"I have yet to see any problem,
however complicated, which, when
you look at it in the right way, did
not become still more complicated."

Poul Anderson
(1926-2001)

Food for Thought ...
Developmental Neurotoxicity – Challenges in the 21st Century and *In Vitro* Opportunities

Lena Smirnova [1], Helena T. Hogberg [1], Marcel Leist [2], and Thomas Hartung [1,2]

Centers for Alternatives to Animal Testing (CAAT) at [1]Johns Hopkins Bloomberg School of Public Health, USA, and [2]University of Konstanz, Germany

Summary

In recent years neurodevelopmental problems in children have increased at a rate that suggests lifestyle factors and chemical exposures as likely contributors. When environmental chemicals contribute to neurodevelopmental disorders developmental neurotoxicity (DNT) becomes an enormous concern. But how can it be tackled? Current animal test-based guidelines are prohibitively expensive, at $1.4 million per substance, while their predictivity for human health effects may be limited, and mechanistic data that would help species extrapolation are not available. A broader screening for substances of concern requires a reliable testing strategy, applicable to larger numbers of substances, and sufficiently predictive to warrant further testing. This review discusses the evidence for possible contributions of environmental chemicals to DNT, limitations of the current test paradigm, emerging concepts and technologies pertinent to in vitro DNT testing and assay evaluation, as well as the prospect of a paradigm shift based on 21st century technologies.

Keywords: environmental exposure, developmental neurotoxicity, species extrapolation, predictivity

Introduction

Developmental neurotoxicity (DNT) is probably the least tested health effect of chemicals: only about 150 substances have been subjected to the internationally agreed guideline studies. We lack DNT data for almost all chemicals, including environmental pollutants, industrial chemicals, drugs, consumer products, and food additives. Epidemiological studies in this field can hardly prove causal relationships unless effects are dramatic; only a handful of compounds, therefore, have been established as definitive DNToxicants in man (Grandjean and Landrigan, 2006): methyl mercury, lead, arsenic, PCBs, toluene, and ethanol. This group was recently expanded to include six additional developmental neurotoxicants – manganese, fluoride, chlorpyrifos, dichlorodiphenyltrichloroethane, tetrachloroethylene, and the polybrominated diphenyl ethers (Grandjean and Landrigan, 2014). This relatively small number of DNToxicants contrasts strongly with the potential risk: the fact that the developing brain in children and fetuses is much more vulnerable to chemical perturbation than the adult brain, leads to major concerns about deficient DNT data. The high sensitivity of the developing brain is due to the still immature blood/brain-barrier, increased absorption versus low body weight, and diminished ability to detoxify exogenous chemicals (Adinolfi, 1985; Tilson, 2000; NRC, 2000). Moreover, CNS development is a complex process involving many different events, such as differentiation of progenitor cells, proliferation and cell migration, synaptogenesis, myelination, cell death, synthesis of neurotransmitters, and formation of receptors. These events occur within strictly controlled timeframes and, therefore, each event creates a different window of vulnerability to xenobiotic exposure (Rice and Barone, 2000; Rodier, 1994, 1995). Once neurodevelopment is disturbed there is little potential for repair and it often leads to permanent consequences.

In addition, it is believed that environmental chemicals contribute to the observed increase in neurodevelopmental disor-

http://dx.doi.org/10.14573/altex.1403271

ders such as lowered IQ, learning disabilities, attention deficit hyperactivity disorder (ADHD) and, in particular, autism (Kuehn, 2010; Sagiv et al., 2010; Grandjean and Landrigan, 2006; Landrigan, 2010).

The main reason for the lack of data lies in the current guidelines for DNT (OECD TG 426 and US EPA 712-C-98-239) (OECD, 2007; US EPA, 1998) themselves – the guidelines are based entirely on *in vivo* experiments, which are costly, time consuming, and unsuitable for testing a larger number of chemicals. The testing of one chemical takes about three months, uses approximately 1000 rat pups, and costs about $ 1.4 million. For these reasons, there is currently no regulatory request for DNT studies prior to registration of new chemicals and recommendations for DNT testing are only based on certain triggers such as structural similarity with known reproductive toxicants, concerns for endocrine disruption, results from other toxicity studies, and the anticipated use and human exposure patterns. Data that can detect these triggers, however, are often lacking as well. Furthermore, if a DNT study is performed, the data can be difficult to interpret and rarely contribute to regulation and risk assessment.

