytotoxic	Absolutely essential to define baseline variation and thus the relevant benchmark response for positive hits	1
6d Endpoint-specific controls (EC)	To provide plausibility and to help initial test setup: EC show that pathways considered to be relevant for test endpoint indeed affect the test endpoint. EC help to correlate (by concentration and time) compound effect on pathway (activity measure to be established) and on test endpoint (standard test readout). EC may be chemicals or siRNA; pathways may be defined from literature or experimentally (gene expression)	E.g., actin is required for migration, thus an actin inhibitor should affect migration endpoint	1
7 Data evaluation		*Here: referring to main endpoint(s)*	**4**
7a Outliers	Procedure for handling and documentation should be established		1
7b Concentration -dependence	Higher confidence in concentration-dependent data; no-effect concentrations must be included (full range curve); data need sufficiently dense spacing around benchmark concentration; preferably provide statistical significance for key data points		1
7c Benchmark response	Give rationale for definition (statistical (after FDR correction) or biological). Provide power estimate if conclusions are drawn from negatives.		1

Criteria	Description	Examples / Why is it important	Max. score
7d Curve fitting	Indicate detailed procedure used for curve fitting; preferentially force fitted curve through 100% at negative control conditions (full function)	E.g., sigmoidal, linear or exponential curve fit	1
8 Testing strategy			**4**
8a Hazard prediction	Which hazard is assessed? Which question does the test method answer?		1
8b Link to an AOP	Does the test give input to a mechanistic concept, e.g., an AOP?	Helps to position test in battery; helps to interpret results	1
8c Role in battery	Full score for stand-alone tests. For tests that are not stand-alone, information on their relation to other tests in a battery is required.	Information is required on how the test data would be used in a battery and under which conditions this is possible.	1
8d Comparison to similar tests	Does the test fill a gap in a battery? Is it providing advantages compared to another test for the same hazard?	Avoid overlapping tests being performed; ensure adequate testing battery/strategy	1
9 Robustness			**4**
9a Reproducibility	Data available on normal variation; information on factors affecting test variation is given	Historic control data on positive controls show normal range; known artefacts and shortcomings	1
9b Intra-lab	Data available from different operators and different test runs over longer time		1
9c Inter-lab	Data available on transferability / reproducibility in another lab		1
9d Historical controls	Data for PC and NC over time		1
10 Test benchmarks			**4**
10a Sensitivity (of the test)	Signal noise ratio (S/N) defined; Sensitivity information available	S/N based on adequate data sets. The S/N is used to determine the limit of detection. Additional measures: True positive rate, hit rate, sensitivity to detect a panel of positive controls, etc.	1
10b Specificity (of the test)	Tested with sufficient number and quality of negative controls	Additional measures: true negative rate, etc.	1
10c Acceptance criteria	Clearly defined and documented; Normal range of variation known	E.g., a given positive control has to reduce the main endpoint by at least 25%, otherwise test plate is discarded.	1
10d Response characteristics	Should the response be linear? What are the upper and lower limits?	Additional measures: mono-directional or bi-directional deviation defined, information on accuracy, precision, limit of quantification, etc.	1
11 Prediction model			**4**
11a Definition	Information should be available and clear (including rationale for model, i.e., its particular strengths). Information and rationale should be given for use of sharp thresholds or probabilistic approach.	Information on how many classes of toxicants are predicted. Positives and non-positives; or strong, medium, weak positives. Information on uncertainty of prediction should be given, at least for positives (note that uncertainty of negatives is often not defined). E.g., you can define a sharp threshold (all above 4 is positive) or you can define a probabilistic approach (above 4 has a 70% likelihood to be positive)	1
11b Rationale	Reason, and mathematical basis / plausibility for prediction model given	Reason for the choice and value of thresholds	1

Criteria	Description	Examples / Why is it important	Max. score
11c Confirmation	Experimental testing of prediction model; confirmation of function/predictivity		1
11d Limitations	Information on limitations of prediction model and on how exceptions and special cases are to be handled	Strange curve shapes, solubility issues, assay interferences, etc. How special chemical classes are handled	1
12 Applicability domains			**3**
12a Chemicals	Is information on the types of chemicals that fall into the prediction model / testing range available?		1
12b Pathways	Type of pathways that are relevant for the test (to be disturbed or to be detected)		1
12c AOP	Information contributed to an AOP KE/MIE; element of a KE testing battery		1
13 Screening hits			**4**
13a Hit definition	Transparent, pre-defined criteria (including curve-fitting/statistical procedure)	Usually, non-hits are discarded. If statements on non-hits are made, they need definition and power calculation.	1
13b Hit confirmation (prim.)	Independent test run(s) in “same” test method; full concentration-response	Often loose (soft) criteria for hits, and no correction for false discovery rate. Confirmation assays can counteract such problems; use of new cells and new compound stocks provides additional robustness.	1
13c Hit confirmation (sec.)	Additional test (different from primary test method) confirming hit on same endpoint as screen	E.g., migration may be measured by tracking cells (primary test) and then (secondary test) by a Boyden chamber method.	1
13d Screen documentation	Acceptability criteria, performance of positive controls, internal robustness controls		1

3.2 Scoring system for readiness criteria

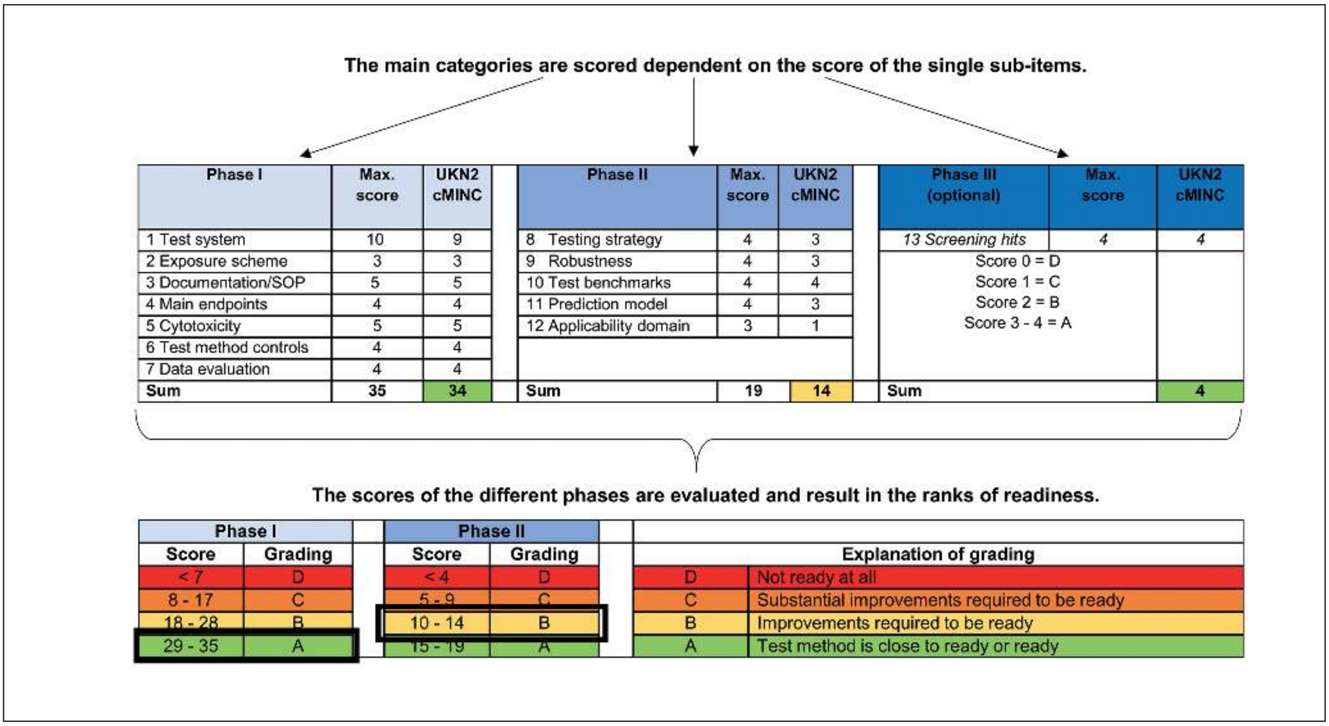
According to the OECD GD 211 (OECD, 2014b), the new generation of *in vitro* test methods may be very useful for some regulatory purposes, even if they are not yet officially validated. For instance, they may be used to provide additional/supplementary mechanistic information on top of standard testing results. Moreover, such tests may be used in companies or regulatory authorities for internal decision making, or for screening programs with the aim of prioritizing substances for further testing (Browne et al., 2017). Although there is guidance on what needs to be considered for test method validation, not many tools are available that provide an actual measure of readiness.

Since readiness needs to be quantified to a certain extent, a simple scoring system was established with the intention of providing a rough quantification of readiness levels. In the future, such a system may be further refined, concerning the criteria considered, the weight given to the criteria, and especially by providing guidance on how the scoring is performed. Here, the system was kept simple, by assigning a maximum score to each criterion (see fourth column in Tab. 2), and by establishing a simple tool for clustering of scores (Fig. 3). The scores were assigned on the basis of publicly available information extracted from publications. The process may be facilitated in the future by a process that assembles all relevant information in a “readiness dossier”, including data not easily found in publications, e.g., provided by test developers and applicants.

This clustering of scoring criteria is an important concept, as it provides individual scores for phases of test development. Phase I concerns all criteria that can be fulfilled during initial test method development. Phase II criteria refer to the test method performance based on, e.g., evaluation of replicates to conclude on robustness and reproducibility. Phase III is optional as a proper screening is not always feasible for each test method, i.e., 2nd and 3rd tier methods. This allows a distinction of readiness for, e.g., academic research purposes, screening and prioritization, or regulatory risk assessment. The example of the UKN2 test shows that a method can have a high readiness level for screening, but still needs further improvement of hazard assessment of individual compounds in the context of a risk assessment process (Fig. 3).

3.3 Exemplary DNT test methods and their preliminary evaluation for readiness

To demonstrate the application of readiness scoring for DNT assays, a set of 17 test methods was selected, and the scoring was performed. Notably, the information used had to be extracted



The main categories are scored dependent on the score of the single sub-items.

Phase I	Max. score	UKN2 cMINC
1 Test system	10	9
2 Exposure scheme	3	3
3 Documentation/SOP	5	5
4 Main endpoints	4	4
5 Cytotoxicity	5	5
6 Test method controls	4	4
7 Data evaluation	4	4
Sum	**35**	**34**

Phase II	Max. score	UKN2 cMINC
8 Testing strategy	4	3
9 Robustness	4	3
10 Test benchmarks	4	4
11 Prediction model	4	3
12 Applicability domain	3	1
Sum	**19**	**14**

Phase III (optional)	Max. score	UKN2 cMINC
13 Screening hits	*4*	*4*
Score 0 = D Score 1 = C Score 2 = B Score 3 - 4 = A		
Sum		**4**

The scores of the different phases are evaluated and result in the ranks of readiness.

Phase I		Phase II	
Score	**Grading**	**Score**	**Grading**
< 7	D	< 4	D
8 - 17	C	5 - 9	C
18 - 28	B	10 - 14	B
29 - 35	A	15 - 19	A

Explanation of grading	
D	Not ready at all
C	Substantial improvements required to be ready
B	Improvements required to be ready
A	Test method is close to ready or ready

Fig. 3: Scoring system for readiness criteria
Overview of the scoring system for the readiness criteria. The 13 criteria are sorted into three phases. Each area has various sub-items and the number of points that can be obtained is indicated in Table 2. Phase I includes the basic features of the test method as they would be provided by academic researchers. They include biological plausibility of the test method, features of the test system, and the availability of controls. A high number of points can be obtained for test system description (10 out of 35), as this is very important at early stages of test development. However, still two thirds of the points come from other areas not to be neglected. Phase II relates to the implementation of a test for practical applications in industry or for regulatory purposes. Here, the relation to a testing strategy, good robustness, and the availability of a prediction model are important. Phase III is optional as not each test method is used for a screening approach. Notably, not all points apply to all tests. In the preliminary rating scheme suggested here, these items are then scored positive automatically (labeled in italics in Tab. 2). Each phase is evaluated independently, and then categorized into one of four readiness classes (A-D). In the figure, an example is given for the rating of the cMINC (UKN2) test method. It would score as “A” (largely ready) in phase I, and as “B” in phase II. For phase III, it would score as “A”.

from the published literature, and thus some information may have been missed or may not have been taken into account. It is also important to note that some methods were not developed specifically for regulatory use. In such cases, information retrieval was from multiple publications, and there were uncertainties and ambiguities concerning several criteria. A more formalized process of information retrieval might lead to higher scores. The selection of scored test methods was meant to give a representative overview of what is available to test interference of chemicals with various neurodevelopmental processes. The selection does not purport to be complete.

The individual scoring information can be found in Table S1[3]. A summary overview is given in Table 3. In the following, some additional details are given on the test methods that have been considered here.

Differentiation of pluripotent stem cells into neural precursor cells (UKN1 test)

This test is exemplary for tests examining processes in the embryonic (very early) phase of brain development. A very early step in embryonic development is the lineage specification of the cells of the inner cell mass into the three germ layers, endoderm, mesoderm and ectoderm (Leist et al., 2008b). The ectoderm is further divided into neural ectoderm, which gives rise to the central nervous system, and the non-neural ectoderm.

The UKN1 test method mimics this early neuroectoderm lineage specification. Human pluripotent or embryonic stem cells (hPSC or hESC) are differentiated into early neuroectoderm progenitor cells. This stage is reached after 6 days under the given assay conditions (prevention of SMAD signaling) (Balmer et al., 2012; Balmer and Leist, 2014). The differentiation is extensively

---

[3] doi:10.14573/altex.1712081s

Tab. 3: Overview of the readiness levels of exemplary DNT test methods

Different DNT test methods were scored according to the criteria presented in Table 1 and Figure 3. The readiness levels for Phase I-III varied from D (not at all ready), over C, B to A (largely ready), see Fig. 3. For detailed data see Table S1[1]. The overall readiness was estimated semi-quantitatively from the pattern of readiness in the different phases. Notably, the term overall readiness has to be used with care, as readiness depends on the purpose, and it is given here only to provide an orientation on the availability of methods in the field. This is exemplified by the cMINC (UKN2 method) which scores A in Phase I and B in Phase II. According to this, the method is not ready for regulatory risk assessment. However, it scores A for screening, and is thus ready for initial prioritization of compounds.

Abbreviations: UKN1, PSC differentiation into NPC/NSC embryonic phase differentiation; NPC1, hNPC proliferation; NPC2, hNPC migration; NPC3, hNPC neuronal differentiation; NPC4, hNPC differentiated neurons; NPC5, hNPC oligodendrocyte differentiation; NPC6, hNPC oligodendrocyte maturation and TH disruption; UKN2, NCC proliferation and migration; MESn, Morphological ESC to neurons; 3Dr, astrocytes, oligodendrocytes, myelination, microglia in 3D rat; 3Dh, astrocytes, oligodendrocytes, myelination, microglia in 3D human, foetal phase; UKN4 (NeuriTox), neurite outgrowth of central neurons; UKN5 (PeriTox), neurite outgrowth of peripheral neurons; NSR, neuronal subtype ratio, neuronal maturation; Syn, synaptogenesis; Nnff, neuronal network formation and function; ZFE, zebrafish

Readiness/ Test method	Phase I	Phase II	Phase III	Overall readiness
UKN1	A	B	B	B+
NPC1	A	A	A	A
NPC2	A	A	A	A
NPC3	A	A	B	A-
NPC4	A	B	C	B
NPC5	A	A	B	A-
NPC6	A	B	B	B+
UKN2 (cMINC)	A	B	A	A-
MESn	C	D	D	D+
UKN4 (NeuriTox)	A	A	A	A
UKN5 (PeriTox)	A	B	A	A-
NSR	C	D	D	D+
SYN	B	B	B	B
Nnff	B	A	B	B+
3Dr	A	A	A	A
3Dh	B	C	C	C+
ZFE	B	B	A	B+

characterized by whole transcriptome analysis, showing that the differentiation protocol results in a homogenous neuroepithelial progenitor (NEP) cell population with an anterior gene expression patterning. The process has been extremely well characterized on the level of transcriptome and epigenetic changes (Shinde et al., 2015, 2016; Rempel et al., 2015; Balmer and Leist, 2014; Weng et al., 2012, 2014). A change in this gene expression pattern indicates a wrong differentiation track and may help to measure KEs such as neural tube patterning or neural differentiation (Rempel et al., 2015; Tonk et al., 2015; Krug et al., 2013).

The evaluation of UKN1 with our suggested criteria list revealed that the system is ready concerning phase I. For phase II, the transferability to another laboratory is missing, as well as a final confirmed prediction model. It is a challenge to set up a prediction model based on gene expression data alone. Therefore, anchoring of data to a functional endpoint (rosette formation) will be included (Waldmann et al., 2017). Regarding the screening issue of phase III, this test method reaches a readiness level of "B", which means improvements are required.

Primary hNPC proliferation assay (NPC1)

Various assays are available to study KEs belonging within the fetal phase of brain development. Exemplary are the NPC tests, the PeriTox, and the NeuriTox assay. NPC proliferation is a fundamental neurodevelopmental KE that, when disturbed, like in Zika virus-infected primary NPC, leads to microcephaly in children (Tang et al., 2016; Devakumar et al., 2018).

Proliferation of primary hNPC of fetal origin (Lonza), grown as neurospheres in 3D, is studied by measuring the increase in sphere size over 14 days using phase contrast microscopy (Baumann et al., 2014, 2015; Gassmann et al., 2010, 2012; Moors et al., 2009; Schreiber et al., 2010; Tofighi et al., 2011) and/or by measuring bromodeoxyuridine (BrdU) incorporation after 3 days *in vitro* (DIV) using a luminescence-based BrdU assay (Roche) and a luminometer (Baumann et al., 2014, 2015). Briefly, neurospheres with a diameter of 300 µm are plated one sphere/well in a 96-well plate with or without chemical in epidermal growth factor (EGF)- and fibroblast growth factor (FGF)-containing defined medium. For neurosphere diameter assessment, phase-contrast microscopic images are taken on plating day (day 0) as well as on days 7 and 14. Changes in sphere diameter are measured and monitored with ImageJ for each individual sphere. The same set-up is used for the BrdU assay, where BrdU incorporation into the DNA of hNPC is measured using a luminometer. The endpoint-specific control for this assay is withdrawal of growth factors, significantly reducing hNPC proliferation. This assay is part of a "high content DNT test", the "Neurosphere Assay" (NPC1-6), and is also set up with hiPSC-derived neurospheres as well as with spheres generated from prepared rat, mice or rabbit brains (Baumann et al., 2016; Barenys et al., 2017, unpublished data)

Scoring of the assay with our suggested list of criteria revealed that the system is ready (scoring A) concerning phase I. For phase II the assay also scored A, although the transferability to another laboratory is missing and the prediction model needs finalization. This is currently under development with a large data

---

[1] doi:10.14573/altex.1712081s

set. Concerning the screening issue of phase III, this test method also reaches A.

Primary hNPC migration assay (NPC2)

Cortex development takes place during the fetal phase of development. It involves radial glia migration leading to the development of a scaffold that is subsequently used by neurons as a set of "highways" to migrate and reach their final cortical destination. In contrast to rodents, human brain is gyrencephalic and radial glia composition of gyrencephalic species differs from that of non-folded brain surface species (Borrell et al., 2014). Thus, NPC migration is a fundamental neurodevelopmental KE that, when disturbed, e.g., in methylmercury exposed children, leads to alterations in cortex development (Choi et al., 1989).

Primary hNPCs of fetal origin (Lonza) are grown as neurospheres in 3D (see NPC1). Plating of size-defined (300 µm diameter) spheres on a poly D-Lysin/laminin matrix in a 96-well plate or 8-chamber slide format in the absence of growth factors initiates radial NPC migration out of the sphere. The first cells migrating out of the neurosphere display radial glia morphology and are nestin, SOX-2 and PAX-6 positive (Moors et al., 2007, 2009, 2012; Edoff et al., 2017). Their migration is dependent on laminin-integrin interaction (Barenys et al., 2017), which is also known to be crucial for radial glia migration *in vivo* (Belvindrah et al., 2007). Moreover, treatment with bone morphogenic proteins (BMPs) causes glial fibrillary acidic protein (GFAP) enrichment accompanied by morphological changes towards star-like astrocyte cell shapes (Baumann et al., 2015). These data support the concept that these cells are radial glia cells (Moors et al., 2007, 2009; Baumann et al., 2016; Edoff et al., 2017). Secondarily, neurons and oligodendrocytes arise, the former migrating on the glia carpet (Schmuck et al., 2017). Hence, this multicellular secondary 3D model (Alépée et al., 2014) can be used for measuring a) radial glia cell, b) early neuronal, and c) oligodendrocyte migration. Radial glia cell migration is measured after 24 and/or 72 h by determination of the distance the cells cover from the sphere core to the furthest migrated cell using phase contrast microscopy (Baumann et al., 2015, 2016; Gassmann et al., 2010, 2012; Moors et al., 2007, 2009; Schmuck et al., 2017; Barenys et al., 2017; Tofighi et al., 2011; Edoff et al., 2017) or applying high content image analysis (HCIA) and the Omnisphero program[4] to DAPI-stained spheres (Baumann et al., 2016; Schmuck et al., 2017). When the latter approach is multiplexed with β(III)tubulin-stained neurons or O4-stained oligodendrocytes, the Omnisphero program quantifies not only radial glia cells, but also neuronal and oligodendrocyte migration simultaneously (Schmuck et al., 2017). Migration cues differ between radial glia cells and neurons, as EGF stimulates radial glia and does not affect neuronal migration at very low concentrations, while at higher exposure levels both cell types are responsive to the EGF cue. EGF also stimulates migration *in vivo* (Puehringer et al., 2013). This assay assesses early fetal neuronal and oligodendrocyte differentiation at the same time, yet these are described as separate assays (NPC3 and NPC5, see below), as they can also be studied without migration measures. The NPC2 assay for total cell migration is also established for NPC prepared from rat, mouse or rabbit brains (Baumann et al., 2016; Barenys, unpublished data).

Scoring of the assay with our suggested criteria list revealed that the NPC2 assay is ready (scoring A) concerning phase I. For phase II the assay also scored A, only the prediction model needs finalization. This is currently under development with a large data set. Concerning the screening issue of phase III, this test method also scores A.

Primary hNPC neuronal differentiation assay (NPC3)

Primary hNPC of fetal origin (Lonza) grow as neurospheres in 3D (see NPC1). Plating of size-defined (300 µm diameter) spheres on a poly D-Lysin/laminin matrix in a 96-well plate or 8-chamber slide format in the absence of growth factors initiates radial NPC migration out of the plated sphere (NPC2) accompanied by consecutive cell differentiation into nestin+ radial glia, β(III)tubulin+ neurons and O4+ oligodendrocytes (Moors et al., 2012; Edoff et al., 2017; Baumann et al., 2014, 2015) over a period of one to five days (Schmuck et al., 2017). Neuronal cells are identified by positive β(III)tubulin staining within the migration area of each neurosphere three or five days after plating either manually or by using the Omnisphero platform (Schmuck et al., 2017; 5). With this program, DAPI-stained nuclei are identified. An algorithm specifically created for small, young neurons with short neurites identifies β(III)tubulin+ neurites and secondarily finds the belonging nucleus by its association with the skeletonized neurite. By comparing this Omnisphero algorithm to the Neuronal Profiler Bioapplication (NPBA), a program that is customized for studying neuronal morphology with the Cellomics Array Scan (Thermo Scientific), we reduced the false-positive neuronal identification rate from 40% to < 10%. NPC3 can be multiplexed with NPC4 (neuronal morphology, see below) or NPC2 (radial glia and neuronal migration, see above); in the latter, information on neuronal (β(III)tubulin+ cell) positioning is further processed to values of neuronal migration (Schmuck et al., 2017). In addition, multiplexing of NPC3 with NPC2 and NPC5 (oligodendrocyte differentiation and positioning, see below) after five days *in vitro* reveals information on neuronal and oligodendrocyte differentiation and migration within one assay (Schmuck, unpublished data).

Scoring of the NPC3 assay with our suggested criteria list revealed that the assay is ready (scoring A) concerning phase I. For phase II the assay also scored A, only the prediction model needs finalization. This is currently under development with a large data set. Concerning the screening issue of phase III, this test method reaches level B of readiness.

Neuronal morphology (neurite number, average and total neurite length, neurite branching) of young neurons differentiated from fetal hNPC (NPC4)

The outgrowth of neurites is a major process during brain development. It is needed for the formation of dendrites and axons and is therefore a pre-requisite for connectivity of neurons. A disturbed or impaired neurite outgrowth during human brain

---

[4] http://www.omnisphero.com

development is thought to be one reason for the development of autism spectrum disorders. Therefore, this test method was developed in order to more rapidly assess chemical toxicity on the growth of neurites.

The NPC4 assay is an extension of the NPC3 assay when NPC3 is evaluated with the Omnisphero software (see above) because it quantifies morphological measures of stained, human fetal NPC differentiated, young β(III)tubulin+ neurons. Skeletonized neurites are evaluated for their number, length and branching (Schmuck et al., 2017). The test is a HCIA assay, which has been extensively characterized with two individual software programs *versus* manual evaluation of all endpoints, and thus there is high confidence in the outcome.

Scoring of the assay NPC4 with our suggested criteria list revealed that the assay is ready (scoring A) concerning phase I. For phase II the assay scored B, and in phase III this test method reaches level C of readiness.

Oligodendrocyte differentiation (NPC5)

Primary hNPCs of fetal origin (Lonza) grow as neurospheres in 3D (see NPC1). Plating of size-defined (300 µm diameter) spheres on a poly D-Lysin/laminin matrix in a 96-well plate or 8-chamber slide format in the absence of growth factors initiates radial NPC migration out of the plated sphere (NPC2) accompanied by consecutive cell differentiation into nestin+ radial glia, β(III)tubulin+ neurons and O4+ oligodendrocytes (Moors et al., 2012; Edoff et al., 2017; Baumann et al., 2014, 2015) over a period of one to five days (Schmuck et al., 2017). Oligodendrocytes are identified by positive O4 staining within the migration area of each neurosphere five days after plating either manually or by using the Omnisphero platform (Schmuck et al., 2017; 5). Thus, DAPI-stained nuclei that co-localize for the epitope O4 are identified. The number of identified O4+ oligodendrocytes divided by the number of total nuclei in the migration area reveals % of differentiated oligodendrocytes (Baumann et al., 2016; Barenys et al., 2017; Dach et al., 2017; Schmuck et al., 2017). The endpoint-specific control BMP reduces oligodendrocyte differentiation and accelerates astrocyte maturation in hNPC (Baumann et al., 2016) similar to its effects *in vivo* (Bond et al., 2012). NPC5 can be multiplexed with NPC2 (migration), NPC3 (neuronal differentiation) and NPC4 (neurite morphology).

Scoring of NPC5 with our suggested criteria list revealed that the assay is ready (scoring A) for phase I. For phase II the assay scored A, and in phase III this test method reaches level B of readiness.

Oligodendrocyte maturation – Thyroid hormone (TH) disruption assay (NPC6)

Maturation of O4+ oligodendrocytes differentiated from hNPC is studied by quantifying myelin basic protein (*MBP*) mRNA expression divided by the % O4+ cells as assessed within NPC5. This ratio is defined as the oligodendrocyte maturation quotient (Q_M). During NPC development, Q_M strongly increases upon treatment of cultures with triiodothyronine (T3; Dach et al., 2017). Human TH disruptors are identified by interfering with this process, i.e., when $Q_{M\text{ solvent control}} < Q_{M\text{ TH + compound}} < Q_{M\text{ TH}}$. Oligodendrocyte toxicants can be distinguished from TH disruptors when % oligodendrocyte decreases accompanied by no change or reduction in $Q_{M\text{ TH + compound}}$ in a concentration-dependent manner, respectively. This assay can also be performed in mouse and rat NPC, but the mode of action of TH and its disruptors is different in rodent compared to human NPC (Dach et al., 2017).

Scoring of the assay NPC6 with our suggested criteria list revealed that the assay is ready (scoring A) concerning phase I. For phase II the assay scored B, and in phase III this test method also reaches level B of readiness.

The cMINC neural crest cell migration assay (UKN2 test)

Neural crest cells differentiate during neurulation from the cells of the neural roof plate. These cells eventually give rise to over 100 different cell types in the human body, including the peripheral nervous system, melanocytes, cardiomyocytes or facial connective tissue (Huang and Saint-Jeannet, 2004). One major feature of neural crest cells is that they migrate to the different parts of the developing embryo and differentiate to the according cell type once they arrive at their final destination. A large percentage of developmental disorders (e.g., congenital heart defects, orofacial clefts, Hirschsprung's disease) is caused by NC cell (NCC) deficits. These kinds of alterations can be induced by genetic factors (Lee et al., 2009) or exposure to pharmaceuticals (e.g., valproic acid, Fuller et al., 2002) and pesticides (e.g., triadimefon, Menegola et al., 2005).

For the migration inhibition of neural crest cells (cMINC assay), human pluripotent stem cells are differentiated into HNK-1+/DLL- neural crest cells. The cells are then further expanded for up to 30 days before freezing. The differentiated cells are thawed and seeded in 96-well plates supplemented with a silicon stopper that creates a 2 mm cell-free area. Migration is initiated by removal of the stopper, and the number of viable cells is measured after 48 h (Nyffeler et al., 2017a).

The evaluation of the MINC assay revealed an A-score for readiness for phase I and III as an extensive screen, including screen confirmation, was performed using the NTP library of chemicals (Nyffeler et al., 2017b). For full readiness in phase II, the transferability into other laboratories has to be shown and further responsible pathways and AOPs are missing.

An additional feature of the assay is that other endpoints such as proliferation have been established and may be easily incorporated into standard testing.

Neuronal differentiation of pluripotent stem cells (various publications, MESn)

The UKN1 test method models early stages of embryonic neurodevelopment by the differentiation of early anterior determined NPC. However, increasing numbers of differentiation protocols are published that enable differentiation of hESC or iPSC into other neuronal cell types. Each of these cellular systems is ready in terms of academic research and could serve as a starting point to develop new toxicological test methods.

In these test systems, human ESC are differentiated directly to neurons. It is important that this stage of brain development is covered by a DNT test battery as several compounds such as ethanol, methylmercury and lead have shown to induce perturbations during this time window. The most common approach to assess morphological neuronal differentiation is by immunohistochemistry for neuronal specific proteins such as neurofilaments, β(III)tubulin and Map2. Most studies combine the imaging approach with other quantitative measurements, e.g., Western blot (protein detection) or RT-PCR (mRNA expression). Several groups have developed protocols for the differentiation process; however, there is no harmonization between these different protocols. Furthermore, very few groups have tested more than one compound or generated concentration-dependent data and the main endpoints often show effects at cytotoxic concentrations.

The performance criteria have been scored as the mean of five studies from different academic labs (He et al., 2012; Pal et al., 2011; Senut et al., 2014; Stummann et al., 2009; Talens-Visconti et al., 2011). Publications that described a promising test system but did not test any compounds were not included. The score for phase I (C) indicates that the test method needs substantial improvements to be ready; the score of phase II (D) and III (D) shows that the test method is not ready at all for application. The main shortcoming of this test method is the few compounds tested, while the test system itself is promising and relevant for DNT. Once data is generated from reference compounds, this test method would likely be useful in a DNT testing battery. Similar tests have also been developed for murine ESC (Zimmer et al., 2011a,b; Kuegler et al., 2010) and may be used for species comparison. An interesting development is also the use of a 3D hiPSC-based system that has promising toxicological performance parameters (Schwartz et al., 2015)

The NeuriTox neurite outgrowth of CNS neurons test (UKN4)

For the establishment of this test method, immortalized primary cells derived from an 8-week old mesencephalon were used (Scholz et al., 2011). These cells are kept in a progenitor status by overexpression of v-myc under the control of a TET-off promotor. Upon silencing of v-myc expression, the neuronal progenitors differentiate into mature post-mitotic neurons in 6 days. In order to assess effects of chemicals on neurite outgrowth, the differentiating cells are plated after two days of differentiation into 96-well plates and are treated for 24 h (Krug et al., 2013). Then the cells are stained with Hoechst and calcein and imaged with an automated microscope. The viable cells and the neurite area are determined by double positivity and measurement of calcein-positive pixels.

The evaluation of the UKN4 test method revealed a full readiness for phase I criteria (scoring A); for phase II the transferability of the method needs to be shown (scoring A). The cellular system including the differentiation has already been transferred into many different laboratories. A first screening was performed with the 80 compounds of the NTP library (Delp et al., 2018). In phase III this test also reached level A of readiness.

The PeriTox neurite outgrowth of PNS neurons test (UKN5)

Besides the neurite outgrowth of CNS neurons, also the neurites of PNS neurons are sensitive targets of chemicals. A prominent example is the development of neuropathies during chemotherapy with platinum compounds (Quasthoff and Hartung, 2002). In addition, acrylamide is a known toxicant that induces neuropathies in humans.

In order to differentiate immature human dorsal root ganglia cells, human pluripotent stem cells are differentiated for 8 days resulting in neural crest cells. These progenitor cells can be frozen. After thawing the neural crest cells immediately start to grow neurites. One hour after thawing the cells are treated for 24 h with test chemicals and stained with Hoechst and calcein. For imaging and quantification of viable cells and neurite area the principle is the same as in the UKN4 test method (see above) (Hoelting et al., 2016).

The evaluation revealed that the PeriTox test has a full readiness score for phase I (scoring A), whereas for phase II the transferability to another laboratory has to be shown and a final prediction model needs to be developed and confirmed (scoring B). A first screening was performed with the 80 compounds of the NTP library (Delp et al., 2018). In phase III this test received scoring A.

Development of neuronal subtypes (e.g., different neurotransmitters, NSR)

Perinatal exposure to low doses of toxicants such as lead and methylmercury can alter neuronal functions rather than leading to morphological alterations or to a net cell loss (Neal and Guillarte, 2010; Gimenez-Llort et al., 2001; Zimmer et al., 2011a,b). This effect may precede neurobehavioral and neurophysiological abnormalities that may also manifest long-term after exposure to the toxicant in later life (Tamm and Ceccatelli, 2017; Heyer and Meredith, 2017). Possible explanations concerning the molecular mechanisms are that such toxicants may interfere with expression of functionally relevant genes. Also, dysregulation of genes involved in the neurotransmitter metabolism can lead to, e.g., an altered ratio of neuronal subtypes. This might affect the patterning of the body axis or, later on, the homeostasis of the neurotransmitter system and eventually may affect neuronal function and connectivity, which could have implications in the adult organism.

Approaches used to evaluate different neuronal subtypes are based on gene and protein expression of specific marker enzymes involved in the synthesis of specific neurotransmitters (i.e., glutamate decarboxylases (GAD1), tyrosine hydroxylases, neurotransmitter transporters such as dopamine transporter (DAT), glutamate aspartate transporter (GLT) or the serotonin transporter (5-HTT)). Further, a toxicant may affect the expression of receptors of specific neurotransmitters. Profiling of relevant genes and/or proteins associated with neurotransmitter signaling have been performed on biased candidate genes by RT-qPCR (Zimmer et al., 2011a,b) and on whole transcriptome level during the maturation of neurons (Zimmer et al., 2011a,b). Together with functional endpoints, i.e., measurements of calcium flux, whole patch clamp or microelectrode arrays (see *Neuronal network formation and function* below), this provides further indication of the ability of toxicants to disrupt neuronal activity due to previously altered gene expression.

Differentiating mESC have shown some potential to address this issue at a stage where most neuronal precursors are formed and maturation of neuronal subtypes takes place. The main endpoints addressed have been differentiation of neuronal subtypes and expression of specific neurotransmitter receptors and transporters (Zimmer et al., 2011a,b; Sanchez-Martin et al., 2013). Importantly, adverse effects of tested toxicants (MeHg, Pb) on these endpoints were not related to growth inhibition or cytotoxicity (Zimmer at al., 2011a,b; Sanchez-Martin et al., 2013). Although these test systems are of murine origin, they are very useful and helpful to investigate such toxic mechanisms, especially because human systems are rare. The test system as described in Zimmer et al. (2011a,b) (NSR: neuronal subtype ratio) and in Sanchez-Martin et al. (2013) was initially not developed as a test method and therefore would need further development to fulfill the readiness criteria as suggested here. The NSR test system reached scoring C for phase I and scoring D for phase II and III.

The modern trend in toxicology is to use human cellular systems to investigate such toxic effects (Daneshian et at., 2016). So far, protocols to obtain glutamatergic, γ-aminobutyric acid (GABA)ergic, dopaminergic or region-specific neuronal subtypes from hESC and hPSC have been published (Daadi et al., 2012; Gut et al., 2013; Begum et al., 2015), although no compounds have been tested for an effect on the differentiation process.

A further trend in toxicology is to use 3-dimensional models (3D) to investigate the more complex cellular structure of the nervous system. These models are of high interest in neurotoxicology and may be an opportunity to investigate possible shifts in neuronal subtypes. Moreover, they might be good test systems for investigations of cellular composition of neural cells, including neurons and glial cells. Several human models have been developed recently using various techniques and based on different cell sources such as cell lines (Smirnova et al., 2016; Simão et al., 2016), ESCs (Lancaster et al., 2013; Sandström von Tobel et al., 2014; Sandström et al., 2017a,b), and iPSCs (Pamies, 2017a,b,c; Dang, 2016). These models have the capacity to differentiate into various neuronal subtypes and different glial cells (see *Glial cell differentiation and maturation* below), making them suitable test systems for neurotoxicity and DNT. However, very few compounds have been tested in these systems, and previously developed assays generally need to be optimized to the 3D condition. Therefore, there is currently no well-developed DNT test available using these human models. There will likely be a rapid increase in the use of these systems for DNT in the near future, especially as many groups have shown the relevance of using these systems to study neurological diseases and pathologies, e.g., Alzheimer (Choi et al., 2014, 2016), microcephaly (Lancaster et al., 2013) and Zika infections (Dang et al., 2016; Qian et al., 2016)

Neuronal maturation/neuronal network formation – Synaptogenesis (SYN)

The synapse formation assay allows to measurement of changes in the number of synapses induced by exposure to a compound that occurs during synaptogenesis. Impairment of synaptogenesis is an important KE in the existing AOPs relevant to DNT (Bal-Price and Meek, 2017; Bal-Price et al., 2017) since this key neurodevelopmental process is affected by different classes of chemicals (e.g., Shi et al., 2011; Viberg, 2009; Harrill et al., 2011a,b, 2015a,b). Several approaches exist to measure synaptogenesis *in vitro* including (i) a commercially available kit based on HCIA (Thermo Fisher Scientific), referring to previously published data (e.g., Harrill et al., 2015a,b); (ii) synapse microarrays (Shi et al., 2011) and (iii) protein (Viberg, 2009; Kim and Lee, 2012) or mRNA analyses (Laurenza et al., 2013). These assays allow quantification of presynaptic (e.g., synaptophysin, synapsin1, synaptobrevin, synaptogamin) and postsynaptic markers (PSD95, gephyrin, drebrin) at protein or mRNA levels as well as evaluation of their co-localization (HCIA).

The effects of chemicals on synapse function are routinely evaluated using whole-cell patch-clamp recording (Bal et al., 2010) or microelectrode arrays (MEA) applied to neuronal networks (e.g., Hogberg et al., 2011; Vassallo et al., 2017) as described in this report (see *Neuronal network formation and function* below).

However, to apply a synaptogenesis assay for routine chemical screening, it needs further development of the performance criteria, i.e., threshold for hits and data interpretation procedure.

The scores for phases I-III (all B) indicate that the test method is already well developed and standardized, however it still needs further optimization to fully satisfy the regulatory requirements. The test system itself is critical to complete a DNT *in vitro* testing battery.

Neuronal network formation and function (Nnff)

These methods resemble early phases of brain development during which neuronal contacts are formed and become active. A few groups have used these methods to establish effects of developmental exposure to several compounds (including MeHg, several insecticides and domoic acid) on the development of neuronal activity (Brown et al., 2016; Dingemans et al., 2016; Hogberg et al., 2011; Robinette et al., 2011). Primary cortical culture from rat neonates grown on microelectrode arrays (MEAs) that develop into spontaneously active neuronal networks over time (Cotterill et al., 2016; Brown et al., 2016; Dingemans et al., 2016; Wagenaar et al., 2006) is the most established cell model for such measurements.

However, there is not yet much harmonization between these different protocols in terms of exposure window or exposure duration. However, for at least one of these protocols, the procedure has been published with a small set of positive controls (Brown et al., 2016), and a set of 86 compounds has been screened that included 60 compounds known to cause DNT in mammals, of which nearly 82% altered at least one parameter of network formation (Frank et al., 2017). In addition to chronic/developmental exposure, neuronal networks grown on MEAs are routinely used for acute exposure studies to determine effects on neuronal network function, which by now has been done for > 1000 compounds (Strickland et al., 2017) using multiwell MEAs (mwMEAs). More recently, human iPSC-derived neuronal networks have been grown on MEAs (Tukker et al., 2016; Pamies et al., 2017a), although the degree of characterization of these human-based models and the number of compounds tested is currently limited. Regardless of the cell model used, MEAs can be multiplexed with cell viability assays such as LDH leakage, MTT and CellTiter Blue assays to distinguish neurotoxicity from cytotoxicity (Wallace et al., 2015). The scores for phase I (B), phase II (A) and phase III (B) indicate that improvements are still required to be ready, mainly regarding controls and harmonization of exposure paradigms and methods of analysis. Once done, this test method would be a useful inclusion in a DNT testing battery.

Glial cell differentiation and maturation: assays to evaluate the potential role of astrocytes, oligodendrocytes, myelination, microglia, and neuroinflammation (3Dr, 3Dh)

Regarding glial cells, two types of disturbances may occur: (a) impaired development of the respective cell type; and (b) inflammatory over-activation of glial cells during the developmental period. The latter disturbance may have long-term consequences for brain structure and function: for instance, chronic neuroinflammation triggered during brain development was shown to be associated with Alzheimer's pathology when aging (Krstic et al., 2012; Krstic and Knuesel, 2013), suggesting that the consequences of such DNT effects may only be revealed after a long asymptomatic delay (AOP-12[5]). Assays to evaluate glial differentiation (astrocytes and oligodendrocytes) can be performed in 2D or 3D rodent models. Alternatively, cells may be differentiated from human ESC or iPSC, or from neural progenitor cells (Alépée et al., 2014).

Microglial cells in the brain are derived from yolk sack myeloid progenitors (Gomez Perdiguero et al., 2013). Microglial differentiation *per se*, in the brain, has not been studied as a DNT endpoint but, since microglia have an essential role in the neuroinflammatory process and in the removal of other dying cells (Hirt et al., 2000), their reactive potential may differ depending on their maturation state or tissue environment (Sandström et al., 2017a; Lund et al., 2006). Maturation of astrocytes can be assessed by a progressive decrease of vimentin expression and a progressively increased expression of GFAP and glutamine synthase (GS), as specific markers of astrocytes (Molofsky and Deneen, 2015). Toxicity to differentiating astrocytes would lead to a decrease of GFAP or GS levels, but it could also manifest by a re-expression of vimentin and mainly by an increased expression of GFAP over control level, as a sign of astrocyte reactivity (astrocyte activation is a typical sign of neuro-inflammation).

Oligodendrocyte differentiation and maturation can be evaluated by measuring the sequential expression of markers of different stages of differentiation (i.e., first SOX10, followed by NG2 and O4, Gal-C, CNP, then MBP and finally MOG) (Rowitch, 2004). In mixed cultures, oligodendrocyte maturation can also be quantified by studying MBP expression. Completion of the myelination process can be assessed by the presence of compact myelin sheets visualized by electron microscopy (Pamies et al., 2017a).

Neuro-inflammation is mainly measured by glial reactivity, evidenced by increased expression of microglial and astrocyte specific markers (CD11b, Iba1, Isolectin B4, GFAP) and morphological changes, accompanied by increased expression and release of pro-inflammatory cytokines (IL-1β, TNF-α, IL-6). Reactive glial cells can acquire neurotoxic (M1, A1) or neuroprotective (M2, A2) phenotypes (Kigerl et al., 2009; Liddelow et al., 2017; Shinozaki et al., 2017). Development-dependent changes in the expression of M1/2 phenotype markers of microglial cells have been observed upon toxicant exposure (Sandström et al., 2017a).

Various test systems for glial differentiation (3Dr, 3Dh) have been evaluated for their readiness. The more complex 3D culture systems are required for measurements of myelination and neuro-inflammation, processes depending on complex cell-cell interactions. Using the suggested criteria list, the 3D culture systems derived from hESC or iPSC (Sandström et al., 2017b; Hogberg et al., 2013; Pamies et al., 2017a,c) were scored "B" for phase I and "C" for phase II and III. High readiness (A for phase I and II, B for phase III) was achieved by the 3D rat brain cell culture system (Monnet-Tschudi et al., 1993, 1996, 1999, 2000; Zurich et al., 2000, 2002, 2012).

Zebrafish assays

The zebrafish behavioral assays at early developmental stages (0-5 days post fertilization (dpf), considered non-animal testing according to EU legislation) have shown their potential as a whole organism approach to predict human DNT, complementary to *in vitro* assays (Nishimura et al., 2015; Padilla et al., 2011; Garcia et al., 2016; Fritsche et al., 2015). These tests may be incorporated in a test battery in different ways (Fig. 4). The behavioral endpoints are readouts that integrate early events of central nervous system (CNS) development and functioning in a metabolically competent *in vivo* model system. Zebrafish brain development, anatomical features such as the blood-brain barrier, and physiology of early life stages are well described (Fleming et al., 2013; Mueller and Wullimann, 2016; Schmidt et al., 2013), while genetic and functional homology with humans has been demonstrated (Howe et al., 2013; Khan et al., 2017; Parker et al., 2013). Many brain subdivisions found in the developing mammalian brain are identifiable in the developing zebrafish, and neurotransmitters including GABA, glutamate, serotonin, dopamine, noradrenalin, and acetylcholine are found in the neurons of zebrafish at 1-5 dpf with spatio-temporal expression highly consistent with those in the mouse (Panula et al., 2010).

The zebrafish genome has been mapped and approximately 70%-80% of zebrafish genes share homology with the human genome, and 82% of genes associated with disease in humans can be related to at least one zebrafish orthologue (Howe et al., 2013).

The stereotypic motor activity of the developing zebrafish includes three sequentially appearing behaviors that are in line with neurodevelopment: a transient period of alternating tail coiling, followed by responses to touch, and the appearance of organized free swimming of larvae (Nishimura et al., 2015). Behavioral assays for DNT in zebrafish include one or more of these three basic behaviors (Chen et al., 2012b; He et al., 2016; Selderslaghs et al., 2010, 2013; Jin et al., 2016) or some variants including a light stimulus in the photomotor response test (PMR) or light/dark challenge (Ali et al., 2012; Jarema et al., 2015; Noyes et al., 2015). These behaviors appear comparable at a functional level with human behavior, with links to neural circuitry underlying

---

[5] https://aopwiki.org/aops/12

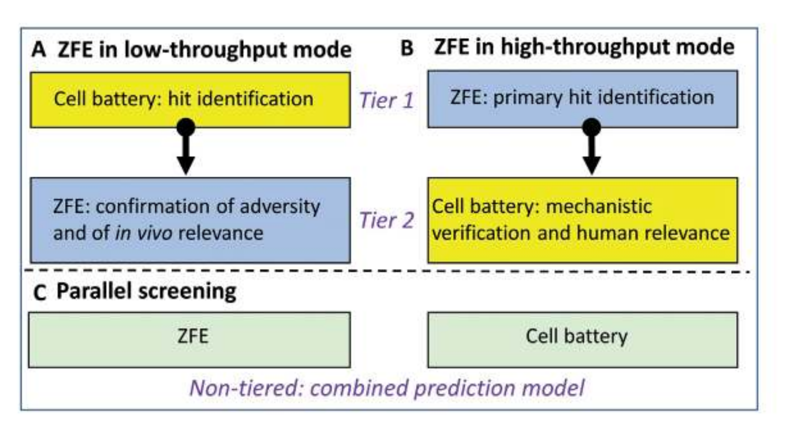


Fig. 4: Incorporation of ZFE model in a low- and high-throughput mode battery of tests
The zebrafish embryo (ZFE) test may be incorporated in various ways into a DNT test battery, depending on resources, lab automation and the purpose of testing. If ZFE testing allows only low-throughput, it may be used as second tier to further examine hits from other *in vitro* tests by a more complex whole-animal based test. Conversely, ZFE testing available as a high-throughput system may be used to identify primary hits that are further characterized and/or confirmed for human relevance by human cell-based *in vitro* tests. As a third approach, ZFE testing may be run in parallel with *in vitro* tests to feed data into an overall decision model.

the basic form of behavioral regulation. Consistent with mammals, neural networks generate, e.g., periodic motor commands for rhythmic movements and visual challenges can result in anxiety-like behavioral effects (Nishimura et al., 2015).

Many different zebrafish behavioral assays were reviewed by Legradi et al. (2015) and Planchart et al. (2016), concluding that there is a need for a harmonized protocol with recommendations for, e.g., inclusion of embryo teratogenic endpoints, positive and negative controls, and a standard exposure scenario. Nevertheless, the robustness of the behavioral endpoints has been demonstrated through comparison among different assays of a small number of chemicals (i.e., three compounds: ethanol, valproate and pentylenetetrazole) in 7, 3 and 4 studies, respectively giving similar results (Legradi et al., 2015).

The scoring for readiness considered aforementioned publications, covering screening of between 1 and 60 compounds, demonstrating compliance with a majority of performance criteria with B for phase I and II, and A for phase III. Zebrafish behavioral analyses are promising tools, complementary to cellular assays, which will benefit from further protocol harmonization and defining screening hits. The behavioral assays might be strengthened through inclusion of mechanism-based assays (axon growth, gene expression profiles, neurotransmitter activity) in relation to observed adverse outcomes (Chen et al., 2012a; He et al., 2016; Jin et al., 2016) and link to other human-based cellular model systems within the DNT battery.

ReNcell CX-based proliferation assay
ReNcell CX cells (Millipore, Temecula, CA) are a myc-immortalized cell line derived from a 14-week gestation human fetal cortex growing as a monolayer. For the proliferation assay, cells are plated in laminin-coated 96-well plates. ReNcell CX cell proliferation is determined by quantifying DNA replication using the Cellomics BrdU Cell Proliferation Kit for high-content screening (Thermo- Fisher Scientific, Pittsburgh, PA) using the Cellomics ArrayScan. Proliferation is assessed after 4, 24, and 48 h of compound treatment in a high content format (Breier et al., 2008; Radio et al., 2015).

4 Key neurodevelopmental processes covered by a battery of DNT *in vitro* assays

Over the last decade, there has been a thorough effort to identify neurodevelopmental KEs that are essential for brain development (Fig. 1) and can reliably be tested in an *in vitro* assay format. This task is complex as the developmental period of the brain is the longest compared to other organs – spanning from post-conceptual week four until the mid-20 years of age – and during the different phases of neurodevelopment various brain cell types perform distinct yet coordinated tasks. Neurodevelopmental processes in the context of timing and with a focus on human brain development are summarized in Silbereis et al. (2016), which serves as the basis for this chapter. The relevant processes are laid out here and corresponding *in vitro* assays that have the ability to detect changes in such are identified. The list of assays comprising a possible future testing battery can be found in Section 3.3 of this paper and is not repeated here. However, missing assays for certain neurodevelopmental processes are identified.