Thus, fast and reliable identification of DNT effects of chemicals using a battery of high-throughput tests of modern toxicology is a high priority. This will facilitate and speed up the process of risk assessment and identification of possible environmental/gene interaction leading to neurodevelopmental disorders and prevent an increase of such disorders in the future. To improve and speed up DNT testing, experts in the field from industry, academia, and government have discussed the development of alternative approaches for testing for DNT over the last decade with a series of conferences and workshops, many steered by Center for Alternatives to Animal Testing (CAAT) at Johns Hopkins University and the European Commission's European Centre for the Validation of Alternative Methods (ECVAM); in fact the fourth international conference will take place in May 2014[1]. Experts in the field in this series discussed the current status and problems of DNT assessment, identified promising alternative approaches to be included in an integrated testing strategy, and provided recommendations for the future (Coecke et al., 2007; Lein et al., 2007; Crofton et al., 2011; Bal-Price et al., 2010b, 2012).

Consideration 1: "Epidemic" of neurodevelopmental disorders

Today one out of six children is diagnosed with a developmental disorder (Boyle et al., 1994; Decouflé et al., 2001; Schettler, 2001) and in many cases this involves the central nervous system (CNS). Disorders of neurobehavioral development affect 10-15% of all births (Grandjean and Landrigan, 2014). Common neurodevelopmental disorders include learning disabilities, neurodevelopmental delays, autism spectrum disorders (ASD), and attention deficit and hyperactivity disorder (ADHD). ASD affects 1 in 110 individuals in the US (Hu, 2013), increasing to 1 in 68 according to CDC in 2014[2], and 1-64 in the UK (Ratajczak, 2011); ADHD affects 14% of the 4 million children born in the US each year (Landrigan et al., 2012) and learning disabilities affect up to 10% of children attending public schools (Schmid and Rotenberg, 2005). Although the assessment and reporting of these disorders have improved over the last few years, scientific evidence suggests that the incidence of such disorders is actually increasing. ASDs are a major public health concern in the United States with associated morbidity and functional limitations substantially diminishing quality of life accounting for annual direct care and associated indirect costs estimated at $126 (Ema et al., 2007). The rates for autism have doubled over the last decade, quadrupled over the last two (Schmid and Rotenberg, 2005), and reached 1.1% in the US according to CDC[3], hinting at environmental risk factors. Increased recognition might lead to better diagnoses as, for example, suggested by the perfect correlation of children diagnosed with autism and media coverage[4], but the question is, what is the hen and what the egg? In general, autism increases are not believed to be a result of reclassification (Sullivan, 2005). A major concern remains the possibility that exposures to drugs and industrial chemicals have contributed to this increase (Kuehn, 2010; Sagiv et al., 2010; Grandjean and Landrigan, 2006). It was estimated that about 4% of prescription drugs have been withdrawn from the market because of observed adverse neurological effects (Fung et al., 2001), adding to such concerns. Proof-of-concept evidence for a contribution of drugs derives from studies specifically linking autism to exposures in early pregnancy to thalidomide, misoprostol, and valproic acid (Landrigan, 2010).

It is widely accepted that the developing CNS is much more vulnerable to chemical induced injury than the CNS of the adult. The high sensitivity of the developing brain is due to the still immature blood/brain-barrier, increased absorption versus low body weight, and a diminished ability to detoxify exogenous chemicals, making the developing CNS much more sensitive to the chemical exposures, and adverse effects can be more severe and less reversible than those in adults.

Bondy and Campbell (2005) argued: "*It has been calculated that, were newborn infants to experience a loss of 30 IQ points resulting from a transient prenatal exposure to a toxic agent, one would be very unlikely to uncover the cause of this deficit. In the absence of spectacular and obvious physical changes, such as those incurred with prenatal exposure to thalidomide, minor behavioral impairments are very difficult to detect and attribute to a gestational origin... An important distinction is between the effect of a minor insult to an individual and that on society as a whole. Thus, if an exposure to a chemical agent were to cause a drop in IQ of 5 points compared with the IQ achievable under optimal conditions, this would probably not affect an individual greatly. However, a widespread exposure*

[1] http://caat.jhsph.edu/programs/workshops/DNT4/index.html

[2] http://www.cdc.gov/media/releases/2014/p0327-autism-spectrum-disorder.html

[3] http://www.cdc.gov/NCBDDD/autism/data.html

[4] Nate Silver "The signal and the noise", Penguin Press, 2012, page 218, Fig. 7-4.

(for example, such as that existing for lead) could spread such a deficit over the whole population. This would markedly affect the lower and upper ends of a bell-shaped intelligence distribution curve and result in a significantly greater percentage of the population who are not able to care for themselves as well as a decrease in the number of highly gifted individuals."

Grandjean and Landrigan (2014, references there) recently made impressive extrapolations on the societal costs of DNT: "*Loss of cognitive skills reduces children's academic and economic attainments and has substantial long-term economic effects on societies. Thus, each loss of one IQ point has been estimated to decrease average lifetime earnings capacity by about €12000 or US$18000 in 2008 currencies. The most recent estimates from the USA indicate that the annual costs of childhood lead poisoning are about US$50 billion and that the annual costs of methylmercury toxicity are roughly US$5 billion. In the European Union, methylmercury exposure is estimated to cause a loss of about 600000 IQ points every year, corresponding to an annual economic loss of close to €10 billion. In France alone, lead exposure is associated with IQ losses that correspond to annual costs that might exceed €20 billion. Since IQ losses represent only one aspect of developmental neurotoxicity, the total costs are surely even higher*."