During early embryogenesis, embryonic stem cells commit to the neural lineage by becoming neural precursor cells (NPCs). These cells migrate and form the neural plate and subsequently the neural tube as the first defined structures of the brain. Later during development, the neural tube is called the subventricular zone, the area of cell origin (Kolb and Gibb, 2011). Assays capturing effects of chemical exposure on these endpoints include development of NPCs from hESC or hiPSC and stem cell-derived rosette formation. At this time, the rosettes resemble the neuronal tube structure in a two-dimensional (2D) format (Stummann et al., 2009; Colleoni et al., 2011, 2012; Senut et al., 2014; Waldmann et al., 2017). Readouts are either morphological features of rosette formation or changes in gene expression levels below the cytotoxic threshold. On this basis, the transcriptomics-based teratogenicity index was established (Waldmann et al., 2014; Shinde et al., 2016).

During a phase of exponential growth, the neural tube expands to form the critical brain processes that establish the primary organization of the central nervous system. This involves proliferation of NPC, which can be measured with different cell systems in 2D, i.e., hESC (Talens-Visconti et al., 2011; Bai et al., 2013), hiPSC (Souza et al., 2016), myc-immortalized ReNcell CX (Breier et al., 2008; Radio et al., 2015) or 3D, i.e., NPC (Gassmann et al., 2010; Schreiber et al., 2010; Baumann et al., 2015, 2016; Barenys et al., 2017). In the neurulating embryo during neural plate formation, neural crest cells (NCCs) emerge that will later develop into cell types of various tissues (e.g., bone, cartilage, neurons, and melanocytes). For terminal specification, NCCs migrate to their loci of function (Dupin and Sommer, 2012). Disturbances in NCC migration might lead to, e.g., Wardenburg's syndrome, Hirschsprung's disease, craniofacial abnormalities like frontonasal dysplasia and others. Thus, the neural crest cell migration (MINC) assay is an important tool to study effects of chemicals on this endpoint (Dreser et al., 2015; Pallocca et al., 2016; Zimmer et al., 2012, 2014; Hirsch et al., 2017).

For development of individual brain regions and connections between parts, distinct signaling is necessary, as illustrated by brain region-specific transcriptomic profiles in developing human brains (Miller et al., 2014). For human cortical development, differences from other species like rodents include the appearance of a secondary proliferative zone that allows the massive expansion of the human cortex (Kriegstein et al., 2006; Hansen et al., 2010). Outer radial glia (oRG, or basic radial glia (bRG)) cells, which contribute the majority of human radial glia cells and reside in this outer subventricular zone, are thought to produce the greater part of human cortical neurons (Smart et al., 2002; Lewitus et al., 2013). Lack of oRG cells causes lissencephaly, a normal condition in, for example, mice, but a rare, severe brain malformation in humans. An assay which addresses RG cell migration is the human NPC2 assay (see Section 3.3; Moors et al., 2007, 2009; Barenys et al., 2017; Schmuck et al., 2017). Initially migrating cells show RG cell morphology and express nestin and GFAP. Upon BMP treatment, they develop into star-shaped, GFAP expressing astrocytes. More detailed molecular knowledge on the specific type of RG cell differentiated in these cultures will be helpful in the development of brain region-specific *in vitro* models.

The first neurons that already develop in human gestation week (GW) 4 are motoneurons (Bayer and Altman, 2007; O'Rahilly and Muller, 2006). Several methods for the generation of motor neurons from ESC have been established and characterized. With regard to neurogenesis in the context of cortical development, neocortical neurons start to arise from GW7 and, with some exceptions, the majority of neurons are formed prenatally, e.g., neocortical excitatory neuron generation ceases at GW27 (Workman et al., 2013;[6]). Here, one can distinguish between early neurogenesis creating the most essential neuronal circuits mainly from hindbrain rhombomeres (Kiecker and Lumsden, 2005) and later neurogenesis during cortex formation from RG cell populations (Borrell and Götz, 2014). As for the early neurogenesis, methods for *in vitro* neuronal differentiation from hESC or lately hiPSC are established (Stummann et al., 2009; He et al., 2012; Nash et al., 2012; Druwe et al., 2016; Pistollato et al., 2017; Zagoura et al., 2017). For later neurogenesis during corticogenesis, it seems advantageous to employ fetal cells that arise from the 2nd trimester of gestation (Hansson et al., 2000) and which form neurons from RG neural precursors as in 3D neurospheres from primary human fetal NPC, as described in the NPC3 assay (Moors et al., 2009; Baumann et al., 2015; Barenys et al., 2017) or equivalent stem cell-derived neurons with cortical features (Rigamonti et al., 2016).

During brain development, more neurons are generated than needed, and final circuits are shaped by programmed death of surplus neurons that do not reach their target area. This has been modelled in primary neurons by conditions favoring hypo-polarization (Gerhardt et al., 2001; Volbracht et al., 1999), and similarly dedicated test methods may need to be devised for human neurons (Druwe et al., 2015).

In addition to neurogenesis, neuronal migration is a hallmark of cortex formation. Neuronal migration can also be measured by multiplexing the NPC2 and the NPC3 assays (Schmuck et al., 2017). After birth, newly formed and migrated neurons develop further by massively growing out neurites, dendrites and axons, followed by synaptogenesis. These processes are indispensable for neuronal network formation. Different neuronal *in vitro* systems allow measurements of these endpoints ranging from hESC- or hiPSC-derived neuronal monoculture (Harrill et al., 2011a; Druwe et al., 2016) or mixed cultures (Zagoura et al., 2017; Pistollato et al., 2017) to 3D hESC- or hiPSC-derived mixed cultures (He et al., 2012; Rigamonti et al., 2016), LUHMES dopaminergic neuronal monocultures as in the UKN4 assay (Scholz et al., 2011) or primary hNPC-derived mixed cultures using the NPC4 assay (Schmuck et al., 2017). Synaptogenesis, however, has been quantitatively assessed in rat neurons via HCIA (Harrill et al., 2011b). As already mentioned in Section 3.3, several different methods exist to measure synaptogenesis *in vitro* quantitatively, including a commercially available kit based on HCIA (Thermo Fisher Scientific). Recently, synapsin as a pre-synaptic vesicle protein was detected by staining in hiPSC-derived mixed cultures that contain GABAergic, glutamatergic and dopaminergic neurons (Zagoura et al., 2017), however, neither synapse number nor protein expression were quantified. Functionality of synapses in these cultures was displayed by electrical activity on microelectrode arrays (MEA), i.e., spikes and bursts, but do not seem to present synchronized bursting as seen for rat primary cortical culture-derived networks (Brown et al., 2016, 2017) or hESC-derived cultures on MEA chips (Kapucu et al., 2012; Kiiski et al., 2013). Nonetheless, MEA measurements were already successfully applied for *in vitro* DNT testing during chronic exposure to domoic acid (Hogberg et al., 2011), including evaluation of different receptor subtype involvement (Hogberg and Bal-Price, 2011), MeHg

---

[6] http://translatingtime.org

(Dingemans et al., 2016) and several insecticides (Dingemans et al., 2016), and recently a set of 86 environmentally relevant chemicals (Frank et al., 2017). As briefly described in Section 3.3, neuronal morphological and functional maturation (including expression of functional receptors, ion channels, pathways involved in a range of cellular responses and defense mechanisms, etc.) can be evaluated by immunocytochemistry specific protein staining, mRNA expression or pathway specific response measurements using specific agonists or antagonists.

Human stem cell-based protocols need further optimization to improve neuronal and glial maturation in mixed cultures derived from hiPSCs, which will be able to generate reliable and reproducible neuronal network activity. Such cultures should contain various cell types, as *in vivo*, of excitatory and inhibitory synapses originating from different neuronal subtypes grown in the presence of glial cells (astrocytes, oligodendrocytes and microglia).

Indeed, besides neurons, glial cells are integral parts of the CNS, representing 50% of cells in the adult brain (Kuegler et al., 2012). Glia cell (astrocytes and oligodendrocytes) generation from RG by producing astrocyte and oligodendrocyte precursor cells generally follows neurogenesis and continues until after birth (Kleiderman et al., 2016a,b).

Astroglia differentiation is a crucial event during brain development because astrocytes create the brain environment, build up the micro-architecture of the brain parenchyma, maintain brain homeostasis, store and distribute energy substrates, control the development of neural cells, synaptogenesis and synaptic maintenance and provide defense strategies for the brain. There are different astrocyte types with different functions in the brain (Hu et al., 2016). Some *in vitro* systems recapitulate astrocyte development from hESC, hiPSC or hNPC (Talens-Visconti et al., 2011; Zagoura et al., 2017; Moors et al., 2009). There is, however, a lack of precise astrocyte molecular characterization besides the expression of GFAP or vimentin that allows understanding of the astrocyte subtypes' roles in such systems. Compound effects on astrocyte reactivity (Zagoura et al., 2017; Sandström et al., 2017b), development (Moors et al., 2010; Baumann et al., 2015) or susceptibility (Talens-Visconti et al., 2011) are just beginning to contribute to the understanding of different astrocyte subtypes and functions in human cultures *in vitro*. Much more information is available on murine primary astrocytes (Falsig et al., 2006), or the combination of murine astrocytes with human neurons (Efremova et al., 2015, 2017), and fully humanized systems can be optimized to yield similar data.

Compared to other glial subtypes, oligodendrocyte myelin production is protracted in humans (Bradl and Lassmann, 2010). Given the inhibitory action of myelin on synapse formation and neuronal network plasticity, delayed myelination prolongs the development of learning activities, memory, and complex sensory perception. This species difference in timing highlights the importance of using human cells for complex *in vitro* oligodendrocyte or myelination models. Some of the recently developed methods for multiple sclerosis research referring to oligodendrocytes are summarized in Madill et al. (2016). In addition, O4+ cells generated from human fetal NPC neurospheres can be used for oligodendrocyte formation in the NPC5 assay (Moors et al., 2009; Schreiber et al., 2010) and TH-dependent maturation evaluation in the NPC6 assay (Dach et al., 2017) as described in Section 3.3. The formation of mature myelin sheets is still challenging to obtain *in vitro* and the 3D structure is crucial for this process. The 3D rat brain cell system has one of the best-developed tests for this process (Monnet-Tschudi et al., 1999), however, the species difference is of concern. A few human models have recently been developed showing characteristic myelin sheet morphology, but the test method needs to be further developed to fulfill the criteria of the DNT test battery (Sandström et al., 2017b; Pamies et al., 2017a).

5 The status of *in vitro* testing in the field of DNT

5.1 Which chemicals have already been tested in assays that can contribute to a DNT test battery?

An alternative approach towards evaluation of test readiness would be to examine which compounds known to be associated with a DNT hazard have been correctly or incorrectly identified by NAMs. This question can only be answered conclusively by data from an entire test battery, as no single *in vitro* method covers the whole spectrum of DNT-relevant processes. A small step towards this ultimate goal would be taking stock of the available data to see which chemicals have been tested, and which gaps in chemical and biological space would need to be filled. In a subsequent step, generally applicable prediction models would need to be established in order to eventually compare the outcome of *in vitro* testing with knowledge on *in vivo* hazard.

We conducted a literature search investigating which of the 32 compounds listed by Aschner et al. (2017) as DNT toxicants have been tested *in vitro*. The outcome of our survey shows (Tab. 4) that only a few compounds (e.g., methylmercury) have been tested broadly, while for others (e.g., heroin) only limited *in vitro* data are available. However, testing this small subset of compounds will not be sufficient. There are other positive controls, and, even more importantly, large numbers of negative controls need to be identified and tested to establish good prediction models. Thus, an important task for future research activities would be to close such data gaps by encouraging the development and use of a larger test set of chemicals to be used widely within the DNT *in vitro* field.

One step in this direction would be development of a database of all compounds tested to date in DNT alternative assays. So far, a summary list of chemicals already tested has been generated (Fig. 5). This figure illustrates the current status of chemical testing in assays that could be used as part of an IATA for DNT. The table was compiled based on publications describing various assays and requesting that the lead authors of those publications report which compounds they had tested. All chemicals were mapped based on conversion of CAS#s to DSSTOX ID numbers using the EPA Chemistry Dashboard[7]. Chemicals were not considered desalted, so there may be similar desalted chemicals

---

[7] https://comptox.epa.gov/dashboard

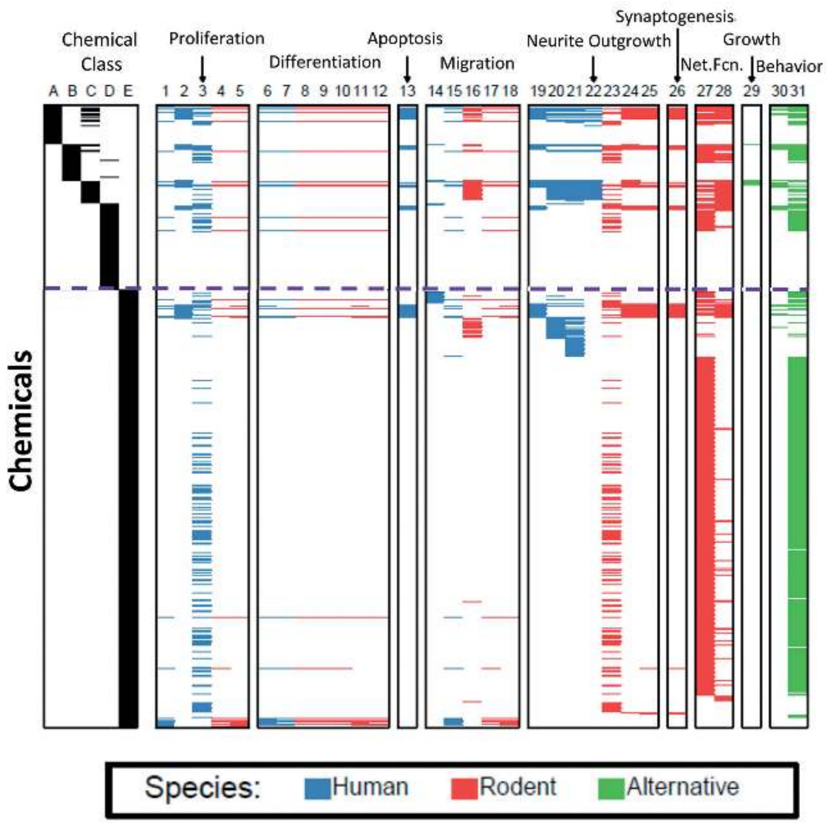


Fig. 5: The current chemical landscape of *in vitro* DNT testing

The heatmap plots chemicals as rows and test status as columns. The first 5 columns provide evidence of the class of chemicals relative to evidence of DNT or priority for testing (see main text Section 5.1). The other columns list assays grouped by neurodevelopmental processes. A brief description of each column is provided below, along with a reference or references, if available. Compounds from columns A-E that have been tested in different assays (columns 1-31), are indicated by a blue (human), red (rodent), or green (alternative species) horizontal line. It should be noted that the information on which compounds have been tested was provided by the laboratories engaged in testing, and that not all of the data for each compound/assay pair have been published. Chemical class columns:
A, Compounds with evidence of DNT from multiple laboratories (Mundy et al., 2015); B, Compounds with evidence of DNT from only one laboratory (Mundy et al., 2015); C, Compounds in the 87 chemical library supplied by the National Toxicology Program; D, Compounds subjected to the literature search in Mundy et al. (2015) that did not have evidence of DNT; E, Other compounds; primarily ToxCast compounds, but also assay positive controls and other miscellaneous compounds.
Assay columns: 1, Proliferation in human neurospheres (Baumann et al., 2016); 2, Proliferation in hNP1 neuroprogenitor cells (Mundy et al., 2010); 3, Proliferation in ReNcellCX human neuroprogenitors (Breier et al., 2008; Radio et al., 2015); 4, Proliferation in mouse neurospheres (Fritsche et al., unpublished data); 5, Proliferation in rat neurospheres (Baumann et al., 2016); 6, Neuronal differentiation in human neurospheres (Baumann et al., 2016); 7, Oligodendrocyte differentiation in human neurospheres (Fritsche et al., unpublished data); 8, Differentiation in mouse neurospheres (Fritsche et al., unpublished data); 9, Neuronal differentiation in mouse neurospheres (Fritsche et al., unpublished data); 10, Oligodendrocyte differentiation in mouse neurospheres (Fritsche et al., unpublished data); 11, Neuronal differentiation in rat neurospheres (Baumann et al., 2016); 12, Oligodendrocyte differentiation in rat neurospheres (Fritsche et al., unpublished data); 13, Apoptosis in human NP1 neural precursors (Druwe et al., 2015); 14, Migration of human neuroprogenitor cells; 15, Migration in human neurospheres (Baumann et al., 2016); 16, Migration in human neural crest cells (Nyffeler et al., 2017a,b); 17, Migration in mouse neurospheres (Fritsche et al., unpublished data); 18, Migration in rat neurospheres (Baumann et al., 2016); 19, Neurite outgrowth in human hN2 neurons. (Harrill et al., 2010); 20, Neurite outgrowth in human peripheral neuroprecursors (Hoelting et al., 2016); 21, Neurite outgrowth in LUHMES neurons (Krug et al., 2013); 22, Neurite outgrowth in human iPS-derived neurons (Ryan et al., 2016); 23, Neurite outgrowth in PC12 cells (Radio et al., 2015); 24, Neurite outgrowth in rat cortical neurons (Harrill et al., 2011a); 25, Maturation of neurites in rat cortical neurons (Harrill et al., 2011b); 26, Synaptogenesis in primary cortical neurons (Harrill et al., 2011b); 27, Neuronal network function – Acute (Strickland et al., 2017, in press); 28, Neuronal network formation – Developmental (Brown et al., 2016); 29, Feeding, larval development and reproduction in *C. elegans* (Behl et al., 2016.); 30, Zebrafish behavioral tests (Cowden et al., 2012; Padilla et al., 2011); 31 Zebrafish behavior 24 hr post-fertilization (Reif et al., 2016).

Tab. 4: Overview of testing status of DNT reference compounds, with respect to NAMs
A subset of chemicals with strong evidence for a DNT hazard *in vivo* (as described in Aschner et al., 2017) was selected. A literature search was performed to retrieve data on *in vitro* testing of these compounds. Data for 11 assays have been compiled here. Pale blue fields indicate that no clear test data have been retrieved. Green fields indicate that the compound has been examined in the respective test method and was found to show a positive effect. Orange fields indicate that the compound has been tested but did not show any effect specific to DNT. In the latter two cases, the literature evidence is indicated.
Each assay allows testing of specific DNT endpoints as indicated below: *NEP differentiation*, neural tube formation; *NPC 1-6*, NPC proliferation, radial glia migration, neuron and glia differentiation, neurite outgrowth; *ReNcell*, NPC proliferation; *UKN2 (cMINC)*, NCC migration; *UKN1*, hESC/hiPS, neuron, astrocyte and oligodendrocyte differentiation, neurite outgrowth. *UKN4 (NeuriTox)*, DA neuron differentiation, neurite outgrowth; *3Dh*, human neuron, astrocyte and oligodendrocyte differentiation, synaptogenesis, myelination, neuronal network formation; *3Dr*, rat neuron, astrocyte and oligodendrocyte differentiation, synaptogenesis, myelination, neuronal network formation, neuroinflammation; *2Dm*, murine neuron and glia differentiation, synaptogenesis, neuronal network formation; *UKN5 (PeriTox)*, neurogenesis; *Zebrafish*, brain development.
Literature as indicated by numbers in the orange and green fields: 1. Zimmer et al., 2014; 2. Dreser et al., 2015; 3. Pallocca et al., 2016; 4. Zhou et al., 2015; 5. Chattopadhyay et al., 2002; 6. Breier et al., 2008; 7. Culbreth et al., 2012; 8. Gulisano et al., 2009; 9. Monnet-Tschudi et al., 1993; 10. Tasneem et al., 2016; 11. Chow et al., 2008; 12. Selderslaghs et al., 2013; 13. Baumann et al., 2016; 14. Lee et al., 2014; 15. Krug et al., 2013b; 16. Monnet-Tschudi et al., 2000; 17. Slotkin and Seidler, 2012; 18. Crumpton et al., 2000; 19. Visan et al., 2012; 20. Dingemans et al., 2016; 21. Lee et al., 2017; 22. McCarthy et al., 2011; 23. Shang et al., 2007; 24. Harrill et al., 2011b; 25. Moors et al., 2012; 26. Ninomiya et al., 2014; 27. Bramanti et al., 2010; 28. Khor et al., 2013; 29. Perez-Gomez and Tasker, 2012; 30. Hogberg et al., 2011; 31. Tiedeken et al., 2005; 32. Palmer et al., 2012; 33. Talens-Visconti et al., 2011; 34. Nash et al., 2012; 35. Guadagnoli et al., 2016; 36. Parker et al., 2014; 37. Benninghoff et al., 2013; 38. Bai et al., 2013; 39. Slikker et al., 2015; 40. Hondebrink et al., 2017; 41. Zimmer et al., 2012; 42. Senut et al., 2014; 43. Zurich et al., 2002; 44. Monnet-Tschudi et al., 1999; 45. Dou et al., 2011; 46. Chen et al., 2012a; 47. Hoareau et al., 2006; 48. Suarez-Isla et al., 1984; 49. Kindlundh-Hogberg et al., 2010; 50. Hondebrink et al., 2016; 51. Santos-Fandila et al., 2015; 52. Stummann et al., 2009; 53. Schmuck et al., 2017; 54. Moors et al., 2009; 55. Wilson et al., 2014; 56. Pallocca et al., 2013; 57. Stiegler et al., 2011; 58. Sandström et al., 2017b; 59. Hoelting et al., 2013; 60. He et al., 2012; 61. Monnet-Tschudi et al., 1996; 62. Popova et al, 2014; 63. Yao et al., 2017; 64. Coronas et al., 2000; 65. Sandström von Tobel et al., 2014; 66. Schreiber et al., 2010; 67. Hirsch et al., 2017; 68. Xiong et al., 2012; 69. Tofighi et al., 2011; 70. Yang et al., 2014; 71. Markus et al., 2010; 72. Colleoni et al., 2011; 73. Orsolits et al., 2013; 74. Addae et al., 2012; 75. Colleoni et al., 2012; 76. Wang et al., 2015; 77. Zimmermann et al., 2015.

Cellular system	**NEP diff.**	**Neuro-spheres**	**ReNcell**	**Neural crest migration**	**hESC / hiPS based diff.**	**CNS neurons**	**3D human cell culture**	**3D rat cell culture**	**2D murine cell culture**	**PNS neurons**	**Zebra fish**
Name of assay	NEP diff.	NPC 1-6	ReNcell	UKN2 (cMINC)	UKN1	UKN4 (Neuri Tox)	3Dh	3Dr	2Dm	UKN5 (PeriTox)	ZFE
COMPOUNDS											
Arsenic				🟩 1-3					🟩 4,5		
Cadmium			🟩 6,7	🟩 2	🟩 8			🟩 9	🟩 10		🟩 11
Chlorpromazine				🟩 1							🟩 12
Chlorpyrifos		🟩 13	🟩 7		🟩 14	🟧 15		🟩 16	🟩 17-20		
Cocaine					🟩 21				🟩 22		🟩 23
Dexamethasone			🟩 6		🟩 24-26				🟩 27		🟩 28
Diphenyldantoin											
Domoic acid									🟩 29,30		🟩 31
Ethanol					🟩 32-34				🟩 35		🟩 36
Haloperidol						🟧 15			🟧 37		
Heroin											
Hexachlorophene											
Ketamine					🟩 38				🟩 39,40		
Lead			🟩 6	🟩 41	🟩 24,42			🟩 43,44	🟩 10,19		🟩 45,46
Lindane											
MAM		🟩 13							🟩 47		
Maneb											

Cellular system	**NEP diff.**	**Neuro-spheres**	**ReNcell**	**Neural crest migration**	**hESC / hiPS based diff.**	**CNS neurons**	**3D human cell culture**	**3D rat cell culture**	**2D murine cell culture**	**PNS neurons**	**Zebra fish**
Name of assay	NEP diff.	NPC 1-6	ReNcell	UKN2 (cMINC)	UKN1	UKN4 (Neuri Tox)	3Dh	3Dr	2Dm	UKN5 (PeriTox)	ZFE
COMPOUNDS											
Manganese									48		
MDMA									49,50		
Methanol											51
Methyl mercury	52	13,53,54	6	41	24,55,56	15,57	58-60	61	10,19,20, 62		12
MPTP											63
Nicotine									50,64		
Paraquat						15	58	65			
PBDE		66		3,67					68		
PCB				2	69				70		
Perfluorate-PFOA											
Perfluorate-PFOS											
Terbutaline									71		
Toluene											
Trans retinoic acid	72		6,7		24				73,74		
Triethyl tin											
Valproic acid	75	13		2,3,41					76		77

mapped in more than one place (e.g., amphetamine sulfate will be in a separate row from amphetamine hydrochloride). The first set of columns A-E, provides an idea of the compound space that has been tested. Column A lists compounds identified as having *in vivo* studies from two or more laboratories indicating the ability to cause DNT in mammals, and column B for chemicals with only one laboratory (Mundy et al., 2015). Column C is a set of ~91 high priority chemicals provided to investigators by the NIEHS National Toxicology Program. Column D is a list of chemicals from Mundy et al. (2015) for which no evidence was found for DNT. In most cases, the lack of evidence of DNT was likely due to a lack of any test data, so false negatives may be likely for the chemicals in column D. The remainder of the compounds in column E were primarily from ToxCast testing and/or assay-specific positive controls. The remaining columns group assays run by different laboratories in a manner consistent with KEs in the development of the nervous system; e.g., proliferation assays, differentiation assays, etc. If an investigator reported that a chemical had been tested in a particular assay, then it is indicated by a colored horizontal bar in the appropriate column. Note that this is an indication that the compound has been tested in a particular assay, not a determination of whether that compound was positive or negative in the assay, and that data may be published or unpublished at this time. Clearly, future work to populate a database with hit-call for these chemicals is needed.

Several important observations are immediately evident from Figure 5. First, for most assays, the total number of compounds tested is small and ranges from 25-100. A larger number of chemicals has been tested only in a smaller number of assays. Examples include ~2000 chemicals for acute network function (column 27) and zebrafish behavior (column 31); ~1000 chemicals for neural cell proliferation (column 3) and neurite outgrowth (column 23).

Importantly, there are many data gaps in the testing of compounds for which there is information about their ability to cause DNT (compounds above the dashed line). Of the compounds with evidence for DNT, there are two subsets that have not been tested in any *in vitro* assay. The first consists of a variety of compounds which could be tested, but to date have not been, including some pesticides (e.g., fenvalerate, cyhalothrin, ivermectin), metals (e.g., arsenic, manganese dioxide) and pharmaceuticals (e.g., naloxone, naltrexone, propranolol). The second untested set includes compounds that currently would be difficult to test *in vitro*, including gases (carbon monoxide, carbon disulfide), volatiles (e.g., xylenes, trichloroethylene, tetrachloroethylene) or semi-volatiles (e.g., methanol, xylenes). This latter group highlights a need for

Tab. 5: Examples of signaling pathways and disturbed neurodevelopmental processes involved in diverse neurodevelopmental pathologies
Exposure to compounds that disrupt certain signaling pathways during brain development may impair key neurodevelopmental processes resulting in diverse neurodevelopmental pathologies. This table presents a few selected examples of signaling pathway dysfunction.

Abbreviations: AKT, protein kinase B (PKB) signaling pathway; BDNF, brain derived neurotrophic factor; CREB, cAMP responsive element binding protein; ESC, embryonic stem cells; NCC, neural crest cells; NPC, neural precursor cells; PSC, pluripotent stem cells; TH, thyroid hormone.

		Exemplary signaling pathways important for normal brain development			
		CREB	**TH**	**BDNF**	**AKT**
	Neurodevelopmental pathologies associated with signaling pathway dysfunction	Decreased activity: involved in deficit of cognitive function (AOP 13, 54, 42)	Decreased levels: deficit in cognitive function (AOP 54, 42)	Decreased levels: impairment of learning and memory (AOP 12, 13, 54)	Pathway activation: human brain overgrowth syndromes in humans
Disturbed neurodevelopmental processes	*Differentiation of PSC into NPC; NPC proliferation*				**X** NPC proliferation
	NCC proliferation and migration	**X** Proliferation			**X**
	Differentiation of ESCs towards neurons	**X**			
	Radial glia migration		**X**		
	Neurite outgrowth	**X**	**X**	**X**	**X**
	Neuronal differentiation				**X**
	Synaptogenesis	**X**	**X**	**X**	
	Neuronal network formation and function		**X**	**X**	**X**
	Glial cell differentiation; myelination		**X** Oligodendrocytes, myelination		

the optimization of the experimental set up of available *in vitro* DNT test systems for reliable exposure to volatile chemicals.

Also apparent from the Figure 5 heatmap is that among the different key neurodevelopmental events, data are particularly lacking for differentiation and migration assays, while proliferation, network function and behavioral assays in zebrafish have broader coverage of compound space. Finally, of the currently available assays, none focus on glial endpoints, so there is clearly a need to develop glial-specific assays (see discussion above).

6 How can the field of NAM-based approaches to DNT testing develop in the short-term *versus* mid-term / long-term?

Here, examples are given for different types of approaches. The examples define knowledge gaps and research needs of areas that are not yet ready but have large potential.

6.1 How ready is the pathway concept for immediate use?

It is well documented that DNT compounds impair key neurodevelopmental processes leading to diverse pathologies through impairment of certain signaling pathways. As described in Fritsche et al. (2017b), signaling pathways are known to be involved in fundamental neurodevelopmental processes including NPC proliferation (e.g., BDNF-ERK-CREB, RTK-PI3K-AKT), NPC apoptosis (e.g., RXR activation, PGE2, RXR), radial glia proliferation (e.g., miRNA-17-92), neuronal and glial migration (e.g., mitogen-activated protein kinase (MAPK), PI3K, BDNF/TrkB, Reelin-Dab, PLCγ1), astrocyte differentiation (e.g., mTORC1-STAT3, Notch signaling), oligodendrocyte differentiation and myelin formation (TH), neuronal differentiation (e.g., mTORC1, BDNF-ERK-CREB, TH, PKC), synaptogenesis (e.g., NMDA receptor activation, calcium signaling, BDNF-Trk, BDNF-ERK-CREB), and neuronal network formation (e.g., PIP metabolism, TH, BDNF-TrkB, BDNF-ERK-CREB).

These pathways, if disturbed sufficiently, will lead to adverse neurodevelopmental outcomes and are therefore thought to serve as anchors for DNT *in vitro* assay development. In combination with basic information on chemical effects on signaling pathways (e.g., via ToxPi; Reif et al., 2010, 2013), DNT *in vitro* testing results concerning key neurodevelopmental processes can be used to inform AOPs on the cellular level, and will thus be fundamental for the establishment of DNT AOP networks. Some of them, such as impaired neuronal differentiation, increased neuronal apoptosis, decreased synaptogenesis, or altered neuronal network formation, have already been identified as KEs in the existing DNT AOPs (Table 2A in Bal-Price and Meek, 2017). Selected signaling pathways involved in a variety of neurodevelopmental processes are described below (Tab. 5).

CREB (cAMP responsive element binding protein) signaling pathway

The CREB pathway is crucial in the development of the central nervous system (CNS), including neuronal survival, neurite outgrowth, precursor proliferation and neuronal differentiation (Lonze and Ginty, 2002; Lesiak et al., 2014) during brain development. It regulates cell density, neuronal morphology, synaptic connectivity (e.g., potentiates transmitter release, promotes dendrogenesis), neuronal excitability, glutamatergic and GABAergic neurotransmission.

It also plays an important role in learning and memory formation through converging BDNF-ERK-CREB signaling cascades in brain development, especially amygdala dependent learning and neuronal plasticity (Ehrlich and Josselyn, 2016). CREB acts as an effector of multiple signaling cascades to transmit signals from the synapse to the nucleus, affecting transcription of plasticity-regulated genes.

A wide range of stimuli can activate CREB signaling in neurons, including hormones, neurotransmitters, growth factors and Ca^{2+}, but also stress (Lonze and Ginty, 2002). In addition, CREB is a phosphorylation target of AKT, which is activated by BDNF and tyrosine kinase B (TrkB) receptors via the phosphatidylinositol-3 kinase (PI3K) pathway. Phosphorylation of CREB allows it to interact with transcriptional coactivators to promote transcription of genes enabling structural and functional plasticity of neurons (Ehrlich and Josselyn, 2016).

Due to the variety of their functions, CREB as well as BDNF and ERK have been linked to a range of psychiatric disorders including autism spectrum disorders. The relevance of the CREB pathway for neurotoxicity has been demonstrated, showing that perturbation of the CREB signaling pathway leads to neurotoxicity (Schuh et al., 2002; Zuo et al., 2009; Brunelli et al., 2012) including *in vivo* DNT upon exposure to fluoride and arsenic (Zhu et al., 2017), lead (Toscano et al., 2002), paraquat+maneb (Li et al., 2016) and using human PSC-derived mixed neuronal/glial cultures (Pistollato et al., 2014).

BDNF (Brain-derived neurotrophic factor) signaling pathway

The neurotrophin BDNF plays an important role during brain development. BDNF is critical for the formation of appropriate synaptic connections in the brain since it regulates dendritic morphogenesis and axon guidance and its growth (reviewed in Park and Poo, 2013). Responses of growing axons to extracellular gradients of BDNF trigger activation of the PI3K, MAPK and phospholipase C-γ (PLC-γ) (for review see Huang and Reichardt, 2003; Huber et al., 2003).

The biological functions of BDNF are mediated by binding to TrkB receptor, which leads to the activation of three major intracellular signaling pathways, including MAPK, PI3K and PLCγ1 (Soulé et al., 2006). TrkB-mediated signaling regulates gene transcription in the nucleus through the activation of several transcription factors that regulate neurite outgrowth, synaptogenesis, synapse maturation, stabilization and synaptic plasticity (Nelson and Alkon, 2015; Nagappan and Lu, 2005). Experimental evidence shows that loss of BDNF through transgenic models or pharmacological manipulation leads to impaired long-term potentiation (LTP) (Monteggia et al., 2004) and decreased learning and memory (Lu et al., 2008). The important role for BDNF in LTP and learning and memory is suggested from numerous studies in rodents. Hippocampal LTP is impaired in mice lacking BDNF in their neurons, and BDNF enhances LTP in the hippocampus and visual cortex (reviewed in Mattson, 2008).

In humans, a common single-nucleotide polymorphism in the BDNF gene results in poor performance on learning and memory tasks and contributes to the pathogenesis of depression and anxiety disorders (reviewed in Cohen and Greenberg, 2008). Similarly, transgenic mice with this mutation display deficits in learning and memory tasks as well as anxiety-related behaviors (reviewed in Cohen and Greenberg, 2008). BDNF has also been shown to play a pivotal role in a variety of learning paradigms in a variety of animal models such as mice, monkeys, zebra finches and chicks (reviewed in Tyler et al., 2002). It is suggested that BDNF, ERK and CREB play an important role in neuronal plasticity through regulation of gene expression to adapt to environmental changes.

As documented in DNT AOP 13[8] (AOP-Wiki: Chronic binding of antagonist to N-methyl-D-aspartate receptors (NMDARs) during brain development induces impairment of learning and memory abilities) and AOP 54[9] (Inhibition of Na+/I-symporter (NIS) leads to learning and memory impairment), a reduced level of BDNF has been defined as the upstream KE that triggers downstream KEs such as reduced presynaptic glutamate release, increased neuronal cell death, and aberrant dendritic morphology, leading to decreased synaptogenesis and decreased neuronal network function resulting in impairment of learning and memory in children, the adverse outcome in these two AOPs.

Experimental support for a relationship between reduced BDNF levels and affected downstream KE can be triggered by lead exposure as described in detail in the DNT AOP 13.

TH (thyroid hormone) signaling pathway

The thyroid hormones (TH) triiodothyronine (T3) and thyroxine (T4) are essential for brain development, maturation, and function as they regulate the early key developmental processes such as neurogenesis, neuronal migration, proliferation, myelination and neuronal and glial differentiation and maturation (de Escobar et al., 2004; Bernal, 2015). Normal human brain development and thus cognitive function rely on sufficient TH presence during the perinatal period.

The developing brain depresses neurogenesis, and TH administration stimulates it. T3 acts through TRα1 nuclear receptor to

---

[8] https://aopwiki.org/aops/13

[9] https://aopwiki.org/aops/54

increase the commitment of neural stem cells to migrating neuroblasts. Neuronal migration in the cerebral cortex, hippocampus and cerebellum is extremely sensitive to TH, and even minor deficiencies are associated with migration defects (Berbel et al., 2001). Among possible mechanisms is the action on the radial glia. The radial glia extend long processes to the cerebral wall, providing a scaffold that serves for cell migration. Maturation of radial glia in the fetal rat brain is delayed in the hippocampus of hypothyroid rats. Thyroid hormones may influence neuronal migration in the cerebral cortex through regulation of the expression of the *Reln* gene in interneurons.

TH also controls the expression of many genes encoding proteins with roles in terminal neuronal and glial differentiation (Morte et al., 2010). Among them are cell cycle regulators, cytoskeletal proteins, neurotrophins and neurotrophin receptors and extracellular matrix proteins. A striking phenotype in the hypothyroid neonatal brain is a reduction in myelination (Adamo et al., 1990) as TH is involved in oligodendrocyte differentiation (Nygard et al., 2003). After prolonged neonatal hypothyroidism, the number of myelinated axons in adult rats is abnormally low, which corresponds with decreased expression of the major constituents of myelin (myelin basic protein (MBP), proteolipid protein (Plp), 2', 3'-cyclic nucleotide 3'-phosphodiesterase (CNPase) and myelin associated glycoprotein (MAG) (Bernal, 2015).

In humans, developing brain hypothyroidism based on TH transporter mutations that cause a lack of TH uptake through the blood-brain-barrier into the developing brain causes severe neurodevelopmental deficits as seen in the Allan-Herndon-Dudley Syndrome. These patients show delayed myelination due to less oligodendrocyte formation or maturation or a combination of both (Tonduti et al., 2013; López-Espíndola et al., 2014). Hence, neurodevelopmental effects due to disturbance of TH homeostasis can be due to either systemic TH disruption, i.e., due to thyroid dysfunction or altered TH metabolism, or both. These differences in modes-of-action need consideration when studying TH disruption *in vitro*.

With regard to local TH disrupting effects on developing brain cells, TH effects on O4+ oligodendrocyte formation and maturation was recently studied in human and mouse NPC differentiating into three major brain cell types, neurons, oligodendrocytes and astrocytes. While TH stimulates formation and maturation of mouse NPC-derived O4+ cells *in vitro*, TH guides only oligodendrocyte maturation in the human *in vitro* system. The suspected TH disruptor BDE-99 disrupted TH-dependent O4+ cell maturation only in mouse NPC, while it reduced generation of human O4+ cells independent of TH signaling in human NPC (Dach et al., 2017). This work proposed the "oligodendrocyte maturation assay" as a test for distinguishing between human neural TH disruptors and oligodendrocyte toxicants (Dach et al., 2017).

As described in DNT AOPs (AOP 54: Inhibition of Na+/I- symporter (NIS) leads to learning and memory impairment[9] and AOP 42: Inhibition of thyroperoxidase (TPO) and subsequent adverse neurodevelopmental outcomes in mammals[10]), a reduced level of TH in the blood results in lower TH levels in the brain, which leads to alterations in gene expression and subsequent protein levels (e.g., decreased levels of BDNF) that are associated with alterations in neuroanatomical structures and physiological functions, which ultimately lead to impairment of cognitive function (AO). This has been shown for chemicals that inhibit NIS (e.g., perchlorate) or TPO (e.g., propylthiouracil, methimazole). Experimental support for a relationship between decreased TH levels and KEs that lead to this AO is described in detail in the AOP-Wiki. Recently the OECD published a scoping document where currently available *in vitro* and *ex vivo* assays for evaluation of disturbance of thyroid functions, including TH signaling pathways, are characterized (OECD, 2014a)

AKT signaling pathway

AKT regulates a variety of general cellular processes, including cell proliferation and growth, autophagy, apoptosis and migration. AKT activity is hereby steered by RTK-PI3K-stimulation, with RTK-PI3K-AKT further activating the mammalian target of rapamycin (mTOR; Hennessy et al., 2005; Yu and Cui, 2016; Zheng et al., 2011), glycogen synthase kinase 3β (GSK3β), and β-catenin (Manning and Toker, 2017; Fang et al., 2007). The pivotal role of this RTK-PI3K-AKT signaling pathway in brain development is well established because dysregulation of this assembly in either direction leads to several neurodevelopmental diseases, such as megalocephaly, microcephaly, autism spectrum disorders, intellectual disability, schizophrenia, and epilepsy (reviewed in Hevner et al., 2015; Wang et al., 2017). On the cellular level, elevation of the PI3K-AKT-mTOR signaling pathway stimulates NPC proliferation, neuronal hypertrophy, and excessive dendritic branching, whereas suppression has the opposite consequences (Costa-Mattioli and Monteggia 2013; Huber et al., 2015; Lipton and Sahin, 2014; Zhou and Parada, 2012).

In the organism, AKT is represented by three isoforms, AKT1, AKT2 and AKT3 in a tissue-specific manner. The effects of altered AKT1-3 abundance in mouse brains (Easton et al., 2005) as well as transgenic modulations of AKT1 and 3 in mice (Easton et al., 2005; Tschopp et al., 2005; Tokuda et al., 2011) indicate that AKT1 and 3 are the isoforms mainly responsible for guidance of neurodevelopmental processes. AKT3 knockout mice display a selective reduction in brain size (Easton et al., 2005; Tschopp et al., 2005), whereas mice with an activating AKT3 mutation have larger brains and a thicker corpus callosum. AKT1 deficiency also leads to decreased brain size, however, by a distinct mechanism: while Akt3 -/- mutants display a reduction in both cell size and cell number, Akt1 -/- mice only show reduced cell numbers (Easton et al., 2005).

In human fetal brains, AKT3 expression is by far overrepresented compared to the two other isoforms (Wu et al., 2009), pointing to a major involvement of AKT3 in human brain development. The significance of this RTK-PI3K-AKT pathway for human brain development *in vivo* is demonstrated by the

---

[10] https://aopwiki.org/aops/42

neurodevelopmental effects of mutations overstimulating its signaling. These can be grouped into mutations causing overstimulation of RTK (Cohen and Kreiborg, 1990; Faivre et al., 2002; Hevner, 2005; D'Ercole and Ye 2008), PI3K and AKT (Flores-Sarnat et al., 2003; Salamon et al., 2006), or AKT downstream signaling (Fraser et al., 2004; Li et al., 2002) that are responsible for diverse brain overgrowth disorders. These data strongly support the notion that compounds interfering with the RTK-PI3K-AKT-mTOR signaling cascade by stimulation or inhibition will lead to an adverse neurodevelopmental outcome.

Because of the function of AKT in regulation of brain size, NPC might be a useful cell method for studying functional effects of impaired AKT signaling. It was recently shown that the AKT signaling machinery is functional in human NPC (Iaconelli et al., 2017). In addition, neuronal differentiation models might be adequate to study AKT effects on neuronal mass and dendrite branching.

Exemplary signaling pathways and disturbed neurodevelopmental processes involved in neurodevelopmental pathologies
Based on the *in vivo* data cited above, a few examples of neurodevelopmental pathologies associated with specific pathway dysfunction that are involved in deregulation of certain neurodevelopmental processes are illustrated in Table 5. These neurodevelopmental pathologies are correlated to environmental chemical exposures as described in the relevant DNT AOPs. As shown, cognitive functional deficits (including impairment of learning and memory) in children is the most frequent adverse outcome associated with the disturbance of these selected signaling pathways and damaged neurodevelopmental processes. Most of these dysregulated neurodevelopmental processes could also be studied using the *in vitro* test methods evaluated in this manuscript.

6.2 Towards an ontology-based concept of future DNT testing

Individual alternative tests should obviously be characterized for their variability, reproducibility and transferability (Hartung et al., 2004). In addition, the biological domain of the assay and its chemical applicability domain are crucial aspects for characterization of the range and limitations of use of each assay. These technical characteristics are paramount to allow interpretation of results of any assay in any context. However, the classical approach of assessing predictive performance (predictivity, sensitivity, specificity) on the level of individual test systems needs reconsideration in view of innovative approaches that employ testing strategies involving combinations of tests rather than a single individual assay replacing an animal study (Piersma et al., 2013; Leist et al., 2014).

The concept of ontologies provides a basis for transition to a biology-based system of animal-free hazard and risk assessment (Brinkley et al., 2013). For computational toxicology, ontologies can be defined as networks of factors that are connected by their quantitative relationships. They can, for example, be used as a matrix to describe physiology from the molecular, via cellular, tissue and organ to the organism level. For toxicological application, only part of this physiological interaction network needs to be described, and the level of detail can be limited to essentials. Thus, the ontology is fit for purpose if it covers the subnetwork of adverse outcome pathways (AOP) that can be triggered by toxicant exposure (Vinken, 2013). An AOP is defined as the linear, one-directional route from molecular initiating event (MIE) triggered by a compound, via a number of causally linked KE steps from the molecular, via the cellular and tissue to the organism level, leading to a defined adverse outcome. The toxicity pathway network can be understood as a compilation of all AOPs, including their interrelationships. This may include stimulating and repressing interactions and feedback information, together describing the pathway from compound exposure to adverse effects at the organism level (Tonk et al., 2015). From this AOP network it should be possible to select a limited number of rate-limiting KEs in the network that are sufficient to predict all toxicant-induced adverse health effects. These KEs then need to be represented in a limited combination of animal-free assays. The challenge then remains to develop a computational model which combines the outcomes of these assays and translates them into a predictor of toxicity. In developmental toxicity, such models are emerging, so far describing individual developmental processes and their perturbation by chemical exposure (Kleinstreuer et al., 2013; Hutson et al., 2017; Leung et al., 2016). Thus, first steps are being taken on the way to full coverage of toxicity pathways in computational systems toxicology.

The above ontology-derived selection of *in vitro* assays can be employed in different ways, depending on available knowledge on the chemicals of interest. This information may include biological activity, physicochemical properties, structure-activity relationships of related compounds, and expected use patterns. Case by case, relevant assays can be selected and carried out in battery or tiered approaches to optimally and pragmatically collect the necessary information about affected AOPs and its consequences for hazard and risk assessment. Such flexible approaches can be described in IATA, as formulated by OECD (Tollefsen et al., 2014). IATA-based approaches are inherently flexible, are designed on the principles of "fit-for-purpose" and "case by case", and require scientific justification based on all available knowledge. It is therefore paramount that the ontology underlying these approaches be comprehensive as to monitoring all possible toxicity pathways, and be fine-tuned to model the human situation. *In vitro* assays included in IATA should ideally be based on human derived cell cultures to avoid interspecies differences (Fritsche et al., 2017a). This will increase scientific confidence in the reliability of the system as a whole as to sufficient coverage of the entire spectrum of toxicology.

The validation of such testing strategies for DNT or combinations of assays requires a novel approach. Validation studies on the predictivity of individual assays in the past were based on limited numbers of compounds, and have shown limited relevance for alternative groups of chemicals (Marx-Stoelting et al., 2009). In addition, the notion that reductionist *in vitro* assays cannot represent the complexity of the intact organism has also hampered acceptance of alternative methods (Piersma et al., 2014). In contrast, the animal study protocols, which were introduced half a century ago as models for human hazard and risk assessment, have been accepted without validation, but their introduction was based on general agreement in the scientific arena that these were the best possible models for the human situation. Likewise, one could contemplate introducing ontology-based testing strategies without the validation procedure as regards predictivity that is currently common practice for individual alternative assays. Given that ontology-based testing strategies are designed to cover the entire network of toxicological mechanisms, and moreover can be fine-tuned to human physiology, these strategies should be considered inherently superior to animal testing procedures, based on their sufficient coverage of the human biology that is targeted by toxicant exposures. Of course, these approaches are still in their infancy and need considerable further development. However, as proofs of principle emerge for defined aspects of the toxicological spectrum, these approaches merit further development in the interest of improved chemical hazard and risk assessment, using animal-free methods fine-tuned to the species of interest, which is the human.

6.3 Towards the development of Integrated Approaches to Testing and Assessment (IATA)

IATA are structured strategies that integrate and weight different types of data, based on the "fit-for-purpose" principle to address questions of hazard, safety or risk assessment within a specific regulatory decision context (Tollefsen et al., 2014). They incorporate multiple sources of information from different levels of biological organization obtained by a variety of methods ((Q)SAR, read-across, *in chemico*, *in vitro* but also human data, *ex vivo*, *in vivo*, etc. or omics technologies (e.g., proteomics, toxicogenomics, metabolomics)) (Tollefsen et al., 2014; OECD, 2016a) to assess whether the existing information is sufficient to address the purpose-specific regulatory decision.

To begin, problem formulation should be clearly defined as it will influence the IATA construction in terms of data requirements, types of testing (e.g., *in vitro*, *in chemico*, *in vivo*), non-testing methods ((Q)SAR, read-across), data integration approaches and acceptable level of uncertainty (e.g., screening and prioritization *versus* hazard or risk assessment). Taking into consideration the huge gap of knowledge (only 19 known human DNT compounds identified so far; Evans et al., 2016), the most urgent issue is to develop IATA for chemical screening and prioritization purposes (the problem formulation) that could serve as a promising tool, permitting initial identification of substances with DNT potential among thousands of non-tested chemicals to which humans are exposed. Having such data that sufficiently covers the biology of the system (especially as to toxicity pathways) will improve confidence that IATA is useful in identification of DNT compounds.

IATA construction should be initiated by gathering all existing information (human data, *in vivo*, *in vitro*, non-testing data) on a chemical that is evaluated through weight of evidence assessment based on expert judgment. However, if the existing information is not adequate to address the regulatory need (problem formulation), the IATA will identify data gaps that can be used to guide the generation of new data.

It is strongly advised that an IATA should be mechanistically informed (Tollefsen et al., 2014; Worth and Patlewicz, 2016; OECD 2016a,b), referring to the pathways of toxicity through which chemicals trigger the cascade of KE resulting in an adverse outcome. This information can be captured using the AOP framework. For some human adverse outcomes (e.g., skin sensitization), various mechanistically informed DAs have already been developed based on AOPs (AOP-informed IATA). AOP-informed IATA for skin sensitization incorporates methods anchored against KE identified in the published AOP in conjunction with non-testing approaches ((Q)SAR, read-across) (Patlewicz et al., 2014; Fitzpatrick and Patlewicz, 2017; OECD, 2016b).

Currently, in the area of DNT, only a few DNT AOPs are available. Notably, these differ clearly from adult neurotoxicity AOPs (e.g., Schildknecht et al., 2017), which supports the notion that DNT assessment requires very different approaches and concepts compared to the evaluation of toxic hazard for the adult nervous system. Further development of a sufficient number of AOPs that are relevant to DNT will take time, as more mechanisms of DNT need to be unraveled. This situation should, however, not delay development and implementation of a testing strategy such as an IATA. Therefore, it was suggested during the recent OECD/EFSA DNT workshop that besides the KEs defined in the existing DNT AOPs, the fundamental neurodevelopmental processes critical for normal brain development could serve as a basis for developing a battery of test methods for DNT testing (Fritsche et al., 2017a). This assumes that nervous system development will be impaired when key biological processes are sufficiently disturbed (Lein et al., 2005; Smirnova et al., 2014). In other words, the assays anchored to AOP KEs and key neurodevelopmental processes will serve to predict adverse DNT outcomes. Based on this assumption, readiness of *in vitro* assays anchored to these critical DNT processes (Fig. 1) have been evaluated (Tab. 3) to decide which assays are ready to be included in IATA. The information presented in Table 3 suggests that assays permitting evaluation of cell migration, proliferation, neurite outgrowth, synaptogenesis and neuronal network formation and function are ready to be used for screening purposes. The acceptable level of uncertainty for screening can be higher when compared to other regulatory purposes such as hazard or risk assessment. It is advisable that this battery of *in vitro* DNT tests is based on *in vitro* neuronal/glial models originating from human IPSCs in order to be as close as possible to human biology.