Similarly, Ganz (2007) calculated the societal costs per autism case at $3.2 million, mainly because of lost productivity and the need for adult care.

Altogether, this makes DNT stand out as a prototypic emerging health effect. Its incidence is apparently increasing, as is public awareness. A precautionary reflex suggests increased testing and banning of possible health threats. However, we also can propose an alternative scenario: What is typically happening with emerging health threats is positive *feed-forward*. Research will increase, funds will be made available and, with the bias of reporting positive associations this goes viral. The experts are called on and they will be unlikely to state that their field is of lesser importance. Only with time will self-critical evaluations arrive, comparing predictions with real developments and sorting the signal from the noise. By then, costly political and economical decisions will have been made. The balance between both approaches is best met with the strategic evaluation of the more likely threats with quality-assured tools. Learning from these helps to understand the extent of the threat and furthers the quality of the tools used to assess it.

Consideration 2: Epidemiological studies of neurodevelopmental disorders

What could be better than studying humans under natural exposure conditions (i.e., observational epidemiological studies (Rice, 2005)) when nailing a human health effect such as DNT of chemical exposure? Cynical reply: It took epidemiology some 50 years to show that smoking induces lung cancer. There are limitations in epidemiological studies – often foremost, there are the limitations of costs and (repeated) access to reasonably representative cohorts. We have to distinguish here between retrospective and prospective epidemiological studies (DiPietro, 2010). Retrospective studies have the advantage that the cases (e.g., disease group) are easily defined, but bias in selecting a control group retrospectively is very problematic. In contrast, prospective studies do not have such bias, as at the time point of enrollment it is not clear who belongs to which group. However, prospective studies, especially for rare diseases, often require extremely large study groups to find a sufficient number of cases. Further challenges lie in biosampling, especially from babies and young children and mothers during pregnancy. This further limits biomonitoring of exposure as well as biological phenotyping of mother/child and their responses to exposure.

Rice (2005) noted the lack of strategies for choice of tests, data analysis, and interpretation of results for neurodevelopmental epidemiological studies. Research on environmental causes of ASD has been limited to date by the lack of prospective studies that include the perinatal window, ASD phenotypes, and simultaneous biosample and epidemiological data. The low frequency of ASD cases impairs prospective studies, but by defining high-risk groups (enriched risk pregnancy cohort) this can be improved (as has been identified as a priority in the 2011 Interagency Autism Coordinating Committee report (IACC, 2011)). A positive example is the Early Autism Risk Longitudinal Investigation (EARLI) cohort[5] – an enriched risk, prospective pregnancy cohort that follows mothers of a child with ASD at the start of a subsequent pregnancy (Newschaffer et al., 2012).

The key problem in case of DNT, however, is that we know if at all only a few DNToxicants. Furthermore, for environmental epidemiology the exposure metrics are extremely difficult (Sim, 2002). Probabilistic exposure determination appears to be a valuable option (Jager et al., 2001; Gustafson and McCandless, 2010). Noteworthy, for pesticides, where exposure can be measured more easily, a recent extensive review of "*the epidemiologic studies did not strongly implicate any particular pesticide as being causally related to adverse neurodevelopmental outcomes in infants and children*" (Burns et al., 2013), in contrast to the available animal studies. A most interesting new avenue is the concept of the Human Exposome (Rappaport, 2011; Wild, 2011), which tries to identify biomarkers of human exposure typically in blood or urine. Evidence of possible low molecular weight biomarkers for ASD is emerging (Al-Gadani et al., 2009; Pastural et al., 2009; James et al., 2004; Ratajczak, 2011; Walsh et al., 2011; Hammock et al., 2012; Austin and Shandley, 2008; Ming et al., 2005). This will help epidemiological and clinical studies, but there are still study design challenges (Fowke, 2009), or as John M. Cowden ironically put it: "*There are three kinds of epidemiologist: those who can count and those who can't*."[6]

[5] http://www.earli.org

[6] http://wwwnc.cdc.gov/eid/article/16/1/09-0030_article.htm

Consideration 3: Is there a need for DNT studies?