The above selected *in vitro* assays are supported by recently developed DNT AOPs (Bal-Price and Meek, 2017; Bal-Price et al., 2015b) in which impairment of these critical neurodevelopmental processes has been identified as late KEs, leading to adverse outcome, e.g., learning and memory deficit in children (AOP 13[8]; AOP 54[9]). Interestingly enough, these AOPs (Tab. 2A in Bal-Price and Meek, 2017) are triggered by various MIEs and different early KEs, but KEs close to adverse outcome such

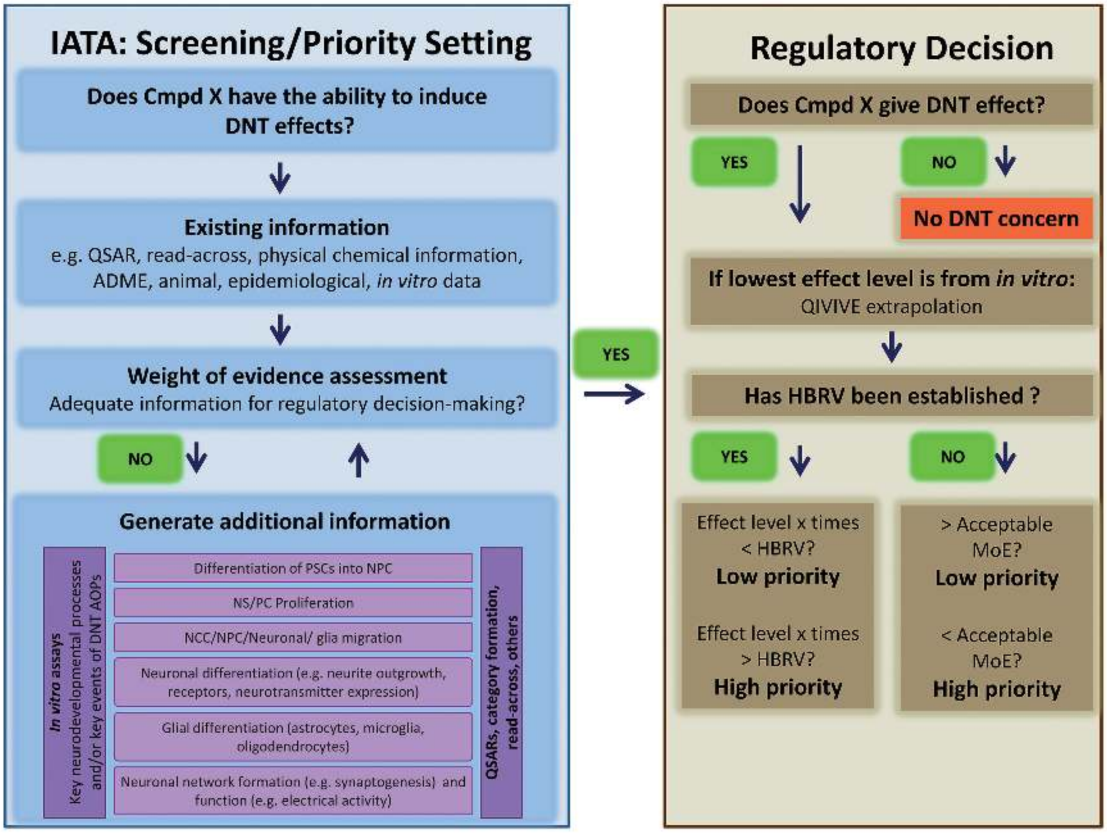


Fig. 6: An IATA designed for DNT screening/prioritization purposes

The IATA was designed for screening/prioritization purposes and was coupled to a decision tree for the DNT regulatory decision making. The IATA integrates multiple sources of existing information (human data, *in vivo*, *in vitro* and non-testing data) and guides the targeted generation of new data when required. If further testing is required, the battery of *in vitro* DNT tests that permit evaluations of key neurodevelopmental processes and KE identified in the relevant AOPs, combined with non-testing methods (e.g., QSARs and read-across) are proposed to be included in the DNT IATA for chemical screening and prioritization. Abbreviations: ADME, absorption, distribution, metabolism and excretion; Cmpd, compound; DNT, developmental neurotoxicity; HBRV, health-based reference value; MoE, margin of exposure; QIVIVE, quantitative *in vitro in vivo* extrapolation; QSAR, quantitative structure activity relationship.

as neuronal differentiation, synaptogenesis or neuronal network formation and function are shared common KEs (CKEs) in several AOPs. Therefore, the assays that permit *in vitro* evaluation of these common KEs are relevant candidates for inclusion in an IATA battery of DNT tests. The existing DNT AOPs (Bal-Price and Meek, 2017; Bal-Price et al., 2015b) provide a mechanistic understanding of the linked KEs and adverse outcomes, thus increasing scientific confidence in the relevance of the selected *in vitro* test methods and providing a mechanistic/biological context for IATA development (Tollefsen et al., 2014). Further development of AOPs relevant to DNT is strongly encouraged, as AOP-informed IATA will play a pivotal role in shifting emphasis from traditional DNT toxicity testing that is entirely based on animals to more tailored, hypothesis-based and predictive approaches taking into account existing mechanistic information at various levels of biological organization.

Since there are only few identified DNT compounds, the outlined IATA (Fig. 6, 7) is proposed for screening and prioritization of chemicals of unknown DNT effects. The first stage in that IATA workflow aims to gather existing information on chemical form and structure, the relevant route of entry, and whether it passes, e.g., the placenta or blood-brain barrier (Schultz et al., 2015). If there is not enough existing information, then the IATA refers to the scenario where new data must be generated to take a decision. The purpose of this IATA is priority setting, i.e., is the compound of DNT concern or not? This is a problem formulation relevant of course for chemicals regulated under the US Toxic Substance Control Act (TSCA) for which no data is available. However, it is also relevant in very data-rich scenarios such as pesticides, since it has been concluded that the triggers for requiring DNT studies in the pesticide regulations are not sensitive enough and do not have adequate biological

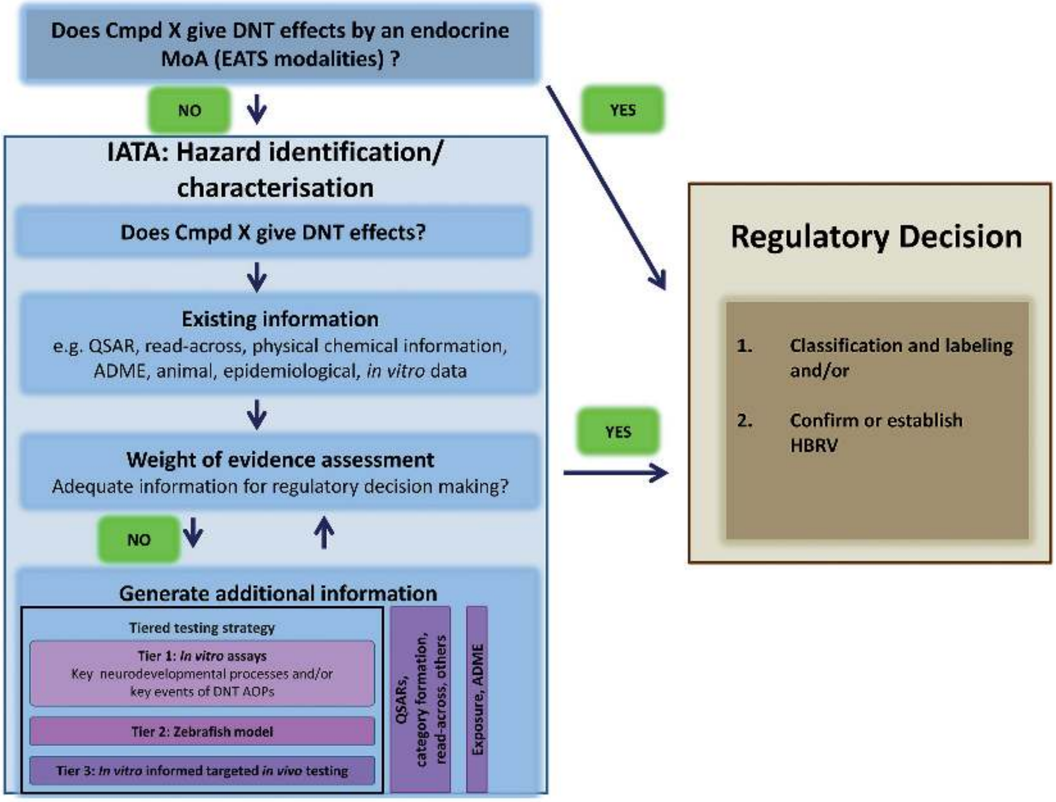


Fig. 7: Incorporation of potential endocrine effects into an IATA for DNT hazard identification/characterization
Before applying the IATA, it would be important to determine whether any DNT hazard could potentially be due to an endocrine mediated mode of action (MoA). Assays and models are in place (or under development) for regulatory purposes (for estrogen, androgen, steroid and thyroid (EATS) modalities). For the regulatory decision making, any further characterization of DNT effects by the proposed IATA should be integrated with the EATS information. Abbreviations: ADME, absorption, distribution, metabolism and excretion; AOP, adverse outcome pathway; Cmpd, compound; DNT, developmental toxicity; EATS, estrogen, androgen, steroid and thyroid modalities; HBRV, health-based reference value; IATA, integrated approach to testing and assessment; QSAR, quantitative structure activity relationship.

coverage in terms of toxicity pathways, since very different and even unique pathways operate during the development of the nervous system (Fritsche et al., 2017b). Consequently, DNT data are often not available, and therefore screening and priority setting is also warranted. In Figure 6 an outline for a decision tree is proposed. Obviously, if no effects are detected, then there is no immediate concern with regard to DNT. If DNT effects are detected in *in vitro* assay(s), then there might be a need to extrapolate the *in vitro* concentrations to *in vivo* concentrations (QIVIVE) (Yoon et al., 2012); as a default the lowest effect level should be chosen. Depending on the regulatory context, other data may be available and a health-based reference value (HBRV) may already exist (as for pesticides) or not. In both scenarios, a decision can be made on comparing the effect levels to a risk-management-defined acceptable safety margin and the compound can be deemed of low or high priority. In the latter case, further hazard and risk characterization or exposure data are required.

An IATA for DNT hazard identification and characterization is also envisaged (Fig. 7). Since DNT effects can also be mediated by endocrine modes of action (e.g., AOP 54) and assays and models are already in place to detect effects at least for estrogen, androgen and steroidogenesis modalities and partly for the thyroid (McCarthy, 2008; Bernal, 2015), it would be relevant to first establish whether such modes of action are involved. If this is not the case, then the IATA for hazard identification and characterization of non-ED mediated DNT effects should be applied. In this case, if further information is needed for regulatory decision making, a tiered testing strategy should be applied, where the *in vitro* DNT battery would be the first tests to be conducted. If further data are needed, then higher tiers would include testing in alternative species (e.g., zebrafish) and, if necessary, ultimately rodent models. In such a scenario, it is obviously crucial that there is confidence in the adequacy of the biological coverage of the first (lower) tier tests. The advantage of such an approach is that the data collected in the lower tiers could probably inform on the relevant *in vivo* testing and thus a targeted design focusing only on producing required information by applying certain selected endpoints would be adequate – thus avoiding the full-scale, costly TG 426 study. The regulatory decision has to integrate all other relevant data and if DNT effects occur this could result in proposals for classification and labelling and/or establishment of HBRVs.

For regulatory decisions, if the compound has no effect in the lower tier tests, there would most likely not be a concern if the compound is within the applicability domain of the assay/QSAR. If DNT effect(s) are observed, the lowest effect concentration from the most sensitive assay should be extrapolated into *in vivo* concentrations by quantitative *in vitro* to *in vivo* extrapolation (QIVIVE). For this, test methods and algorithms for prediction of toxicokinetic properties (not covered in this report) would be essential (e.g., Wetmore, 2015; Meek and Lipscomb, 2015). The required data do not necessarily need to be derived from animals (Daneshian et al., 2015); there are complex *in vitro* models available that predict metabolism and distribution of toxicants (e.g., Schildknecht et al., 2015; Gordon et al., 2015). There are also new high-throughput toxicokinetics models available that can be run with simple *in vitro* derived kinetics parameters (Pearce et al., 2017).

The IATA integrates multiple sources of existing information (human data, *in vivo*, *in vitro* and non-testing data) and guides the targeted generation of new data when required. In the tiered testing strategy, it is proposed to first test in the battery of *in vitro* DNT assays (see Fig. 4) and, if relevant, to further test in zebrafish assays. If further *in vivo* testing is required (rodent test), the design of these tests could be informed by the *in vitro* DNT battery / zebrafish assays and in this way a more tailored and cost-effective test than the TG 426 or TG 443 could be conducted. For any further regulatory decision making including classification and labelling and/or establishment of health-based reference doses (HBRD), the data derived from the IATA should be integrated with other effect data.

References

Adamo, A. M., Aloise, P. A., Soto, E. F. and Pasquini, J. M. (1990). Neonatal hyperthyroidism in the rat produces an increase in the activity of microperoxisomal marker enzymes coincident with biochemical signs of accelerated myelination. *J Neurosci Res 25*, 353-359. doi:10.1002/jnr.490250312

Addae, C., Yi, X., Gernapudi, R. et al. (2012). All-trans-retinoid acid induces the differentiation of encapsulated mouse embryonic stem cells into GABAergic neurons. *Differentiation 83*, 233-241. doi:10.1016/j.diff.2012.03.001

Adeleye, Y., Andersen, M., Clewell, R. et al. (2015). Implementing toxicity testing in the 21st century (TT21C): Making safety decisions using toxicity pathways, and progress in a prototype risk assessment. *Toxicology 332*, 102-111. doi:10.1016/j.tox.2014.02.007

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

Ali, S., Champagne, D. L. and Richardson, M. K. (2012). Behavioral profiling of zebrafish embryos exposed to a panel of 60 water-soluble compounds. *Behav Brain Res 228*, 272-283. doi:10.1016/j.bbr.2011.11.020

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

Aschner, M., Ceccatelli, S., Daneshian, M. et al. (2017). Reference compounds for alternative test methods to indicate developmental neurotoxicity (DNT) potential of chemicals: Example lists and criteria for their selection and use. *ALTEX 34*, 49-74. doi:10.14573/altex.1604201

Bai, X., Yan, Y., Canfield, S. et al. (2013). Ketamine enhances human neural stem cell proliferation and induces neuronal apoptosis via reactive oxygen species-mediated mitochondrial pathway. *Anesth Analg 116*, 869-880. doi:10.1213/ANE.0b013e3182860fc9

Bal, R., Erdogan, S., Theophilidis, G. et al. (2010). Assessing the effects of the neonicotinoid insecticide imidacloprid in the cholinergic synapses of the stellate cells of the mouse cochlear nucleus using whole-cell patch-clamp recording. *Neurotoxicology 31*, 113-120. doi:10.1016/j.neuro.2009.10.004

Balmer, N. V., Weng, M., Zimmer, B. et al. (2012). Epigenetic changes and disturbed neural development in a human embryonic stem cell-based model relating to the fetal valproate syndrome. *Hum Mol Genet 21*, 4104-4114. doi:10.1093/hmg/dds239

Balmer, N. V. and Leist, M. (2014). Epigenetics and transcriptomics to detect adverse drug effects in model systems of human development. *Basic Clin Pharmacol Toxicol 115*, 59-68. doi:10.1111/bcpt.12203

Balmer, N. V., Klima, S., Rempel, E. et al. (2014). From transient transcriptome responses to disturbed neurodevelopment: Role of histone acetylation and methylation as epigenetic switch between reversible and irreversible drug effects. *Arch Toxicol 88*, 1451-1468. doi:10.1007/s00204-014-1279-6

Bal-Price, A. K., Hogberg, H. T. and Buzanska, L. (2010). In vitro developmental neurotoxicity (DNT) testing: Relevant models and endpoints. *Neurotoxicology 31*, 545-554. doi:10.1016/j.neuro.2009.11.006

Bal-Price, A. K., Coecke, S., Costa, L. et al. (2012). Advancing the science of developmental neurotoxicity (DNT): Testing for better safety evaluation. *ALTEX 29*, 202-215. doi:10.14573/altex.2012.2.202

Bal-Price, A., Crofton, K. M., Leist, M. et al. (2015a). International STakeholder NETwork (ISTNET): Creating a developmental neurotoxicity (DNT) testing road map for regulatory purposes. *Arch Toxicol 89*, 269-287. doi:10.1007/s00204-015-1464-2

Bal-Price, A., Crofton, K. M., Sachana, M. et al. (2015b). Putative adverse outcome pathways relevant to neurotoxicity. *Crit Rev Toxicol 45*, 83-91. doi:10.3109/10408444.2014.981331

Bal-Price, A. and Meek, M. E. B. (2017). Adverse outcome pathways: Application to enhance mechanistic understanding of neurotoxicity. *Pharmacol Ther 179*, 84-95. doi:10.1016/j.pharmthera.2017.05.006

Bal-Price, A., Lein, P. J., Keil, K. P. et al. (2017). Developing and applying the adverse outcome pathway concept for understanding and predicting neurotoxicity. *Neurotoxicology 59*, 240-255. doi:10.1016/j.neuro.2016.05.010

Barenys, M., Gassmann, K., Baksmeier, C. et al. (2017). Epigallocatechin gallate (EGCG) inhibits adhesion and migration of neural progenitor cells in vitro. *Arch Toxicol 91*, 827-837. doi:10.1007/s00204-016-1709-8

Baumann, J., Barenys, M., Gassmann, K. and Fritsche, E. (2014). Comparative human and rat "neurosphere assay" for developmental neurotoxicity testing. *Curr Protoc Toxicol 59*, 12.21.1-24. doi:10.1002/0471140856.tx1221s59

Baumann, J., Dach, K., Barenys, M. et al. (2015). Application of the neurosphere assay for DNT hazard assessment: Challenges and limitations. In *Methods in Pharmacology and Toxicology*. Humana Press. doi:10.1007/7653_2015_49

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

Bayer, S. A. and Altman, J. (2007). *Atlas of Human Central Nervous System Development*. Volumes 1-5. CRC Press.

Begum, A. N., Guoynes, C., Cho, J. et al. (2015). Rapid generation of sub-type, region specific neurons and neural networks from human pluripotent stem cell-derived neurospheres. *Stem Cell Res 15*, 731-741. doi:10.1016/j.scr.2015.10.014

Behl, M., Hsieh, J. H., Shafer, T. J. et al. (2015). Use of alternative assays to identify and prioritize organophosphorus flame retardants for potential developmental and neurotoxicity. *Neurotoxicol Teratol 52*, 181-193. doi:10.1016/j.ntt.2015.09.003

Behl, M., Rice, J. R., Smith, M. V. et al. (2016). Editor's highlight: Comparative toxicity of organophosphate flame retardants and polybrominated diphenyl ethers to Caenorhabditis elegans. *Toxicol Sci 154*, 241-252. doi:10.1093/toxsci/kfw162

Belvindrah, R., Graus-Porta, D., Goebbels, S. et al. (2007). Beta1 integrins in radial glia but not in migrating neurons are essential for the formation of cell layers in the cerebral cortex. *J Neurosci 27*, 13854-13865. doi:10.1523/JNEUROSCI.4494-07.2007

Benninghoff, J., Grunze, H., Schindler, C. et al. (2013). Ziprasidone – not haloperidol – induces more de-novo neurogenesis of adult neural stem cells derived from murine hippocampus. *Pharmacopsychiatry 46*, 10-15. doi:10.1055/s-0032-1311607

Berbel, P., Auso, E., Garcia-Velasco, J. V. et al. (2001). Role of thyroid hormones in the maturation and organisation of rat barrel cortex. *Neuroscience 107*, 383-394. doi:10.1016/S0306-4522(01)00368-2

Bernal, J. (2015). Thyroid hormones in brain development and function. In L. J. De Groot, G. Chrousos, K. Dungan et al. (eds.), *Endotext*. South Dartmouth, MA: MDText.com, Inc. https://www.ncbi.nlm.nih.gov/books/NBK285549/

Bond, A. M., Bhalala, O. G. and Kessler, J. A. (2012). The dynamic role of bone morphogenetic proteins in neural stem cell fate and maturation. *Dev Neurobiol 72*, 1068-1084. doi:10.1002/dneu.22022

Borrell, V. and Götz, M. (2014). Role of radial glial cells in cerebral cortex folding. *Curr Opin Neurobiol 27*, 39-46. doi:10.1016/j.conb.2014.02.007

Bradl, M. and Lassmann, H. (2010). Oligodendrocytes: Biology and pathology. *Acta Neuropathol 119*, 37-53. doi:10.1007/s00401-009-0601-5

Bramanti, V., Tomassoni, D., Avitabile, M. et al. (2010). Biomarkers of glial cell proliferation and differentiation in culture. *Front Biosci (Schol Ed) 2*, 558-570. doi:10.2741/85

Breier, J. M., Radio, N. M., Mundy, W. R. and Shafer, T. J. (2008). Development of a high-throughput screening assay for chemical effects on proliferation and viability of immortalized human neural progenitor cells. *Toxicol Sci 105*, 119-133. doi:10.1093/toxsci/kfn115

Brinkley, J. F., Borromeo, C., Clarkson, M. et al. (2013). The ontology of craniofacial development and malformation for translational craniofacial research. *Am J Med Genet C Semin Med Genet 163C*, 232-245. doi:10.1002/ajmg.c.31377

Brown, J. P., Hall, D., Frank, C. L. et al. (2016). Evaluation of a microelectrode array-based assay for neural network ontogeny using training set chemicals. *Toxicol Sci 154*, 126-139. doi:10.1093/toxsci/kfw147

Brown, J. P., Lynch, B. S., Curry-Chisolm, I. M. et al. (2017). Assaying spontaneous network activity and cellular viability using multi-well microelectrode arrays. *Methods Mol Biol 1601*, 153-170. doi:10.1007/978-1-4939-6960-9_13

Browne, P., Judson, R. S., Casey, W. et al. (2017). Correction to screening chemicals for estrogen receptor bioactivity using a computational model. *Environ Sci Technol 51*, 9415-9415. doi:10.1021/acs.est.7b03317

Brunelli, L., Llansola, M., Felipo, V. et al. (2012). Insight into the neuroproteomics effects of the food-contaminant non-dioxin like polychlorinated biphenyls. *J Proteomics 75*, 2417-2430. doi:10.1016/j.jprot.2012.02.023

Chattopadhyay, S., Bhaumik, S., Nag Chaudhury, A. and Das Gupta, S. (2002). Arsenic induced changes in growth development and apoptosis in neonatal and adult brain cells in vivo and in tissue culture. *Toxicol Lett 128*, 73-84. doi:10.1016/S0378-4274(01)00535-5

Chen, J., Chen, Y., Liu, W. et al. (2012a). Developmental lead acetate exposure induces embryonic toxicity and memory deficit in adult zebrafish. *Neurotoxicol Teratol 34*, 581-586. doi:10.1016/j.ntt.2012.09.001

Chen, X., Huang, C., Wang, X. et al. (2012b). BDE-47 disrupts axonal growth and motor behavior in developing zebrafish. *Aquat Toxicol 120-121*, 35-44. doi:10.1016/j.aquatox.2012.04.014

Choi, B. H. (1989). The effects of methylmercury on the developing brain. *Prog Neurobiol 32*, 447-470. doi:10.1016/0301-0082(89)90018-X

Choi, S. H., Kim, Y. H., Hebisch, M. et al. (2014). A three-dimensional human neural cell culture model of Alzheimer's disease. *Nature 515*, 274-278. doi:10.1038/nature13800

Choi, S. H., Kim, Y. H., Quinti, L. et al. (2016). 3D culture models of Alzheimer's disease: A road map to a "cure-in-a-dish". *Mol Neurodegener 11*, 75. doi:10.1186/s13024-016-0139-7

Chow, E. S., Hui, M. N., Lin, C. C. and Cheng, S. H. (2008). Cadmium inhibits neurogenesis in zebrafish embryonic brain development. *Aquat Toxicol 87*, 157-169. doi:10.1016/j.aquatox.2008.01.019

Claudio, L., Kwa, W. C., Russell, A. L. and Wallinga, D. (2000). Testing methods for developmental neurotoxicity of environmental chemicals. *Toxicol Appl Pharmacol 164*, 1-14. doi:10.1006/taap.2000.8890

Coecke, S., Goldberg, A. M., Allen, S. et al. (2007). Workgroup report: Incorporating in vitro alternative methods for developmental neurotoxicity into international hazard and risk assessment strategies. *Environ Health Perspect 115*, 924-931. doi:10.1289/ehp.9427

Cohen, M. M. Jr. and Kreiborg, S. (1990). The central nervous system in the Apert syndrome. *Am J Med Genet 35*, 36-45. doi:10.1002/ajmg.1320350108

Cohen, S. and Greenberg, M. E. (2008). Communication between the synapse and the nucleus in neuronal development, plasticity and disease. *Annu Rev Cell Dev Biol 24*, 183-209. doi:10.1146/annurev.cellbio.24.110707.175235

Colleoni, S., Galli, C., Gaspar, J. A. et al. (2011). Development of a neural teratogenicity test based on human embryonic stem cells: Response to retinoic acid exposure. *Toxicol Sci 124*, 370-377. doi:10.1093/toxsci/kfr245

Colleoni, S., Galli, C., Gaspar, J. A. et al. (2012). Characterisation of a neural teratogenicity assay based on human ESCs differentiation following exposure to valproic acid. *Curr Med Chem 19*, 6065-6071. doi:10.2174/0929867311209066065

Coronas, V., Durand, M., Chabot, J. G. et al. (2000). Acetylcholine induces neuritic outgrowth in rat primary olfactory bulb cultures. *Neuroscience 98*, 213-219. doi:10.1016/S0306-4522(00)00143-3

Costa-Mattioli, M. and Monteggia, L. M. (2013). mTOR complexes in neurodevelopmental and neuropsychiatric disorders. *Nat Neurosci 16*, 1537-1543. doi:10.1038/nn.3546

Cotterill, E., Hall, D., Wallace, K. et al. (2016). Characterization of early cortical neural network development in multiwell microelectrode array plates. *J Biomol Screen 21*, 510-519. doi:10.1177/1087057116640520

Cowden, J., Padnos, B., Hunter, D. et al. (2012). Developmental exposure to valproate and ethanol alters locomotor activity and retino-tectal projection area in zebrafish embryos. *Reprod Toxicol 33*, 165-173. doi:10.1016/j.reprotox.2011.11.111

Crofton, K. M., Makris, S. L., Sette, W. F. et al. (2004). A qualitative retrospective analysis of positive control data in developmental neurotoxicity studies. *Neurotoxicol Teratol 26*, 345-352. doi:10.1016/j.ntt.2004.02.007

Crofton, K. M., Mundy, W. R., Lein, P. J. et al. (2011). Developmental neurotoxicity testing: Recommendations for developing alternative methods for the screening and prioritization of chemicals. *ALTEX 28*, 9-15. doi:10.14573/altex.2011.1.009

Crofton, K. M., Mundy, W. R. and Shafer, T. J. (2012). Developmental neurotoxicity testing: A path forward. *Congenit Anom (Kyoto) 52*, 140-146. doi:10.1111/j.1741-4520.2012.00377.x

Crumpton, T. L., Seidler, F. J. and Slotkin, T. A. (2000). Developmental neurotoxicity of chlorpyrifos in vivo and in vitro: Effects on nuclear transcription factors involved in cell replication and differentiation. *Brain Res 857*, 87-98. doi:10.1016/S0006-8993(99)02357-4

Culbreth, M. E., Harrill, J. A., Freudenrich, T. M. et al. (2012). Comparison of chemical-induced changes in proliferation and apoptosis in human and mouse neuroprogenitor cells. *Neurotoxicology 33*, 1499-1510. doi:10.1016/j.neuro.2012.05.012

D'Ercole, A. J. and Ye, P. (2008). Minireview: Expanding the mind: Insulin-like growth factor I and brain development. *Endocrinology 149*, 5958-5962. doi:10.1210/en.2008-0920

Daadi, M. M., Grueter, B. A., Malenka, R. C. et al. (2012). Dopaminergic neurons from midbrain-specified human embryonic stem ceel-derived neural stem cells engrafted in a monkey model of Parkinson's disease. *PLoS One 7*, e41120. doi:10.1371/journal.pone.0041120

Dach, K., Bendt, F., Huebenthal, U. et al. (2017). BDE-99 impairs differentiation of human and mouse NPCs into the oligodendroglial lineage by species-specific modes of action. *Sci Rep 7*, 44861. doi:10.1038/srep44861

Daneshian, M., Busquet, F., Hartung, T. et al. (2015). Animal use for science in Europe. *ALTEX 32*, 261-274. doi:10.14573/altex.1509081

Daneshian, M., Kamp, H., Hengstler, J. et al. (2016). Launch of a large integrated European in vitro toxicology project: EU-ToxRisk. *Arch Toxicol 90*, 1021-1024. doi:10.1007/s00204-016-1698-7

Dang, J., Tiwari, S. K., Lichinchi, G. et al. (2016). Zika virus depletes neural progenitors in human cerebral organoids through activation of the innate immune receptor TLR3. *Cell Stem Cell 19*, 258-265. doi:10.1016/j.stem.2016.04.014

de Escobar, G. M., Obregon, M. J. and del Rey, F. E. (2004). Maternal thyroid hormones early in pregnancy and fetal brain development. *Best Pract Res Clin Endocrinol Metab 18*, 225-248. doi:10.1016/j.beem.2004.03.012

Delp, J., Gutbier, S., Klima, S. et al. (2018). A high-throughput approach to identify specific neurotoxicants / developmental toxicants in human neuronal cell function assays. *ALTEX 35*, 235-253. doi:10.14573/altex.1712182

Delrue, N., Sachana, M., Sakuratani, Y. et al. (2016). The adverse outcome pathway concept: A basis for developing regulatory decision-making tools. *Altern Lab Anim 44*, 417-429.

Devakumar, D., Bamford, A., Ferreira, M. U. et al. (2018). Infectious causes of microcephaly: Epidemiology, pathogenesis, diagnosis, and management. *Lancet Infect Dis 18*, e1-e13. doi:10.1016/S1473-3099(17)30398-5

Dingemans, M. M., Schütte, M. G., Wiersma, D. M. et al. (2016). Chronic 14-day exposure to insecticides or methylmercury modulates neuronal activity in primary rat cortical cultures. *Neurotoxicology 57*, 194-202. doi:10.1016/j.neuro.2016.10.002

Dou, C. and Zhang, J. (2011). Effects of lead on neurogenesis during zebrafish embryonic brain development. *J Hazard Mater 194*, 277-282. doi:10.1016/j.jhazmat.2011.07.106

Dreser, N., Zimmer, B., Dietz, C. et al. (2015). Grouping of histone deacetylase inhibitors and other toxicants disturbing neural crest migration by transcriptional profiling. *Neurotoxicology 50*, 56-70. doi:10.1016/j.neuro.2015.07.008

Druwe, I., Freudenrich, T. M., Wallace, K. et al. (2015). Sensitivity of neuroprogenitor cells to chemical-induced apoptosis using a multiplexed assay suitable for high-throughput screening. *Toxicology 333*, 14-24. doi:10.1016/j.tox.2015.03.011

Druwe, I., Freudenrich, T. M., Wallace, K. et al. (2016). Comparison of human induced pluripotent stem cell-derived neurons and rat primary cortical neurons as in vitro models of neurite outgrowth. *Appl In Vitro Toxicol 2*, 26-36. doi:10.1089/aivt.2015.0025

Dupin, E. and Sommer, L. (2012). Neural crest progenitors and stem cells: From early development to adulthood. *Dev Biol 366*, 83-95. doi:10.1016/j.ydbio.2012.02.035

Easton, R. M., Cho, H., Roovers, K. et al. (2005). Role for Akt3/protein kinase bγ in attainment of normal brain size. *Mol Cell Biol 25*, 1869-1878. doi:10.1128/MCB.25.5.1869-1878.2005

Edoff, K., Raciti, M., Moors, M. et al. (2017). Gestational age and sex influence the susceptibility of human neural progenitor cells to low levels of MeHg. *Neurotox Res 32*, 683-693. doi:10.1007/s12640-017-9786-x

Efremova, L., Schildknecht, S., Adam, M. et al. (2015). Prevention of human dopaminergic neurodegeneration in an astrocytes co-culture system allowing endogenous drug metabolism. *Br J Pharmacol 172*, 4119-4132. doi:10.1111/bph.13193

Efremova, L., Chovancova, P., Adam, M. et al. (2017). Switching from astrocytic neuroprotection to neurodegeneration by cytokine stimulation. *Arch Toxicol 91*, 231-246. doi:10.1007/s00204-016-1702-2

EFSA Panel on Plant Protection Products and their Residues. (2013). Scientific opinion on the developmental neurotoxicity potential of acetamiprid and imidacloprid. *EFSA Journal 11*, 3471. doi:10.2903/j.efsa.2013.3471

Ehrlich, D. E. and Josselyn, S. A. (2016). Plasticity-related genes in brain development and amygdala-dependent learning. *Genes Brain Behav 15*, 125-143. doi:10.1111/gbb.12255

EPA (2016). Process for Evaluating & Implementing Alternative Approaches to Traditional in vivo Acute Toxicity Studies for FIFRA Regulatory Use. US EPA Office of Pesticide Programs, 4 February 2016. https://www.epa.gov/sites/production/files/2016-03/documents/final_alternative_test_method_guidance_2-4-16.pdf

Evans, R. M., Martin, O. V., Faust, M. and Kortenkamp, A. (2016). Should the scope of human mixture risk assessment span legislative/regulatory silos for chemicals? *Sci Total Environ 543*, 757-764. doi:10.1016/j.scitotenv.2015.10.162

Faivre, L., Gosset, P., Cormier-Daire, V. et al. (2002). Overgrowth and trisomy 15q26.1-qter including the IGF1 receptor gene: Report of two families and review of the literature. *Eur J Hum Genet 10*, 699-706. doi:10.1038/sj.ejhg.5200879

Falsig, J., Porzgen, P., Lund, S. et al. (2006). The inflammatory transcriptome of reactive murine astrocytes and implications for their innate immune functions. *J Neurochem 96*, 893-907. doi:10.1111/j.1471-4159.2005.03622.x

Fang, D., Hawke, D., Zheng, Y. et al. (2007). Phosphorylation of beta-catenin by AKT promotes beta-catenin transcriptional activity. *J Biol Chem 282*, 11221-11229. doi:10.1074/jbc.M611871200

Fitzpatrick, J. M. and Patlewicz, G. (2017). Application of IATA – A case study in evaluating the global and local performance of a Bayesian network model for skin sensitization. *SAR QSAR Environ Res 28*, 297-310. doi:10.1080/1062936X.2017.1311941

Fleming, A., Diekmann, H. and Goldsmith, P. (2013). Functional characterisation of the maturation of the blood-brain barrier in larval zebrafish. *PLoS One 8*, e77548. doi:10.1371/journal.pone.0077548

Flores-Sarnat, L., Sarnat, H. B., Dávila-Gutiérrez, G. and Alvarez, A. (2003). Hemimegalencephaly: Part 2. Neuropathology suggests a disorder of cellular lineage. *J Child Neurol 18*, 776-785. doi:10.1177/08830738030180111101

Foti, S. B., Chou, A., Moll, A. D. and Roskams, A. J. (2013). HDAC inhibitors dysregulate neural stem cell activity in the postnatal mouse brain. *Int J Dev Neurosci 31*, 434-447. doi:10.1016/j.ijdevneu.2013.03.008

Frank, C. L., Brown, J. P., Wallace, K. et al. (2017). Developmental neurotoxicants disrupt formation of cortical networks on microelectrode arrays: Screening 86 compounds in the neural network formation assay. *Toxicol Sci 160*, 121-135. doi:10.1093/toxsci/kfx169

Fraser, M. M., Zhu, X., Kwon, C. H. et al. (2004). Pten loss causes hypertrophy and increased proliferation of astrocytes in vivo. *Cancer Res 64*, 7773-7779. doi:10.1158/0008-5472.CAN-04-2487

Fritsche, E., Alm, H., Baumann, J. et al. (2015). Literature review on in vitro and alternative developmental neurotoxicity (DNT) testing methods. *EFSA Supporting Publications 12*, 4. doi:10.2903/sp.efsa.2015.EN-778

Fritsche, E., Crofton, K. M., Hernandez, A. F. et al. (2017a). OECD/EFSA workshop on developmental neurotoxicity (DNT): The use of non-animal test methods for regulatory purposes. *ALTEX 34*, 311-315. doi:10.14573/altex.1701171

Fritsche, E. (2017b). Report on Integrated Testing Strategies for the identification and evaluation of chemical hazards associated with the developmental neurotoxicity (DNT), to facilitate discussions at the Joint EFSA/OECD Workshop on DNT. ENV/JM/MONO(2017)4/ANN1. https://bit.ly/2L4PPXD

Fuller, L. C., Cornelius, S. K., Murphy, C. W. and Wiens, D. J. (2002). Neural crest cell motility in valproic acid. *Reprod Toxicol 16*, 825-839. doi:10.1016/S0890-6238(02)00059-X

Gantner, F., Leist, M., Küsters, S. et al. (1996). T-cell stimulus-induced crosstalk between lymphocytes and liver cells results in augmented cytokine-release. *Exp Cell Res 229*, 137-146. doi:10.1006/excr.1996.0351

Garcia, G. R., Noyes, P. D. and Tanguay, R. L. (2016). Advancements in zebrafish applications for 21st century toxicology. *Pharmacol Ther 161*, 11-21. doi:10.1016/j.pharmthera.2016.03.009

Gassmann, K., Abel, J., Bothe, H. et al. (2010). Species-specific differential AhR expression protects human neural progenitor cells against developmental neurotoxicity of PAHs. *Environ Health Perspect 118*, 1571-1577. doi:10.1289/ehp.0901545

Gassmann, K., Baumann, J., Giersiefer, S. et al. (2012). Automated neurosphere sorting and plating by the COPAS large particle sorter is a suitable method for high-throughput 3D in vitro applications. *Toxicol In Vitro 26*, 993-1000. doi:10.1016/j.tiv.2012.04.025

Gerhardt, E., Kügler, S., Leist, M. et al. (2001). Cascade of caspase activation in potassium-deprived cerebellar granule neurons: Targets for treatment with peptide and protein inhibitors of apoptosis. *Mol Cell Neurosci 17*, 717-731. doi:10.1006/mcne.2001.0962

Gimenez-Llort, L., Ahlbom, E., Dare, E. et al. (2001). Prenatal exposure to methylmercury changes dopamine-modulated motor activity during early ontogeny: Age and gender-dependent effects. *Environ Toxicol Pharmacol 9*, 61-70. doi:10.1016/S1382-6689(00)00060-0

Gomez Perdiguero, E., Schulz, C. and Geissmann, F. (2013). Development and homeostasis of "resident" myeloid cells: The case of the microglia. *Glia 61*, 112-120. doi:10.1002/glia.22393

Gordon, S., Daneshian, M., Bouwstra, J. et al. (2015). Non-animal models of epithelial barriers (skin, intestine and lung) in research, industrial applications and regulatory toxicology. *ALTEX 32*, 327-378. doi:10.14573/altex.1510051

Grandjean, P. and Landrigan, P. J. (2006). Developmental neurotoxicity of industrial chemicals. *Lancet 368*, 2167-2178. doi:10.1016/S0140-6736(06)69665-7

Guadagnoli, T., Caltana, L., Vacotto, M. et al. (2016). Direct effects of ethanol on neuronal differentiation: An in vitro analysis of viability and morphology. *Brain Res Bull 127*, 177-186. doi:10.1016/j.brainresbull.2016.09.013

Gulisano, M., Pacini, S. Punzi, T. et al. (2009). Cadmium modulates proliferation and differentiation of human neuroblasts. *J Neurosci Res 87*, 228-237. doi:10.1002/jnr.21830

Gut, I. M., Beske, P. H., Hubbard, K. S. et al. (2013). Novel application of stem cell-derived neurons to evaluate the time- and dose-dependent progression of excitotoxic injury. *PLoS One 8*, e64423. doi:10.1371/journal.pone.0064423

Haddow, J. E., Palomaki, G. E., Allan, W. C. et al. (1999). Maternal thyroid deficiency during pregnancy and subsequent neuropsychological development of the child. *N Engl J Med 341*, 549-555. doi:10.1056/NEJM199908193410801

Hansen, D. V., Lui, J. H., Parker, P. R. and Kriegstein, A. R. (2010). Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature 464*, 554-561. doi:10.1038/nature08845

Hansson, O., Castilho, R. F., Kaminski Schierle, G. S. et al. (2000). Additive effects of caspase inhibitor and lazaroid on the survival of transplanted rat and human embryonic dopamine neurons. *Exp Neurol 164*, 102-111. doi:10.1006/exnr.2000.7406

Harrill, J. A., Freudenrich, T. M., Machacek, D. W. et al. (2010). Quantitative assessment of neurite outgrowth in human embryonic stem cell-derived hN2 cells using automated high-content image analysis. *Neurotoxicology 31*, 277-290. doi:10.1016/j.neuro.2010.02.003

Harrill, J. A., Freudenrich, T. M., Robinette, B. L. and Mundy, W. R. (2011a). Comparative sensitivity of human and rat neural cultures to chemical-induced inhibition of neurite outgrowth. *Toxicol Appl Pharmacol 256*, 268-280. doi:10.1016/j.taap.2011.02.013

Harrill, J. A., Robinette, B. L. and Mundy, W. R. (2011b). Use of high content image analysis to detect chemical-induced changes in synaptogenesis in vitro. *Toxicol In Vitro 25*, 368-387. doi:10.1016/j.tiv.2010.10.011

Harrill, J. A., Chen, H., Streifel, K. M. et al. (2015a). Ontogeny of biochemical, morphological and functional parameters of synaptogenesis in primary cultures of rat hippocampal and cortical neurons. *Mol Brain 8*, 10. doi:10.1186/s13041-015-0099-9

Harrill, J. A., Robinette, B. L., Freudenrich, T. M. and Mundy, W. R. (2015b). Media formulation influences chemical effects on neuronal growth and morphology. *In Vitro Cell Dev Biol Anim 51*, 612-629. doi:10.1007/s11626-015-9873-3

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

Hartung, T. and Leist, M. (2008). Food for thought ... on the evolution of toxicology and phasing out of animal testing. *ALTEX 25*, 91-96. doi:10.14573/altex.2008.2.91

Hartung, T., Kavlock, R. and Sturla, S. J. (2017a). Systems toxicology II: A special issue. *Chem Res Toxicol 30*, 869. doi:10.1021/acs.chemrestox.7b00038

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

He, X., Imanishi, S., Sone, H. et al. (2012). Effects of methylmercury exposure on neuronal differentiation of mouse and human embryonic stem cells. *Toxicol Lett 212*, 1-10. doi:10.1016/j.toxlet.2012.04.011

He, X., Gao, J., Dong, T. et al. (2016). Developmental neurotoxicity of methamidophos in the embryo-larval stages of zebrafish. *Int J Environ Res Public Health 14*, 23. doi:10.3390/ijerph14010023

Hennessy, B. T., Smith, D. L., Ram, P. T. et al. (2005). Exploiting the PI3K/AKT pathway for cancer drug discovery. *Nat Rev Drug Discov 4*, 988-1004. doi:10.1038/nrd1902

Henrichs, J., Bongers-Schokking, J. J., Schenk, J. J. et al. (2010). Maternal thyroid function during early pregnancy and cognitive functioning in early childhood: The generation R study. *J Clin Endocrinol Metab 95*, 4227-4234. doi:10.1210/jc.2010-0415

Hevner, R. F. (2005). The cerebral cortex malformation in thanatophoric dysplasia: Neuropathology and pathogenesis. *Acta Neuropathol 110*, 208-221. doi:10.1007/s00401-005-1059-8

Hevner, R. F. (2015). Brain overgrowth in disorders of RTK-PI3K-AKT signaling: A mosaic of malformations. *Semin Perinatol 39*, 36-43. doi:10.1053/j.semperi.2014.10.006

Heyer, D. B. and Meredith, R. M. (2017). Environmental toxicology: Sensitive periods of development and neurodevelopmental disorders. *Neurotoxicology 58*, 23-41. doi:10.1016/j.neuro.2016.10.017

Hirsch, C., Striegl, B., Mathes, S. et al. (2017). Multiparameter toxicity assessment of novel DOPO-derived organophosphorus flame retardants. *Arch Toxicol 91*, 407-425, doi:10.1007/s00204-016-1680-4

Hirt, U., Gantner, F. and Leist, M. (2000). Phagocytosis of non-apoptotic cells dying by caspase-independent mechanisms. *J Immunol 164*, 6520-6529. doi:10.3390/ijms17122068

Hoareau, C., Hazane, F., Le Pen, G. and Krebs, M. O. (2006). Postnatal effect of embryonic neurogenesis disturbance on reelin level in organotypic cultures of rat hippocampus. *Brain Res 1097*, 43-51. doi:10.1016/j.brainres.2006.04.075

Hoelting, L., Scheinhardt, B., Bondarenko, O. et al. (2013). A 3-dimensional human embryonic stem cell (hESC)-derived model to detect developmental neurotoxicity of nanoparticles. *Arch Toxicol 87*, 721-733. doi:10.1007/s00204-012-0984-2

Hoelting, L., Klima, S., Karreman, C. et al. (2016). Stem cell-derived immature human dorsal root ganglia neurons to identify peripheral neurotoxicants. *Stem Cells Transl Med 5*, 476-487. doi:10.5966/sctm.2015-0108

Hogberg, H. T. and Bal-Price, A. K. (2011). Domoic acid-induced neurotoxicity is mainly mediated by the AMPA/KA Receptor: Comparison between immature and mature primary cultures of neurons and glial cells from rat cerebellum. *J Toxicol 2011*, 543512. doi:10.1155/2011/543512

Hogberg, H. T., Sobanski, T., Novellino, A. et al. (2011). Application of micro-electrode arrays (MEAs) as an emerging technology for developmental neurotoxicity: Evaluation of domoic acid-induced effects in primary cultures of rat cortical neurons. *Neurotoxicology 32*, 158-168. doi:10.1016/j.neuro.2010.10.007

Hogberg, H. T., Bressler, J., Christian, K. M. et al. (2013). Toward a 3D model of human brain development for studying gene/environment interactions. *Stem Cell Res Ther 4, Suppl 1*, S4. doi:10.1186/scrt365

Hondebrink, L., Verboven, A. H. A., Drega, W. S. et al. (2016). Neurotoxicity screening of (illicit) drugs using novel methods for analysis of microelectrode array (MEA) recordings. *Neurotoxicology 55*, 1-9. doi:10.1016/j.neuro.2016.04.020

Hondebrink, L., Kasteel, E. E. J., Tukker, A. M. et al. (2017). Neuropharmacological characterization of the new psychoactive substance methoxetamine. *Neuropharmacology 123*, 1-9. doi:10.1016/j.neuropharm.2017.04.035

Howe, K., Clark, M. D., Torroja, C. F. et al. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature 496*, 498-503. doi:10.1038/nature12111

Hu, X., Yuan, Y, Wang, D. and Su, Z. (2016). Heterogeneous astrocytes: Active players in CNS. *Brain Res Bull 125*, 1-18. doi:10.1016/j.brainresbull.2016.03.017

Huang, E. J. and Reichardt, L. F. (2003). Trk receptors: Roles in neuronal signal transduction. *Annu Rev Biochem 72*, 609-642. doi:10.1146/annurev.biochem.72.121801.161629

Huang, X. and Saint-Jeannet, J. P. (2004). Induction of the neural crest and the opportunities of life on the edge. *Dev Biol 275*, 1-11. doi:10.1016/j.ydbio.2004.07.033

Huber, A. B., Kolodkin, A. L., Ginty, D. D. and Cloutier, J. F. (2003). Signaling at the growth cone: Ligand-receptor complexes and the control of axon growth and guidance. *Annu Rev Neurosci 26*, 509-563. doi:10.1146/annurev.neuro.26.010302.081139

Huber, K. M., Klann, E., Costa-Mattioli, M. et al. (2015). Dysregulation of mammalian target of rapamycin signaling in mouse models of autism. *J Neurosci 35*, 13836-13842. doi:10.1523/JNEUROSCI.2656-15.2015

Hutson, M. S., Leung, M. C. K., Baker, N. C. et al. (2017). Computational model of secondary palate fusion and disruption. *Chem Res Toxicol 30*, 965-979. doi:10.1021/acs.chemrestox.6b00350

Iaconelli, J., Lalonde, J., Watmuff, B. et al. (2017). Lysine deacetylation by HDAC6 regulates the kinase activity of AKT in human neural progenitor cells. *ACS Chem Biol 12*, 2139-2148. doi:10.1021/acschembio.6b01014

Jarema, K. A., Hunter, D. L., Shaffer, R. M. et al. (2015). Acute and developmental behavioral effects of flame retardants and related chemicals in zebrafish. *Neurotoxicol Teratol 52*, 194-209. doi:10.1016/j.ntt.2015.08.010

Jin, Y., Zhu, Z., Wang, Y. et al. (2016). The fungicide imazalil induces developmental abnormalities and alters locomotor activity during early developmental stages in zebrafish. *Chemosphere 153*, 455-161. doi:10.1016/j.chemosphere.2016.03.085

Kadereit, S., Zimmer, B., van Thriel, C. et al. (2012). Compound selection for in vitro modeling of developmental neurotoxicity. *Front Biosci (Landmark Ed) 17*, 2442-2460. doi:10.2741/4064

Kapucu, F. E., Tanskanen, J. M., Mikkonen, J. E. et al. (2012). Burst analysis tool for developing neuronal networks exhibiting highly varying action potential dynamics. *Front Comput Neurosci 6*, 38. doi:10.3389/fncom.2012.00038

Khan, K. M., Collier, A. D., Meshalkina, D. A. et al. (2017). Zebrafish models in neuropsychopharmacology and CNS drug discovery. *Br J Pharmacol 174*, 1925-1944. doi:10.1111/bph.13754

Khor, Y. M., Soga, T. and Parhar, I. S. (2013). Caffeine neuroprotects against dexamethasone-induced anxiety-like behaviour in the Zebrafish (Danio rerio). *Gen Comp Endocrinol 181*, 310-315. doi:10.1016/j.ygcen.2012.09.021

Kiecker, C. and Lumsden, A. (2005). Compartments and their boundaries in vertebrate brain development. *Nat Rev Neurosci 6,* 553-564. doi:10.1038/nrn1702

Kigerl, K. A., Gensel, J. C., Ankeny, D. P. et al. (2009). Identification of two distinct macrophage subsets with divergent effects causing either neurotoxicity or regeneration in the injured mouse spinal cord. *J Neurosci 29*, 13435-13444. doi:10.1523/JNEUROSCI.3257-09.2009

Kiiski, H., Äänismaa, R., Tenhunen, J. et al. (2013). Healthy human CSF promotes glial differentiation of hESC-derived neural cells while retaining spontaneous activity in existing neuronal networks. *Biol Open 2*, 605-612. doi:10.1242/bio.20134648

Kim, Y. G. and Lee, Y. I. (2012). Differential expressions of synaptogenic markers between primary cultured cortical and hippocampal neurons. *Exp Neurobiol 21*, 61-67. doi:10.5607/en.2012.21.2.61

Kindlundh-Hogberg, A. M., Pickering, C., Wicher, G. et al. (2010). MDMA (ecstasy) decreases the number of neurons and stem cells in embryonic cortical cultures. *Cell Mol Neurobiol 30*, 13-21. doi:10.1007/s10571-009-9426-y

Kleiderman, S., Gutbier, S., Tufekci, K. U. et al. (2016a). Conversion of non-proliferating astrocytes into neurogenic neural stem cells: Control by FGF2 and IFN-gamma. *Stem Cells 34,* 2861-2874. doi:10.1002/stem.2483

Kleiderman, S., Sá, J. V., Teixeira, A. P. et al. (2016b). Functional and phenotypic differences of pure populations of stem cell-derived astrocytes and neuronal precursor cells. *Glia 64*, 695-715. doi:10.1002/glia.22954

Kleinstreuer, N., Dix, D., Rountree, M. et al. (2013). A computational model predicting disruption of blood vessel development. *PLoS Comput Biol 9*, e1002996. doi:10.1371/journal.pcbi.1002996

Kolb, B. and Gibb, R. (2011). Brain plasticity and behaviour in the developing brain. *J Can Acad Child Adolesc Psychiatry 20*, 265-276.