Most chemicals are never tested for DNT since there are no general requirements for DNT testing of chemicals or pesticides (EPA requires it only for those thought to have neurotoxic effects) prior to their registration, for example, in REACH. The US National Research Council (NRC) estimated in 2000 that 3% of developmental disabilities are direct consequences of neurotoxic exposures and another 25% are due to environmental exposures plus genetic susceptibility (Landrigan et al., 2004). The NRC (2009) identified large gaps in the testing of chemicals for developmental neurotoxicity, which results in a paucity of systematic data to guide prevention and the huge amount of proof needed for regulation. Very few chemicals, therefore, have been regulated as a result of developmental neurotoxicity. DNT evaluation is based on a weight-of-evidence approach for determining when testing should be recommended. This is done by gathering available data from all toxicity studies as well as information of potential human exposure. The decision about whether a chemical should be recognized to have a possible trigger that would require DNT studies can, for example, be based on observations of neurological effects or induced structural abnormalities of the CNS. Chemical triggers can be adult neurotoxicants, hormonally active peptides and amino acids, or chemicals that are structurally similar to other chemicals with DNT effects. However, due to commonly used experimental designs, the relevant data that is needed to trigger a DNT evaluation is not always available. Chemicals with widespread human exposure would primarily be tested for reproductive toxicity and adult neurotoxicity before making a decision for further DNT studies. However, regulators mainly require reproductive and developmental studies for food-used pesticides and now (under REACH) for high-production volume chemicals (though with limited testing proposals in response from industry (Rovida et al., 2011)). In addition, adult neurotoxicological studies are only demanded if certain triggers have been found, such as if the test substance is an organophosphate or a pesticide with structural similarities to a substance that causes delayed neurotoxicity.

Many chemicals will probably meet the criteria for DNT testing, but since we already know the limitations in the data set for several chemicals, these requirements might not be sufficient to protect children from exposure to potential DNToxicants. Currently, DNT testing has only been performed for less than 200 chemicals (most of them pesticides (Bjørling-Poulsen et al., 2008)) and only a few of these studies contributed to risk assessment (Makris et al., 2009; van Thriel et al., 2012). There are several explanations for this. First, the endpoints were not as sensitive as previous tests giving a NOAEL at a higher concentration. If this is correct for most chemicals, DNT studies should be avoided. Other reasons could be that the data was not complete or difficult to interpret, which made it difficult to use for regulatory purposes. If this occurs often the guidelines are not good enough and should be changed. In fact, that is what most experts from industry, academia, and regulatory bodies believe.

As for the D in DNT – what makes DNT different from neurotoxicity? The human brain is an organ of unmatched complexity. Its development lasts at least until puberty and these processes create windows of vulnerability beyond the neurotoxic effects in adults. There are indeed several examples of altered drug targets and biochemical events in the developing nervous system (Selenica et al., 2007). Nevertheless, most hitherto known DNT compounds are also neurotoxicants. It appears that specific vulnerabilities are rather dose-dependent, i.e., that lower concentrations damage the developing nervous system. Different outcomes from similar initial effects are also

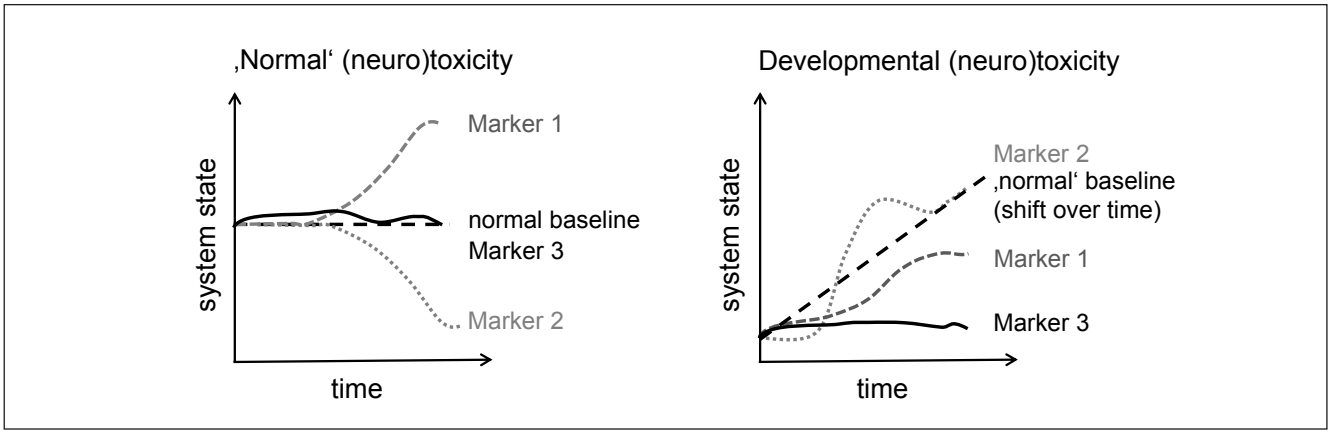

Fig. 1 Developmental neurotoxicity assessment differs conceptually from that of neurotoxicity, and therefore requires special methods for data analysis and visualization