Kriegstein, A., Noctor S. and Martinez-Cerdeno, V. (2006). Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion. *Nat Rev Neurosci 7*, 883-890. doi:10.1038/nrn2008

Krstic, D., Madhusudan, A., Doehner, J. et al. (2012). Systemic immune challenges trigger and drive Alzheimer-like neuropathology in mice. *J Neuroinflammation 9*, 151. doi:10.1186/1742-2094-9-151

Krstic, D. and Knuesel, I. (2013). Deciphering the mechanism underlying late-onset Alzheimer disease. *Nat Rev Neurol 9*, 25-34. doi:10.1038/nrneurol.2012.236

Krug, A. K., Balmer, N. V., Matt, F. et al. (2013). Evaluation of a human neurite growth assay as specific screen for developmental neurotoxicants. *Arch Toxicol 87*, 2215-2231. doi:10.1007/s00204-013-1072-y

Kuegler, P. B., Zimmer, B., Waldmann, T. et al. (2010). Markers of murine embryonic and neural stem cells, neurons and astrocytes: Reference points for developmental neurotoxicitiy testing – A review by the translantic think tank for toxicology (t^4). *ALTEX 27*, 17-42. doi:10.14573/altex.2010.1.16

Kuegler, P. B., Baumann, B. A., Zimmer, B. et al. (2012). GFAP-independent inflammatory competence and trophic functions of astrocytes generated from murine embryonic stem cells. *Glia 60*, 218-228. doi:10.1002/glia.21257

Lancaster, M. A., Renner, M., Martin, C. A. et al. (2013). Cerebral organoids model human brain development and microcephaly. *Nature 501*, 373-379. doi:10.1038/nature12517

Laurenza, I., Pallocca, G., Mennecozzi, M. et al. (2013). A human pluripotent carcinoma stem cell-based model for in vitro developmental neurotoxicity testing: Effects of methylmercury, lead and aluminum evaluated by gene expression studies. *Int J Dev Neurosci 31*, 679-691. doi:10.1016/j.ijdevneu.2013.03.002

Lee, C. T., Chen, J., Kindberg, A. A. et al. (2017). CYP3A5 mediates effects of cocaine on human neocorticogenesis: Studies using an in vitro 3D self-organized hPSC model with a single cortex-like unit. *Neuropsychopharmacology 42*, 774-784. doi:/10.1038/npp.2016.156

Lee, G., Papapetrou, E. P., Kim, H. et al. (2009). Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature 461*, 402-406. doi:10.1038/nature08320

Lee, J. and Freeman J. L. (2014). Zebrafish as a model for investigating developmental lead (Pb) neurotoxicity as a risk factor in adult neurodegenerative disease: A mini-review. *Neurotoxicology 43*, 57-64. doi:/10.1016/j.neuro.2014.03.008

Legradi, J., el Abdellaoui, N., van Pomeren, M. and Legler, J. (2015). Comparability of behavioural assays using zebrafish larvae to assess neurotoxicity. *Environ Sci Pollut Res Int 22*, 16277-16289. doi:10.1007/s11356-014-3805-8

Lein, P., Silbergeld, E., Locke, P. and Goldberg, A. M. (2005). In vitro and other alternative approaches to developmental neurotoxicity testing (DNT). *Environ Toxicol Pharmacol 19*, 735-744. doi:10.1016/j.etap.2004.12.035

Leist, M., Hartung, T. and Nicotera, P. (2008a). The dawning of a new age of toxicology. *ALTEX 25*, 103-114. doi:10.14573/altex.2008.2.103

Leist, M., Bremer, S., Brundin, P. et al. (2008b). The biological and ethical basis of the use of embryonic stem cells for in vitro tests and cell therapy. *ALTEX 25*, 163-190. doi:10.14573/altex.2008.3.163

Leist, M., Efremova, L. and Karreman, C. (2010). Food for thought on considerations and guidelines for basic test method descriptions in toxicology. *ALTEX 27*, 309-317. doi:10.14573/altex.2010.4.309

Leist, M., Hasiwa, N., Daneshian, M. et al. (2012). Validation and quality control of replacement alternatives – Current status and future challenges. *Toxicol Res 1*, 8-22 doi:10.1039/c2tx20011b

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

Leist, M., Ghallab, A., Graepel, R. et al. (2017). Adverse outcome pathways: Opportunities, limitations and open questions. *Arch Toxicol 91*, 3477-3505 doi:10.1007/s00204-017-2045-3

Lesiak, A., Zhu, M., Chen, H. et al. (2014). The environmental neurotoxicant PCB 95 promotes synaptogenesis via ryanodine receptor-dependent miR132 upregulation. *J Neurosci 34*, 717-725. doi:10.1523/JNEUROSCI.2884-13.2014

Leung, M. C. K., Hutson, M. S., Seifert, A. W. et al. (2016). Computational modeling and simulation of genital tubercle development. *Reprod Toxicol 64*, 151-161. doi:10.1016/j.reprotox.2016.05.005

Lewitus, E., Kelave, I. and Huttner, W. B. (2013). Conical expansion of the outer subventricular zone and the role of ceocortical folding in evolution and development. *Front Hum Neurosci 7*, 424. doi:10.3389/fnhum.2013.00424

Li, B., He, X., Sun, Y. and Li, B. (2016). Developmental exposure to paraquat and maneb can impair cognition, learning and memory in Sprague-Dawley rats. *Mol Biosyst 12*, 3088-3097. doi:10.1039/C6MB00284F

Li, L., Liu, F., Salmonsen, R. A. et al. (2002). PTEN in neural precursor cells: Regulation of migration, apoptosis, and proliferation. *Mol Cell Neurosci 20*, 21-29. doi:10.1006/mcne.2002.1115

Liddelow, S. A., Guttenplan, K. A., Clarke, L. E. et al. (2017). Neurotoxic reactive astrocytes are induced by activated microglia. *Nature 541*, 481-487. doi:10.1038/nature21029

Lipton, J. O. and Sahin, M. (2014). The neurology of mTOR. *Neuron 84*, 275-291. doi:10.1016/j.neuron.2014.09.034

Lonze, B. E. and Ginty, D. D. (2002). Function and regulation of CREB family transcription factors in the nervous system. *Neuron 35*, 605-623. doi:10.1016/S0896-6273(02)00828-0

López-Espíndola, D., Morales-Bastos, C., Grijota-Martínez, C. et al. (2014). Mutations of the thyroid hormone transporter MCT8 cause prenatal brain damage and persistent hypomyelination. *J Clin Endocrinol Metab 99*, E2799-2804. doi:10.1210/jc.2014-2162

Lotharius, J., Barg, S., Wiekop, P. et al. (2002). Effect of mutant alpha-synuclein on dopamine homeostasis in a new human mesencephalic cell line. *J Biol Chem 277*, 38884-38894. doi:10.1074/jbc.M205518200

Lu, Y., Christian, K. and Lu, B. (2008). BDNF: A key regulator for protein synthesis dependent LTP and long-term memory? *Neurobiol Learn Mem 89*, 312-323. doi:10.1016/j.nlm.2007.08.018

Lund, S., Porzgen, P., Mortensen, A. L. et al. (2006). The dynamics of the LPS triggered inflammatory response of murine microglia under different culture and in vivo conditions. *J Neuroimmunol 180*, 71-87. doi:10.1016/j.jneuroim.2006.07.007

Madill, M., Fitzgerald, D., O'Connell, K. E. et al. (2016). In vitro and ex vivo models of multiple sclerosis. *Drug Discov Today 21*, 1504-1511. doi:10.1016/j.drudis.2016.05.018

Manning, B. D. and Toker, A. (2017). AKT/PKB signaling: Navigating the network. *Cell 169*, 381-405. doi:10.1016/j.cell.2017.04.001

Markus, T., Hansson, S. R., Cronberg, T. et al. (2010). beta-adrenoceptor activation depresses brain inflammation and is neuroprotective in lipopolysaccharide-induced sensitization to oxygen-glucose deprivation in organotypic hippocampal slices. *J Neuroinflammation 7*, 94. doi:10.1186/1742-2094-7-94

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

Marx-Stoelting, P., Adriaens, E., Ahr, H. J. et al. (2009). A review of the implementation of the embryonic stem cell test (EST). The report and recommendations of an ECVAM/ReProTect workshop. *Altern Lab Anim 37*, 313-328.

Mattson, M. P. (2008). Glutamate and neurotrophic factors in neuronal plasticity and disease. *Ann N Y Acad Sci 1144*, 97-112. doi:10.1196/annals.1418.005

McCarthy, D. M., Zhang, X., Darnell, S. B. et al. (2011). Cocaine alters BDNF expression and neuronal migration in the embryonic mouse forebrain. *J Neurosci 31*, 13400-13411. doi:10.1523/JNEUROSCI.2944-11.2011

McCarthy, M. M. (2008). Estradiol and the developing brain. *Physiol Rev 88*, 91-124. doi:10.1152/physrev.00010.2007

Meek, M. E. and Lipscomb, J. C. (2015). Gaining acceptance for the use of in vitro toxicity assays and QIVIVE in regulatory risk assessment. *Toxicology 332*, 112-123. doi:10.1016/j.tox.2015.01.010

Menegola, E., Broccia, M. L., Di Renzo, F. et al. (2005). Craniofacial and axial skeletal defects induced by the fungicide triadimefon in the mouse. *Birth Defects Res B Dev Reprod Toxicol 74*, 185-195. doi:10.1002/bdrb.20035

Miller, J. A., Ding, S. L., Sunkin, S. M. et al. (2014). Transcriptional landscape of the prenatal human brain. *Nature 508*, 199-206. doi:10.1038/nature13185

Molofsky, A. V. and Deneend, B. (2015). Astrocyte development: A guide for the perplexed. *Glia 63*, 1320-1329. doi:10.1002/glia.22836

Monnet-Tschudi, F., Zurich, M. G. and Honegger, P. (1993). Evaluation of the toxicity of different metal compounds in the developing brain using aggregating cell cultures as a model. *Toxicol In Vitro 7*, 335-339. doi:10.1016/0887-2333(93)90024-Y

Monnet-Tschudi, F., Zurich, M. G. and Honegger, P. (1996). Comparison of the developmental effects of two mercury compounds on glial cells and neurons in aggregate cultures of rat telencephalon. *Brain Res 741*, 52-59. doi:10.1016/S0006-8993(96)00895-5

Monnet-Tschudi, F., Zurich, M. G., Sorg, O. et al. (1999). The naturally occurring food mycotoxin fumonisin B1 impairs myelin formation in aggregating brain cell culture. *Neurotoxicology 20*, 41-48.

Monnet-Tschudi, F., Zurich, M. G., Schilter, B. et al. (2000). Maturation-dependent effects of chlorpyrifos and parathion and their oxygen analogs on acetylcholinesterase and neuronal and glial markers in aggregating brain cell cultures. Toxicol *Appl Pharmacol 165*, 175-183. doi:10.1006/taap.2000.8934

Monteggia, L. M., Barrot, M., Powell, C. M. et al. (2004). Essential role of brain-derived neurotrophic factor in adult hippocampal function. *Proc Natl Acad Sci U S A 101*, 10827-10832. doi:10.1073/pnas.0402141101

Moors, M., Cline, J. E., Abel, J. and Fritsche, E. (2007). ERK-dependent and -independent pathways trigger human neural progenitor cell migration. *Toxicol Appl Pharmacol 221*, 57-67. doi:10.1016/j.taap.2007.02.018

Moors, M., Rockel, T. D., Abel, J. et al. (2009). Human neurospheres as three-dimensional cellular systems for developmental neurotoxicity testing. *Environ Health Perspect 117*, 1131-1138. doi:10.1289/ehp.0800207

Moors, M., Vudattu, N. K., Abel, J. et al. (2010). Interleukin-7 (IL-7) and IL-7 splice variants affect differentiation of human neural progenitor cells. *Genes Immun 11*, 11-20. doi:10.1038/gene.2009.77

Moors, M., Bose, R., Johansson-Hague, K. et al. (2012). Dickkopf 1 mediates glucocorticoid-induced changes in human neural progenitor cell proliferation and differentiation. *Toxicol Sci 125*, 488-495. doi:10.1093/toxsci/kfr304

Morte, B., Díez, D., Ausó, E. et al. (2010). Thyroid hormone regulation of gene expression in the developing rat fetal cerebral cortex: Prominent role of the Ca^{2+}/calmodulin-dependent protein kinase IV pathway. *Endocrinology 151*, 810-820. doi:10.1210/en.2009-0958

Mueller, T. and Wullimann, M. (2016). *Atlas of Early Zebrafish Brain Development. A Tool for Molecular Neurogenetics*. 2nd edition. Academic Press. doi:10.1016/B978-0-12-418669-9.09987-6

Mundy, W. R., Radio, N. M. and Freudenrich, T. M. (2010). Neuronal models for evaluation of proliferation in vitro using high content screening. *Toxicology 270*, 121-130. doi:10.1016/j.tox.2010.02.004

Mundy, W. R., Padilla, S., Breier, J. M. et al. (2015). Expanding the test set: Chemicals with potential to disrupt mammalian brain development. *Neurotoxicol Teratol 52*, 25-35. doi:10.1016/j.ntt.2015.10.001

Nagappan, G. and Lu, B. (2005). Activity-dependent modulation of the BDNF receptor TrkB: Mechanisms and implications. *Trends Neurosci 28*, 464-471. doi:10.1016/j.tins.2005.07.003

Nash, R., Krishnamoorthy, M., Jenkins, A. and Csete, M. (2012). Human embryonic stem cell model of ethanol-mediated early developmental toxicity. *Exp Neurol 234*, 127-135. doi:10.1016/j.expneurol.2011.12.022

Neal, A. P. and Guillarte, R. T. (2010). Molecular neurobiology of lead (Pb^{2+}): Effect on synaptic function. *Mol Neurobiol 42*, 151-160. doi:10.1007/s00210-008-0344-1

Nelson, T. J. and Alkon, D. L. (2015). Molecular regulation of synaptogenesis during associative learning and memory. *Brain Res 1621*, 239-251 doi:10.1016/j.brainres.2014.11.054

Ninomiya, E., Hattori, T., Toyoda, M. et al. (2014). Glucocorticoids promote neural progenitor cell proliferation derived from human induced pluripotent stem cells. *Springerplus 3*, 527. doi:10.1186/2193-1801-3-527

Nishimura, Y., Murakami, S., Ashikawa, Y. et al. (2015). Zebrafish as a systems toxicology model for developmental neurotoxicity testing. *Congenit Anom (Kyoto) 55*, 1-16. doi:10.1111/cga.12079

Noyes, P. D., Haggard, D. E., Gonnerman, G. D. and Tanguay, R. L. (2015). Advanced morphological – Behavioral test platform reveals neurodevelopmental defects in embryonic zebrafish exposed to comprehensive suite of halogenated and organophosphate flame retardants. *Toxicol Sci 145*, 177-195. doi:10.1093/toxsci/kfv044

Nyffeler, J., Karreman, C., Leisner, H. et al. (2017a). Design of a high-throughput human neural crest cell migration assay to indicate potential developmental toxicants. *ALTEX 34*, 75-94. doi:10.14573/altex.1605031

Nyffeler, J., Dolde, X., Krebs, A. et al. (2017b). Combination of multiple neural crest migration assay to identify environmental toxicants from a proof-of-concept chemical library. *Arch Toxicol 91*, 3613-3632. doi:10.1007/s00204-017-1977-y

Nygard, M., Wahlstrom, G. M., Gustafsson, M. V. et al. (2003). Hormone-dependent repression of the E2F-1 gene by thyroid hormone receptors. *Mol Endocrinol 17*, 79-92. doi:10.1210/me.2002-0107

O'Rahilly, R. and Muller, F. (2006). *The Embryonic Human Brain: An Atlas of Developmental Stages*. 3rd edition. John Wiley & Sons, Inc. doi:10.1002/0471973084

OECD (2013). Revised Guidance Document for Developing and Assessing Adverse Outcome Pathways (AOPs). 2nd edition. ENV/JM/MONO(2013)6. *Series on Testing and Assessment No. 184*, OECD Publishing, Paris. https://bit.ly/1RJXOEx

OECD (2014a). New Scoping Document on In Vitro and Ex Vivo Assays for the Identification of Modulators of Thyroid Hormone Signalling. ENV/JM/MONO(2014)23. *Series on Testing and Assessment No. 207*, OECD Publishing, Paris. doi:10.1787/20777876

OECD (2014b). Guidance Document for Describing Non-Guideline In Vitro Test Methods. ENV/JM/MONO(2014)35. *Series on Testing and Assessment No. 211*, OECD Publishing, Paris. doi:10.1787/20777876

OECD (2016a). Guidance Document for the Use of Adverse Outcome Pathways in Developing Integrated Approaches to Testing and Assessment (IATA). ENV/JM/MONO(2016)67. *Series on Testing and Assessment No. 260*, OECD Publishing, Paris. https://bit.ly/2IJNOC3

OECD (2016b). Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to be used within Integrated Approaches to Testing and Assessment for Skin Sensitisation. ENV/JM/MONO(2016)29. *Series on Testing and Assessment No. 256*, OECD Publishing, Paris. doi:10.1787/20777876

OECD (2016c). Guidance Document on the Reporting of Defined Approaches to be Used Within Integrated Approaches to Testing and Assessment. ENV/JM/MONO(2016)28. *Series on Testing and Assessment No. 255*, OECD Publishing, Paris. doi:10.1787/20777876

OECD (2017). Draft Guidance Document on Good In Vitro Method Practices (GIVIMP) for the Development and Implementation of In Vitro Methods for Regulatory Use in Human Safety Assessment. OECD Paris. https://bit.ly/2LaoZgU

Orsolits, B., Borsy, A., Madarasz, E. et al. (2013). Retinoid machinery in distinct neural stem cell populations with different retinoid responsiveness. *Stem Cells Dev 22*, 2777-2793. doi:10.1089/scd.2012.0422

Padilla, S., Hunter, D. L., Padnos, B. et al. (2011). Assessing locomotor activity in larval zebrafish: Influence of extrinsic and intrinsic variables. *Neurotoxicol Teratol 33*, 624-630. doi:10.1016/j.ntt.2011.08.005

Pal, R., Mamidi, M. K., Das, A. K. and Bhonde, R. (2011). Human embryonic stem cell proliferation and differentiation as parameters to evaluate developmental toxicity. *J Cell Physiol 226*, 1583-1595. doi:10.1002/jcp.22484

Pallocca, G., Fabbri, M., Sacco, M. G. et al. (2013). miRNA expression profiling in a human stem cell-based model as a tool for developmental neurotoxicity testing. *Cell Biol Toxicol 29*, 239-257, doi:10.1007/s10565-013-9250-5

Pallocca, G., Grinberg, M., Henry, M. et al. (2016). Identification of transcriptome signatures and biomarkers specific for potential developmental toxicants inhibiting human neural crest cell migration. *Arch Toxicol 90*, 159-180. doi:10.1007/s00204-015-1658-7

Palmer, J. A., Poenitzsch, A. M., Smith, S. M. et al. (2012). Metabolic biomarkers of prenatal alcohol exposure in human embryonic stem cell-derived neural lineages. *Alcohol Clin Exp Res 36*, 1314-1324. doi:10.1111/j.1530-0277.2011.01732.x

Pamies, D., Barreras, P., Block, K. et al. (2017a). A human brain microphysiological system derived from induced pluripotent stem cells to study neurological diseases and toxicity. *ALTEX 34*, 362-376. doi:10.14573/altex.1609122

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

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

Panula, P., Chen, Y. C., Priyadarshini, M. et al. (2010). The comparative neuroanatomy and neurochemistry of zebrafish CNS systems of relevance to human neuropsychiatric diseases. *Neurobiol Dis 40*, 46-57. doi:10.1016/j.nbd.2010.05.010

Park, H. and Poo, M. M. (2013). Neurotrophin regulation of neural circuit development and function. *Nat Rev Neurosci 14*, 7-23. doi:10.1038/nrn3379

Parker, M. O., Brock, A. J., Walton, R. T. et al. (2013). The role of zebrafish (Danio rerio) in dissecting the genetics and neural circuits of executive function. *Front Neural Circuits 7*, 63. doi:10.3389/fncir.2013.00063

Parker, M. O., Annan, L. V., Kanellopoulos, A. H. et al. (2014). The utility of zebrafish to study the mechanisms by which ethanol affects social behavior and anxiety during early brain development. *Prog Neuropsychopharmacol Biol Psychiatry 55*, 94-100, doi:10.1016/j.pnpbp.2014.03.011

Patlewicz, G., Kuseva, C., Kesova, A. et al. (2014). Towards AOP application – Implementation of an integrated approach to testing and assessment (IATA) into a pipeline tool for skin sensitization. *Regul Toxicol Pharmacol 69*, 529-545. doi:10.1016/j.yrtph.2014.06.001

Pearce, R. G., Setzer, R. W., Davis, J. L. and Wambaugh, J. F. (2017). Evaluation and calibration of high-throughput predictions of chemical distribution to tissues. *J Pharmacokinet Pharmacodyn 44*, 549-565. doi:10.1007/s10928-017-9548-7

Perez-Gomez, A. and Tasker, R. A. (2012). Enhanced neurogenesis in organotypic cultures of rat hippocampus after transient subfield-selective excitotoxic insult induced by domoic acid. *Neuroscience 208*, 97-108. doi:10.1016/j.neuroscience.2012.02.003

Piersma, A. H., Bosgra, S., van Duursen, M. B. et al. (2013). Evaluation of an alternative in vitro test battery for detecting reproductive toxicants. *Reprod Toxicol 38*, 53-64. doi:10.1016/j.reprotox.2013.03.002

Piersma, A. H., Ezendam, J., Luijten, M. et al. (2014). A critical appraisal of the process of regulatory implementation of novel in vivo and in vitro methods for chemical hazard and risk assessment. *Crit Rev Toxicol 44*, 876-894. doi:10.3109/10408444.2014.940445

Pistollato, F., Louisse, J., Scelfo, B. et al. (2014). Development of a pluripotent stem cell derived neuronal model to identify chemically induced pathway perturbations in relation to neurotoxicity: Effects of CREB pathway inhibition. *Toxicol Appl Pharmacol 280*, 378-388. doi:10.1016/j.taap.2014.08.007

Pistollato, F., Canovas-Jorda, D., Zagoura, D. and Price, A. (2017). Protocol for the differentiation of human induced pluripotent stem cells into mixed cultures of neurons and glia for neurotoxicity testing. *J Vis Exp 124*, e55702. doi:10.3791/55702

Planchart, A., Mattingly, C., Allen, D. et al. (2016). Advancing toxicology research using in vivo high throughput toxicology with small fish models. *ALTEX 33*, 435-452. doi:10.14573/altex.1601281

Popova, D. and Jacobsson, S. O. (2014). A fluorescence microplate screen assay for the detection of neurite outgrowth and neurotoxicity using an antibody against betaIII-tubulin. *Toxicol In Vitro 28*, 411-418. doi:10.1016/j.tiv.2013.12.009

Puehringer, D., Orel, N., Lüningschrör, P. et al. (2013). EGF transactivation of Trk receptors regulates the migration of newborn cortical neurons. *Nat Neurosci 16*, 407-415. doi:10.1038/nn.3333

Quasthoff, S. and Hartung, H. P. (2002). Chemotherapy-induced peripheral neuropathy. *J Neurol 249*, 9-17. doi:10.1007/PL00007853

Qian, X., Nguyen, H. N., Song, M. M. et al. (2016). Brain-region-specific organoids using mini-bioreactors for modeling ZIKV exposure. *Cell 165*, 1238-1254. doi:10.1016/j.cell.2016.04.032

Radio, N. M., Breier, J. M., Reif, D. M. et al. (2015). Use of neural models of proliferation and neurite outgrowth to screen environmental chemicals in the ToxCast phase I library. *Appl In Vitro Toxicol 1*, 131-139. doi:10.1089/aivt.2014.0009

Raffaele, K. C., Rowland, J., May, B. et al. (2010). The use of developmental neurotoxicity data in pesticide risk assessments. *Neurotoxicol Teratol 32*, 563-572. doi:10.1016/j.ntt.2010.04.053

Ryan, K. R., Sirenko, O., Parham, F. et al. (2016). Neurite outgrowth in human induced pluripotent stem cell-derived neurons as a high-throughput screen for developmental neurotoxicity or neurotoxicity. *Neurotoxicology 53*, 271-281. doi:10.1016/j.neuro.2016.02.003

Reif, D. M., Martin, M. T., Tan, S. W. et al. (2010). Endocrine profiling and prioritization of environmental chemicals using ToxCast data. *Environ Health Perspect 118*, 1714-1720. doi:10.1289/ehp.1002180