The major difference is due to the continuously changing state of the system, i.e., the baseline for omics analysis or morphometric assays is not constant. It shifts, because the types and populations of cells present in a culture or a tissue keep changing. This has important implications. For instance, marker 1 is up-regulated by toxicant exposure in a "normal" neurotoxicological test (NT) system (e.g., adult neurons) and in a developmental neurotoxicity (DNT) test (e.g., developing stem cells). In the first case, this up-regulation is real when compared to baseline. In the second case (DNT), the up-regulation is in fact a relative down-regulation, compared to baseline. Another example is illustrated by changes of marker 3 by a toxicant. This marker does not change in absolute terms in DNT or NT. But in DNT, this marker is relatively down-regulated, compared to baseline.

linked to the degree of reversibility: the effects of some chemicals in the developing nervous system are permanent, while they are only transient in the adult. This applies in particular to neurotoxicants, which act reversibly on synaptic function. A nicotinic agonist or an acetylcholine esterase inhibitor may show transient neurotoxicity in the adult, but these effects are adaptable (receptor down-regulation) and fully reversible. The situation is different during the development of the nervous system, when appropriate signals are required for the formation of correct networks, and disruption of such signals during critical time windows can have permanent effects because of malformed networks. Another distinguishing feature is that the phases of rapid growth are especially sensitive to toxic disruption, as we well know from tumor chemotherapy.

Much of the difference is also explained by variations in exposure and pharmacokinetics: Children drink more water, eat more food, and breathe more air per body weight compared with adults (Landrigan et al., 2004): "*Children in the first 6 months of life drink seven times as much water, whereas children ages 1 through 5 years eat 3 to 4 times more food on a body-weight basis than the average adult. The air intake of a resting infant is twice that of an adult. The implication of these findings for health is that children will have substantially heavier exposures than adults to any environmental contaminants present in water, food, and air*."

Moreover, metabolic detoxification is still immature and the blood-brain barrier shows different properties

There are also major differences between DNT and NT with respect to testing endpoints. This applies both to *in vivo* and *in vitro* endpoints. The situation is illustrated in Figure 1, using transcriptional markers as test endpoints.

All together: (i) lack of DNT studies for most chemicals, (ii) increased vulnerability of developing brain, and (iii) rising case numbers in neurodevelopmental disorders do not leave any doubt of need for DNT studies.

Consideration 4: Current DNT studies and their limitations

Systematic guideline-based testing of new chemicals for toxicity before marketing has only been required since the 1980s and many widely used chemicals were never sufficiently assessed for their human and environmental safety. This has led to upcoming changes in the chemical regulation in the Western societies attempting to close the gap in knowledge of the toxic effects of chemicals. Systematic testing for DNT is still not routinely required by most regulatory agencies and becomes obligatory only if it has been triggered by observations during organ toxicity testing (Bal-Price et al., 2010a,b). The DNT test guidelines were developed to serve as a general framework to assess DNT and address a number of study design issues: They should be suitable for testing of any chemical and provide consistency, but also flexibility in the specific methodology used. This means that chemicals used in different regulatory frameworks (pesticides, insecticides, food additives, cosmetics, industrial chemicals, nanoparticles) do not always have to undergo the same testing. In 1991 the US Environmental Protection Agency (EPA) issued the first guideline for DNT (US EPA OPPTS Developmental Neurotoxicity Testing Guideline 870.6300 § 83-6) that was revised and published in 1998 (US EPA, 1998). The guideline was founded upon an extensive scientific database including between-laboratory "validation" studies (Makris et al., 2009). However, since children are considered a susceptible population they require much more extensive evaluations of potential risks, and US EPA recommended the inclusion of a DNT study for all chemical food-used pesticides (Food Quality Protection Act from 1996). The recommendation was expanded to include all organophosphate insecticides (US EPA in 1999) and in 2002 US EPA required registrants to perform DNT studies for a wide range of pesticides that showed evidence of neurotoxicity.

In 1995 the Organization for Economic Co-operation and Development (OECD) initiated the development of the OECD test guideline 426, using the US EPA guideline as a template, which was adopted by the OECD council (OECD, 2007). TG 426 addressed important issues and incorporated improvements recommended by expert consultation meetings held between 1996 and 2005. However, only a limited number of chemicals have been tested according to these guidelines. More recently, OECD has published the extended one-generation reproductive toxicity test guideline (TG 443) that also allows assessing the potential impact of chemical exposure on the developing nervous system (OECD, 2011). This new guideline is intended to evaluate all developing life stages, both prenatal and post-natal up to puberty, and the extended impact of developmental hits up to adulthood. Testing guideline 443 would cover in a more detailed way specific endpoints, such as nervous system or immune function, according to the indications by previous studies.