Reif, D. M., Sypa, M., Lock, E. F. et al. (2013). ToxPi GUI: An interactive visualization tool for transparent integration of data from diverse sources of evidence. *Bioinformatics 29*, 402-403. doi:10.1093/bioinformatics/bts686

Reif, D., Truong, L., Mandrell, D. et al. (2016). High-throughput characterization of chemical-associated embryonic behavioral changes predicts teratogenic outcomes. *Arch Toxicol 90*, 1459-1470. doi:10.1007/s00204-015-1554-1

Rempel, E., Hölting, L., Waldmann, T. et al. (2015). A transcriptome-based classifier to identify developmental toxicants by stem cell testing: Design, validation and optimization for histone deacetylase inhibitors. *Arch Toxicol 89*, 1599-1618. doi:10.1007/s00204-015-1573-y

Richard, A. M., Judson, R. S., Houck, K. A. et al. (2016). ToxCast chemical landscape: Paving the road to 21st century toxicology. *Chem Res Toxicol 29*, 1225-1251. doi:10.1021/acs.chemrestox.6b00135

Rigamonti, A., Repetti, G., Sun, C. et al. (2016). Large-scale production of mature neurons from human pluripotent stem cells in a three-dimensional suspension culture system. *Stem Cell Reports 6*, 93-1008. doi:10.1016/j.stemcr.2016.05.010

Robinette, B. L., Harrill, J. A., Mundy, W. R. and Shafer, T. J. (2011). In vitro assessment of developmental neurotoxicity: Use of microelectrode arrays to measure functional changes in neuronal network ontogeny. *Front Neuroeng 4*, 1. doi:10.3389/fneng.2011.00001

Rovida, C., Asakura, S., Daneshian, M. et al. (2015). Toxicity testing in the 21st century beyond environmental chemicals. *ALTEX 32*, 171-181. doi:10.14573/altex.1506201

Rowitch, D. H. (2004). Glial specification in the vertebrate neural tube. *Nat Rev Neurosci 5*, 409-419. doi:10.1038/nrn1389

Sachana, M., Munn, S. and Bal-Price, A. (2016). Adverse outcome pathway on chronic binding of antagonist to N-methyl-D-aspartate receptors (NMDARs) during brain development induces impairment of learning and memory abilities. *OECD Series on Adverse Outcome Pathways No. 5*, OECD Publishing, Paris. doi:10.1787/5jlsqs5hcrmq-en

Salamon, N., Andres, M., Chute, D. J. et al. (2006). Contralateral hemimicrencephaly and clinical-pathological correlations in children with hemimegalencephaly. *Brain 129*, 352-365. doi:10.1093/brain/awh681

Sànchez-Màrtin, F. J., Fan, X., Lindquist, D. M. et al. (2013). Lead induces similar gene expression changes in brains of gestationally exposed adult mice and in neurons differentiated from mouse embryonic stem cells. *PLoS One 8*, e80558. doi:10.1371/journal.pone.0080558

Sandström von Tobel, J., Zoia, D., Althaus, J. et al. (2014). Immediate and delayed effects of subchronic Paraquat exposure during an early differentiation stage in 3D-rat brain cell cultures. *Toxicol Lett 230*, 188-197. doi:10.1016/j.toxlet.2014.02.001

Sandström, J., Broyer, A., Zoia, D. et al. (2017a). Potential mechanisms of development-dependent adverse effects of the herbicide paraquat in 3D rat brain cell cultures. *Neurotoxicology 60*, 116-124. doi:10.1016/j.neuro.2017.04.010

Sandström, J., Eggermann, E., Charvet, I. et al. (2017b). Development and characterization of a human embryonic stem cell-derived 3D neural tissue model for neurotoxicity testing. *Toxicol In Vitro 38*, 124-135. doi:10.1016/j.tiv.2016.10.001

Santos-Fandila, A., Vasquez, E., Barranco, A. et al. (2015). Analysis of 17 neurotransmitters, metabolites and precursors in zebrafish through the life cycle using ultrahigh performance liquid chromatography-tandem mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci 1001*, 191-201. doi:10.1016/j.jchromb.2015.07.040

Schildknecht, S., Pape, R., Meiser, J. et al. (2015). Preferential extracellular generation of the active parkinsonian toxin MPP^+ by transporter-independent export of the intermediate $MPDP^+$. *Antioxid Redox Signal 23*, 1001-1016. doi:10.1089/ars.2015.6297

Schildknecht, S., Di Monte, D. A., Pape, R. et al. (2017). Tipping points and endogenous determinants of nigrostriatal degeneration by MPTP. *Trends Pharmacol Sci 38*, 541-555. doi:10.1016/j.tips.2017.03.010

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

Schmidt, R., Strähle, U. and Scholpp, S. (2013). Neurogenesis in zebrafish – From embryo to adult. *Neural Dev 8*, 3. doi:10.1186/1749-8104-8-3

Schmuck, M. R., Temme, T., Dach, K. et al. (2017). Omnisphero: A high-content image analysis (HCA) approach for phenotypic developmental neurotoxicity (DNT) screenings of organoid neurosphere cultures in vitro. *Arch Toxicol 91*, 2017-2028. doi:10.1007/s00204-016-1852-2

Scholz, D., Pöltl, D., Genewsky, A. et al. (2011). Rapid, complete and large-scale generation of post-mitotic neurons from the human LUHMES cell line. *J Neurochem 119*, 957-971. doi:10.1111/j.1471-4159.2011.07255.x

Schreiber, T., Gassmann, K., Götz, C. et al. (2010). Polybrominated diphenyl ethers induce developmental neurotoxicity in a human in vitro model: Evidence for endocrine disruption. *Environ Health Perspect 118*, 572-578. doi:10.1289/ehp.0901435

Schuh, R. A., Lein, P. J., Beckles, R. A. and Jett, D. A. (2002). Noncholinesterase mechanisms ofchlorpyrifos neurotoxicity: Altered phosphorylation of Ca^{2+}/cAMP response element binding protein in cultured neurons. *Toxicol Appl Pharmacol 182*, 176-185. doi:10.1006/taap.2002.9445

Schultz, L., Zurich, M. G., Culot, M. et al. (2015). Evaluation of drug-induced neurotoxicity based on metabolomics, proteomics and electrical activity measurements in complementary CNS in vitro models. *Toxicol In Vitro 30*, 138-165. doi:10.1016/j.tiv.2015.05.016

Schwartz, M. P., Hou, Z., Propson, N. E. et al. (2015). Human pluripotent stem cell-derived neural constructs for predicting neural toxicity. *Proc Natl Acad Sci U S A 112*, 12516-12521. doi:10.1073/pnas.1516645112

Selderslaghs, I. W., Hooyberghs, J., De Coen, W. and Witters, H. E. (2010). Locomotor activity in zebrafish embryos: A new method to assess developmental neurotoxicity. *Neurotoxicol Teratol 32*, 460-471. doi:10.1016/j.ntt.2010.03.002

Selderslaghs, I. W., Hooyberghs, J., Blust, R. and Witters, H. E. (2013). Assessment of the developmental neurotoxicity of compounds by measuring locomotor activity in zebrafish embryos and larvae. *Neurotoxicol Teratol 37*, 44-56. doi:10.1016/j.ntt.2013.01.003

Senut, M. C., Sen, A., Cingolani, P. et al. (2014). Lead exposure disrupts global DNA methylation in human embryonic stem cells and alters their neuronal differentiation. *Toxicol Sci 139*, 142-161. doi:10.1093/toxsci/kfu028

Shang, E. H. and Zhdanova, I. V. (2007). The circadian system is a target and modulator of prenatal cocaine effects. *PLoS One 2*, e587, doi:10.1371/journal.pone.0000587

Shi, P., Scott, M. A., Ghosh, B. et al. (2011). Synapse microarray identification of small molecules that enhance synaptogenesis. *Nat Commun 2*, 510-519. doi:10.1038/ncomms1518

Shinde, V., Klima, S., Sureshkumar, P. S. et al. (2015). Human pluripotent stem cell based developmental toxicity assays for chemical safety screening and systems biology data generation. *J Vis Exp 100*, e52333. doi:10.3791/52333

Shinde, V., Hoelting, L., Perumal, S. S. et al. (2016). Definition of transcriptome-based indices for quantitative characterization of chemically disturbed stem cell development – Introduction of the STOP-Toxukn and STOP-Toxukk. *Arch Toxicol 91*, 839-864. doi:10.1007/s00204-016-1741-8

Shinozaki, Y., Shibata, K., Yoshida, K. et al. (2017). Transformation of astrocytes to a neuroprotective phenotype by microglia via P2Y1 receptor downregulation. *Cell Rep 19*, 1151-1164. doi:10.1016/j.celrep.2017.04.047

Silbereis, J. C., Pochareddy, S., Zhu, Y. et al. (2016). The cellular and molecular landscapes of the developing human central nervous system. *Neuron 89*, 248-268. doi:10.1016/j.neuron.2015.12.008

Simão, D., Terrasso, A. P., Teixeira, A. P. et al. (2016). Functional metabolic interactions of human neuron-astrocyte 3D in vitro networks. *Sci Rep 6*, 33285. doi:10.1038/srep33285

Slikker, W. Jr., Liu, F., Rainosek, S. W. et al. (2015). Ketamine-induced toxicity in neurons differentiated from neural stem cells. *Mol Neurobiol 52*, 959-969. doi:10.1007/s12035-015-9248-5

Slotkin, T. A. and Seidler, F. J. (2012). Developmental neurotoxicity of organophosphates targets cell cycle and apoptosis, revealed by transcriptional profiles in vivo and in vitro. *Neurotoxicol Teratol 34*, 232-241. doi:10.1016/j.ntt.2011.12.001

Smart, I. H., Dehay, C., Giroud, P. et al. (2002). Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. *Cereb Cortex 12*, 37-53. doi:10.1093/cercor/12.1.37

Smirnova, L., Hogberg, H. T., Leist, M. and Hartung, T. (2014). Developmental neurotoxicity – Challenges in the 21st century and in vitro opportunities. *ALTEX 31*, 129-156. doi:10.14573/altex.1403271

Smirnova, L., Harris, G., Delp, J. et al. (2016). A LUHMES 3D dopaminergic neuronal model for neurotoxicity testing allowing long-term exposure and cellular resilience analysis. *Arch Toxicol 90*, 2725-2743. doi:10.1007/s00204-015-1637-z

Soulé, J., Messaoudi, E. and Bramham, C. R. (2006). Brain-derived neurotrophic factor and control of synaptic consolidation in the adult brain. *Biochem Soc Trans 34*, 600-604. doi:10.1042/BST0340600

Souza, B. S., Sampaio, G. L., Pereira, C. S. et al. (2016). Zika virus infection induces mitosis abnormalities and apoptotic cell death of human neural progenitor cells. *Sci Rep 6*, 39775. doi:10.1038/srep39775

Stiegler, N. V., Krug, A. K., Matt, F. and Leist, M. (2011). Assessment of chemical-induced impairment of human neurite outgrowth by multiparametric live cell imaging in high-density cultures. *Toxicol Sci 121*, 73-87. doi:10.1093/toxsci/kfr034

Strickland, J. D., Martin, M., Houck, T. et al. (2017). Screening the ToxCast phase II libraries for neuroactivity using cortical neurons grown on multi-well microelectrode array (mwMEA) plates. *Arch Toxicol 92*, 487-500. doi:10.1007/s00204-017-2035-5

Stummann, T. C., Hareng, L. and Bremer, S. (2009). Hazard assessment of methylmercury toxicity to neuronal induction in embryogenesis using human embryonic stem cells. *Toxicology 257*, 117-126. doi:10.1016/j.tox.2008.12.018

Suarez-Isla, B. A., Pelto, D. J., Thompson, J. M. and Rapoport, S. I. (1984). Blockers of calcium permeability inhibit neurite extension and formation of neuromuscular synapses in cell culture. *Brain Res 316*, 263-270. doi:10.1016/0165-3806(84)90311-0

Talens-Visconti, R., Sanchez-Vera, I., Kostic, J. et al. (2011). Neural differentiation from human embryonic stem cells as a tool to study early brain development and the neuroteratogenic effects of ethanol. *Stem Cells Dev 20*, 327-339. doi:10.1089/scd.2010.0037

Tamm, C. and Ceccatelli, S. (2017). Mechanistic insight into neurotoxicity induced by developmental insults. *Biochem Biophys Res Commun 482*, 408-418. doi:10.1016/j.bbrc.2016.10.087

Tang, H., Hammack, C., Ogden, S. C. et al. (2016). Zika virus infects human cortical neural progenitors and attenuates their growth. *Cell Stem Cell 18*, 587-590. doi:10.1016/j.stem.2016.02.016

Tasneem, S., Farrell, K., Lee, M. Y. and Kothapalli, C. R. (2016). Sensitivity of neural stem cell survival, differentiation and neurite outgrowth within 3D hydrogels to environmental heavy metals. *Toxicol Lett 242*, 9-22. doi:10.1016/j.toxlet.2015.11.021

Terron, A., Bal-Price, A., Paini, A. et al. (2018). An adverse outcome pathway for parkinsonian motor deficits associated with mitochondrial complex I inhibition. *Arch Toxicol 92*, 41-82. doi:10.1007/s00204-017-2133-4

Tiedeken, J. A., Ramsdell, J. S. and Ramsdell, A. F. (2005). Developmental toxicity of domoic acid in zebrafish (Danio rerio). *Neurotoxicol Teratol 27*, 711-717. doi:10.1016/j.ntt.2005.06.013

Tofighi, R., Wan Ibrahim, W. N., Rebellato, P. et al. (2011). Non-dioxin-like polychlorinated biphenyls interfere with neuronal differentiation of embryonic neural stem cells. *Toxicol Sci 124*, 192-201. doi:10.1093/toxsci/kfr221

Tokuda, S., Mahaffey C. L, Monks, B. et al. (2011). A novel Akt3 mutation associated with enhanced kinase activity and seizure susceptibility in mice. *Hum Mol Genet 20*, 988-999. doi:10.1093/hmg/ddq544

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

Tonduti, D., Vanderver, A., Berardinelli, A. et al. (2013). MCT8 deficiency: Extrapyramidal symptoms and delayed myelination as prominent features. *J Child Neurol 28*, 795-800. doi:10.1177/0883073812450944

Tonk, E. C., Pennings, J. L. and Piersma, A. H. (2015). An adverse outcome pathway framework for neural tube and axial defects mediated by modulation of retinoic acid homeostasis. *Reprod Toxicol 55*, 104-113. doi:10.1016/j.reprotox.2014.10.008

Toscano, C. D., Hashemzadeh-Gargari, H., McGlothan, J. L. and Guilarte, T. R. (2002). Developmental Pb^{2+} exposure alters NMDAR subtypes and reduces CREB phosphorylation in the rat brain. *Dev Brain Res 139*, 217-226. doi:10.1016/S0165-3806(02)00569-2

Tschopp, O., Yang, Z. Z. and Brodbeck, D. (2005). Essential role of protein kinase B gamma (PKB gamma/Akt3) in postnatal brain development but not in glucose homeostasis. *Development 132*, 2943-2954. doi:10.1242/dev.01864

Tsuji, R. and Crofton K. M. (2012). Developmental neurotoxicity guideline study: Issues with methodology, evaluation and regulation. *Congenit Anom (Kyoto) 52*, 122-128. doi:10.1111/j.1741-4520.2012.00374.x

Tukker, A. M., de Groot, M. W., Wijnolts, F. M. et al. (2016). Is the time right for in vitro neurotoxicity testing using human iPSC-derived neurons? *ALTEX 33*, 261-271. doi:10.14573/altex.1510091

Tyler, W. J., Alonso, M., Bramham, C. R. and Pozzo-Miller, L. D. (2002). From acquisition to consolidation: On the role of brain-derived neurotrophic factor signaling in hippocampal-dependent learning. *Learn Mem 9*, 224-237. doi:10.1101/lm.51202

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

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. doi:10.1016/j.neuro.2011.10.002

Vassallo, A., Chiappalone, M., De Camargos, L. R. et al. (2017). A multi-laboratory evaluation of microelectrode array-based measurements of neural network activity for acute neurotoxicity testing. *Neurotoxicology 60*, 280-292. doi:10.1016/j.neuro.2016.03.019

Viberg, H. (2009). Neonatal ontogeny and neurotoxic effect of decabrominated diphenyl ether (PBDE 209) on levels of synaptophysin and tau. *Int J Dev Neurosci 27*, 423-429. doi:10.1016/j.ijdevneu.2009.05.007

Vinken, M. (2013). The adverse outcome pathway concept: A pragmatic tool in toxicology. *Toxicology 312*, 158-165. doi:10.1016/j.tox.2013.08.011

Visan, A., Hayess, K., Sittner, D. et al. (2012). Neural differentiation of mouse embryonic stem cells as a tool to assess developmental neurotoxicity in vitro. *Neurotoxicology 33*, 1135-1146. doi:/10.1016/j.neuro.2012.06.006

Volbracht, C., Leist, M. and Nicotera, P. (1999). ATP controls neuronal apoptosis triggered by microtubule breakdown or potassium deprivation. *Mol Med 5*, 477-489.

Vorhees, C. V. and Makris, S. L. (2015). Assessment of learning, memory, and attention in developmental neurotoxicity regulatory studies: Synthesis, commentary, and recommendations. *Neurotoxicol Teratol 52*, 109-115. doi:10.1016/j.ntt.2015.10.004

Wagenaar, D. A., Pine, J. and Potter, S. M. (2006). An extremely rich repertoire of bursting patterns during the development of cortical cultures. *BMC Neurosci 7*, 11. doi:10.1186/1471-2202-7-11

Waldmann, T., Rempel, E., Balmer, N. V. et al. (2014). Design principles of concentration-dependent transcriptome deviations in drug-exposed differentiating stem cells. *Chem Res Toxicol 27*, 408-420. doi:10.1021/tx400402j

Waldmann, T., Grinberg, M., König, A. et al. (2017). Stem cell transcriptome responses and corresponding biomarkers that indicate the transition from adaptive responses to cytotoxicity. *Chem Res Toxicol 30*, 905-922. doi:10.1021/acs.chemrestox.6b00259

Wallace, K., Strickland, J. D., Valdivia, P. et al. (2015). A multiplexed assay for determination of neurotoxicant effects on spontaneous network activity and viability from microelectrode arrays. *Neurotoxicology 49*, 79-85. doi:10.1016/j.neuro.2015.05.007

Wang, L., Liu, Y., Li, S. et al. (2015). Wnt signaling pathway participates in valproic acid-induced neuronal differentiation of neural stem cells. *Int J Clin Exp Pathol 8*, 578-585.

Wang, L., Zhou, K., Fu, Z. et al. (2017). Brain development and Akt signaling: The crossroads of signaling pathway and neurodevelopmental diseases. *J Mol Neurosci 61*, 379-384. doi:10.1007/s12031-016-0872-y

Weng, M., Zimmer, B., Pöltl, D. et al. (2012). Extensive transcriptional regulation of chromatin modifiers during human neurodevelopment. *PLoS One 7*, e36708. doi:10.1371/journal.pone.0036708

Weng, M. K., Nataraj, K., Scholz, D. et al. (2014). Lineage-specific regulation of epigenetic modifier genes in human liver and brain. *PLoS One 9*, e102035. doi:10.1371/journal.pone.0102035

Wetmore, B. A. (2015). Quantitative in vitro-to-in vivo extrapolation in a high-throughput environment. *Toxicology 332*, 94-101. doi:10.1016/j.tox.2014.05.012

Wilson, M. S., Graham, J. R. and Ball, A. J. (2014). Multiparametric high content analysis for assessment of neurotoxicity in differentiated neuronal cell lines and human embryonic stem cell-derived neurons. *Neurotoxicology 42*, 33-48. doi:10.1016/j.neuro.2014.03.013

Workman, A. D., Charvet, C. J., Clancy, B. et al. (2013). Modeling transformations of neurodevelopmental sequences across mammalian species. *J Neurosci 33*, 7368-7383. doi:10.1523/JNEUROSCI.5746-12.2013

Worth, A. P. and Patlewicz, G. (2016). Integrated approaches to testing and assessment. *Adv Exp Med Biol 856*, 317-342. doi:10.1007/978-3-319-33826-2_13

Wu, C., Orozco, C., Boyer, J. et al. (2009). BioGPS: An extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol 10*, R130. doi:10.1186/gb-2009-10-11-r130

Xiong, Y., Ibhazehiebo, K., Iwasaki, T. and Koibuchi, N. (2012). An in vitro method to study the effects of thyroid hormone-disrupting chemicals on neuronal development. *Neurotoxicology 33*, 753-757. doi:10.1016/j.neuro.2012.04.021

Yang, D., Kania-Korwel, I., Ghogha, A. et al. (2014). PCB 136 atropselectively alters morphometric and functional parameters of neuronal connectivity in cultured rat hippocampal neurons via ryanodine receptor-dependent mechanisms. *Toxicol Sci 138*, 379-392. doi:10.1093/toxsci/kft334

Yao, L., Peng, S. X., Xu, Y. D. et al. (2017). Unexpected neuroprotective effects of loganin on 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-induced neurotoxicity and cell death in zebrafish. *J Cell Biochem 118*, 615-628. doi:10.1002/jcb.25749

Yoon, M., Campbell, J. L., Andersen, M. E. and Clewell, H. J. (2012). Quantitative in vitro to in vivo extrapolation of cell-based toxicity assay results. *Crit Rev Toxicol 42*, 633-652. doi:10.3109/10408444.2012.692115

York, R. G., Barnett, J., Girard, M. F. et al. (2005). Refining the effects observed in a developmental neurobehavioral study of ammonium perchlorate administered orally in drinking water to rats. II. Behavioral and neurodevelopment effects. *Int J Toxicol 24*, 451-467. doi:10.1080/10915810500367094

Yu, J. S. and Cui, W. (2016). Proliferation, survival and metabolism: The role of PI3K/AKT/mTOR signalling in pluripotency and cell fate determination. *Development 143*, 3050-3060. doi:10.1242/dev.137075

Zagoura, D., Canovas-Jorda, D., Pistollato, F. et al. (2017). Evaluation of the rotenone-induced activation of the Nrf2 pathway in a neuronal model derived from human induced pluripotent stem cells. *Neurochem Int 106*, 62-73. doi:10.1016/j.neuint.2016.09.004

Zheng, J., Feng, X., Hou, L. et al. (2011). Latanoprost promotes neurite outgrowth in differentiated RGC-5 cells via the PI3K-Akt-mTOR signaling pathway. *Cell Mol Neurobiol 31*, 597-604. doi:10.1007/s10571-011-9653-x

Zhou, H., Liu, Y., Tan, X. J. et al. (2015). Inhibitory effect of arsenic trioxide on neuronal migration in vitro and its potential molecular mechanism. *Environ Toxicol Pharmacol 40*, 671-677. doi:10.1016/j.etap.2015.08.026

Zhou, J. and Parada L. F. (2012). PTEN signaling in autism spectrum disorders. *Curr Opin Neurobiol 22*, 873-879. doi:10.1016/j.conb.2012.05.004

Zhu, Y. P., Xi, S. H., Li, M. Y. et al. (2017). Fluoride and arsenic exposure affects spatial memory and activates the ERK/CREB signaling pathway in offspring rats. *Neurotoxicology 59*, 56-64. doi:10.1016/j.neuro.2017.01.006

Zimmer, B., Kuegler, P. B., Baudis, B. et al. (2011a). Coordinated waves of gene expression during neuronal differentiation of embryonic stem cells as basis for novel approaches to developmental neurotoxicity testing. *Cell Death Differ 18*, 383-395. doi:10.1038/cdd.2010.109

Zimmer, B., Schildknecht, S., Kuegler, P. B. et al. (2011b). Sensitivity of dopaminergic neuron differentiation from stem cells to chronic low-dose methylmercury exposure. *Toxicol Sci 12*, 357-367. doi:10.1093/toxsci/kfr054

Zimmer, B., Lee, G., Balmer, N. V. et al. (2012). Evaluation of developmental toxicants and signaling pathways in a functional test based on the migration of human neural crest cells. *Environ Health Perspect 120*, 1116-1122. doi:10.1289/ehp.1104489

Zimmer, B., Pallocca, G., Dreser, N. et al. (2014). Profiling of drugs and environmental chemicals for functional impairment of neural crest migration in a novel stem cell-based test battery. *Arch Toxicol 88*, 1109-1126. doi:10.1007/s00204-014-1231-9

Zimmermann, F. F., Gaspary, K. V., Leite, C. E. et al. (2015). Embryological exposure to valproic acid induces social interaction deficits in zebrafish (Danio rerio): A developmental behavior analysis. *Neurotoxicol Teratol 52*, 36-41. doi:10.1016/j.ntt.2015.10.002

Zuo, Z., Cai, J., Wang, X. et al. (2009). Acute administration of tributyltin and trimethyltin modulate glutamate and N-methyl-D-aspartate receptor signalling pathway in Sebastiscus marmoratus. *Aquat Toxicol 92*, 44-49. doi:10.1016/j.aquatox.2009.01.008

Zurich, M. G., Honegger, P., Schilter, B. et al. (2000). Use of aggregating brain cell cultures to study developmental effects of organophosphorus insecticides. *Neurotoxicology 21*, 599-605.

Zurich, M. G., Eskes, C., Honegger, P. et al. (2002). Maturation-dependent neurotoxicity of lead aceate in vitro: Implication of glial reactions. *J Neurosc Res 70*, 108-116. doi:10.1002/jnr.10367

Zurich, M. G., Stanzel, S., Kopp-Schneuder, A. et al. (2012). Evaluation of aggregating brain cell cultures for the detection of acute organ-specific toxicity. *Toxicol In Vitro 27*, 1416-1424. doi:10.1016/j.tiv.2012.06.018

Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgements

This work was supported by the DZ foundation, the European Commission (EU-ToxRisk), EFSA, the BMBF and JPI-NutriCog-Selenius.