In the EU, recognized testing methods represent those listed in the Commission Communication No. 2013/C 95/01 (developmental toxicity and neurotoxicity; OECD Test Guideline 426 is listed amongst the developmental toxicity studies). In Europe, the REACH regulation (Registration, Evaluation, Authorisation and Restriction of Chemicals) went into force in June 2007. The aim of REACH is to improve the protection of human health and the environment through the better and earlier identification of toxic properties of chemicals (Hartung, 2010a). The REACH legislation includes a systematic evaluation of chemicals that are produced in significant quantities within the European Union. It requires that producers and importers register all chemicals, produced in volumes greater than 1 ton per year at the European Chemicals Agency (EChA) based in Helsinki, Finland. This includes information on their properties, users' risks, and safe ways of handling. The chemicals of very high concern (e.g., bio-accumulative, carcinogenic, mutagenic, and reproductive toxic compounds) require specific authorizations before usage. Chemicals causing unmanageable risks will be phased out in the European Union by partial or total bans (European Commission, 2006; Hartung, 2010b). In the REACH testing scheme there is no direct requirement for DNT testing, although it is mentioned in the law text. Currently, there is pressure, especially from some Scandinavian regulators, to increase the DNT testing requirements – in case of any noticeable neurobehavioral changes observed during systemic toxicity evaluations, they

recommend testing for neurotoxicity and/or DNT according to the existing OECD test guidelines. Further testing should identify the chemicals with possible DNT effects and finally lead to the restricted use and control of the risk of exposure (Grandjean and Landrigan, 2006). Furthermore, Annex II of the Regulation No 1107/2009 concerning the placing of plant protection products on the market, DNT is considered "*a critical effect of particular significance*." The EU Scientific Committee for Food (SCF) has recommended that appropriate experts should set the criteria for when DNT testing is necessary.

Interestingly, several endocrine disruptors have been identified as DNToxicants (Masuo and Ishido, 2011; Choi et al., 2004; Weiss, 2011; Boas et al., 2009). This might prompt new discussions on DNT test requirements in the context of the ongoing US and planned EU endocrine disruptor screening programs.

Preclinical DNT studies for human pharmaceuticals are based on the International Conference on Harmonization guideline S7A[7] from 2000, which is used by the European Medicines Agency (EMA) and the US Food and Drug Administration (FDA).

Current DNT guidelines are entirely based on *in vivo* animal experiments, where observations are made to detect gross neurological and behavioral abnormalities, including the assessment of physical development. It is important to note, however, that they have not undergone formal validation in ring trials. It is estimated that about 1,000 rat pups are used for one DNT study and as a minimum 140 mated females are needed to produce enough pups. Of these 1,000 pups, approximately 640 are kept for at least 3 weeks and 240 pups are kept up to the young adult stage of 60 days postnatal. Together with the prenatal period, a DNT study lasts for three months. Although OECD acceptance of the extended one-generation study (OECD test guideline 443) reduced cost and time needs compared to two-generation studies TG 416 (though general feasibility is still challenged), it is obvious that animal test strategies remain complex and expensive in terms of scientific resources, time, and animal use.

OECD TG 426 uses rodents treated during gestation and lactation to cover critical pre- and postnatal periods of nervous system development. A core battery of tests was established to detect postnatal developmental disorders in these rats. Gross functional, behavioral, and neuroanatomical abnormalities are assessed during postnatal development and adulthood. The recommended endpoints can be divided into three groups: neurobehavioral testing, neuropathology, and pharmacokinetics. These assessments encompass physical development, behavioral ontogeny, motor activity, motor and sensory function, learning and memory, brain morphometry, and neuropathology. However, for learning and memory assessment the guidelines' methodology is flexible: "*The Guideline allows various approaches with respect to the assignment of animals exposed in utero and through lactation to functional and behavioural tests, sexual maturation, brain weight determination, and neuropathological evaluation. Other tests of neurobehavioural function (e.g., social behaviour), neurochemistry or neuropathology can be added on a case-by-case basis, as long as the integrity of the original required tests is not compromised*" (OECD, 2007). Therefore, its sensitivity varies, which leads to variability and subtle effects might remain undetected (Raffaele et al., 2010). Endpoints in risk assessment are standard motor activity and neuropathology, more or less sensory functions and only very limited approaches (e.g., anxiety test or simple mazes) exploring cognitive functions (Dr Rex Fitzgerald, personal communication, 2014).

We have earlier addressed the general shortcomings of *in vivo* tests (Hartung, 2008a) and will address here mainly those specific to DNT testing. The EFSA Plant Protection Products and their Residues (PPR) Panel recently addressed DNT in an opinion in general terms (though mainly addressing data on two substances only, EFSA, 2013), which gives a good summary of the problems of the test:

"*DNT guidelines are complex, time consuming, costly and not suitable for routine testing of high numbers of chemicals. Some concerns in terms of feasibility and animal welfare have been raised in the scientific literature. Although the protocol of the guidelines is well designed and covers a broad window of exposure, the critical phase for some effects might be missed and not all effects would be found. Furthermore, the interpretation of results is difficult because of knowledge gaps concerning normal brain development on the functional, structural and molecular levels, thus complicating risk assessment of compounds (Beronius et al., 2013). A number of issues related to the interpretation of DNT studies have been raised, such as excessive variability that may mask treatment-related effects and, conversely, minor statistically significant changes that can be considered as treatment-related when in fact they might fall within the normal range (Raffaele et al., 2008). All findings should be considered in the context of the study and interpreted in conjunction with other findings. Even statistically significant findings should be consistent with a pattern of effects (Tyl et al., 2008)... It is not completely clear whether a negative DNT test is sufficient to exclude a DNT potential of a chemical compound... There is considerable flexibility in TG 426 concerning the study design, such as the choice of behavioral tests included in the study, and also the design of the individual tests, e.g., strength of stimulus, intervals between testing trials and sessions, number of trials per day, etc. It is up to expert judgment of the investigator to design, on a case-by-case basis, the most sensitive and appropriate test relevant for the exposure and toxicity of the compound under investigation... This flexibility introduces potential sources of variability in DNT study design... A substantial amount of expertise is also required to interpret DNT study results, as well as to evaluate the reliability and relevance of DNT data for risk assessment*."

Moreover, there are also scientific concerns regarding the relevance of these studies for human health effects. Testing can be performed only at high doses, which are not relevant for human exposure scenarios and which often represent low dose exposures over prolonged time periods. The animal test also cannot reflect inter-individual (epi)genetic differences. Moreover, the interpretation of the behavioral effects and histology of *in vivo* data generated can be difficult and is less quantitative, which make it hard to predict human health effects. The study design is very demanding, often resulting in compromised study reports

[7] http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S7A/Step4/S7A_Guideline.pdf

(Crofton et al., 2004; Tyl et al., 2008). The statistical evaluation creates further challenges, especially because of the multiple testing issues with hundreds of significance tests required (Holson et al., 2008). From an industry perspective (Kaufmann, 2003), the "*extrapolation from basic biology without developmental toxicological validation*," the lack of a recommended optimal set of definitive techniques, and limited understanding of indirect (via maternal toxicity) impacts have been criticized.

Together, this results in low reliability, especially low reproducibility of even the positive control data (Crofton et al., 2004). The human relevance of behavioral study results often causes concern among scientists. Guidance of neurohistopathological examination in the DNT guideline also has some shortcomings, especially relating to methodology (Kuwagata, 2012). Consequently, current guidelines often do not provide sufficient information to facilitate regulatory decision-making.

Consideration 5: The consensus process toward alternative DNT assessment

Starting in 2005, groups of experts in the field and guiding institutions such as CAAT, ECVAM, or the US EPA have organized a series of international workshops/conferences (DNT1 to 3, with DNT4 to take place in May 2014) to discuss the current status and problems of developmental neurotoxicity assessment, identify promising alternative approaches, and provide recommendations for the future (Coecke et al., 2006, 2007; Lein et al., 2007; Hogberg et al., 2009, 2010; Bal-Price et al., 2010a; Crofton et al., 2011; Leist et al., 2008b; Kuegler et al., 2010). Key recommendations from DNT2 in 2008, for example, are summarized in box 1. This included the identification of critical elements of what is today referred to by OECD as adverse outcome pathways (AOP). DNT2 also developed an initial list of reference DNToxicants, and a preliminary uncommented compilation has been published (Tab. 1; Crofton et al., 2011). Similar work was done for example in the European ESNATS consortium (Kadereit et al., 2012). Consolidated lists with richer supporting information have been compiled by a CAAT-Europe workshop and they are currently being compiled for publication. Following DNT3, organized by ECVAM in Italy in 2011, CAAT will host the 2014 conference in Philadelphia. This represents the opportunity to further organize the existing knowledge in a set of AOPs (see below).

The next step toward a DNT testing strategy is based on the increasing understanding of mechanisms. The toxicity pathways identified so far include perturbation of: cell proliferation, precursor cell differentiation, glial reactivity, glial maturation (e.g., myelination), migration, axon/dendritic outgrowth, apoptosis, synapse formation, synapse pruning, neurotransmitter receptor profiles, and neuronal connectivity (Kadereit et al., 2012; Balmer and Leist, 2014; Leist et al., 2012). This now offers the opportunity to guide future research along the emerging AOP by first of all implementing measures of the suggested key events. Focusing on the key biological processes (Fig. 2) has the advantage of allowing for a test strategy without knowing the ends of AOPs, i.e., the molecular initiating event (MIE) and the final

Box 1: Recommendations from DNT-2

- *Key event of neurodevelopment:* Test methods should incorporate one or more endpoints that model key aspects of human neurodevelopment.
- *Endpoint measurement:* All test methods must demonstrate the ability to correctly and accurately measure the intended endpoint
- *Characterization of dynamic range:* One should determine the extent of change that can be detected for a DNT endpoint and whether both increases and decreases from untreated control can be measured.
- *Parametric controls:* Assay parameters that result in predictable changes in the endpoint should be characterized. These experimental parameters can be used to optimize the test method.
- *Response characterization:* The level of change in the response associated with an effect should be characterized. This is the degree of change that if exceeded results in a positive response (a "hit"). There are two ways to determine the positive response level. The first approach, commonly used in pharmaceutical screening, defines a hit as any response greater than 3 SD from the control. This conservative statistical approach is used to ensure a very small number of false positives: false positives would be costly to pursue. In toxicological screening and prioritization for further testing, it may be acceptable to have a higher rate of false positives. Thus, a second approach defines a positive response level based on biological relevance. Professional judgment should be used to balance the biological and statistical relevance of the response level.
- *Concentration range:* Each test method should be designed to characterize the concentration-response relationship. One recommendation is to minimally test five concentrations ranging from the solubility limit to five logs below the solubility limit. Concentration-response is critical to comparison of sensitivity between test methods, or endpoints within a test method.
- *Endpoint selectivity:* The ability of the test method to discriminate the endpoint of concern from other outcomes.
- *Endpoint-selective controls:* Endpoint-selective control chemicals reliably and consistently alter the endpoint by known mechanisms. Both positive and negative control chemicals should be tested. A positive control is a chemical or stressor that is known from previous experience to reliably affect the endpoint. A negative control is a chemical

that reliably causes no effect on the endpoint of interest. A negative control demonstrates the base-line result obtained when a test chemical does not produce a measurable positive result.
- *Training set of chemicals:* A training set of chemicals should be developed that includes chemicals known to reliably affect the endpoint of concern *in vitro*. Additional evidence from *in vivo* studies, if available, is highly recommended. Chemicals that reliably do not affect these endpoints must also be included. The goal of the training set is to evaluate the test method, including: 1) testing the practical ability of the method to efficiently process moderate numbers of chemicals; 2) confirmation of positive and negative controls; and 3) generation of historical control data to characterize the inherent response range for the endpoint.
- *Testing set of chemicals:* The testing set should include a large number of chemicals known to affect endpoints of developmental neurotoxicity *in vivo*, as well as chemicals that reliably do not affect developmental neurotoxic endpoints. This list should be large enough to: 1) demonstrate the ability of the method to rapidly and efficiently test large numbers of chemicals; 2) provide data that can be used in determining future steps in the process of regulatory acceptance of the endpoint and test method as part of an alternative testing strategy.
- *Specificity and Sensitivity:* Sensitivity is defined as the proportion of active substances that are correctly identified by the new test, and specificity is defined as the proportion of inactive substances that are correctly identified. Positive and negative predictivity are the frequencies of correct predictions obtained from the new tests.
- *High throughput:* The test method should hold the potential for automation and be more efficient than the current testing scheme (OECD Test Guideline 426, 2007) in regards to time and resources needed.
- *Documentation:* The test method needs to be fully documented and available to allow for implementation across laboratories.
- *Transferability:* The required resources need to be accessible and widely available to allow for implementation across laboratories.
- *Data sharing through open access databases:* It is extremely important for data from testing methods to be openly reported in publically accessible databases. This will allow inter-laboratory and intra-laboratory comparisons of test methods.

Tab. 1: Draft list of chemicals (from Crofton et al., 2011) to consider when developing new test methods for developmental neurotoxicity

Chemicals on this list have published or regulatory data from humans, non-human primates, or laboratory mammals suggestive of adverse neurological outcomes following developmental exposure. To be included on the list there had to be positive results from more than one laboratory. It is very important to take consideration 6.6 (see text) into account when choosing chemicals from this list for a specific test system. A validation of the usefulness of the compounds for *in vitro* testing has not yet been performed.

Acrylamide	Chlorine dioxide	Heptachlor	Permethrin
Aldicarb	Chlorpromazine	Hexachlorobenzene	Phenylacetate
Allethrin	Colcemid	Hydroxyurea	Phenylalanine
Aluminum	Colchicine	Imminodiproprionitrile (IDPN)	Phthalates
Amino-nicotinamide(6-)	Cytocine arabinoside	Lindane	Propylthiouracil
Amphetamine(d-)	DEET	LSD	Salicylate
Aspartame	Diamorphine hydrochloride	Maneb	Tellurium
Azocytidine	Diazepam	Methadone	Thalidomide
Benomyl	Diazinon	Methanol	Toluene
Benzene	Dieldrin	Methimazole (methylimidazole)	Triamcinolone
Bioallethrin	Diethylene glycol diethyl ether	Methoxyethanol, 2-	Tributyltin chloride
Bis(tri-n-butyltin)oxide	Diethylstilbestrol	Methylazoxymethanol	Trichlorfon
Butylated hydroxy anisol	Epidermal Growth Factor	Monosodium glutamate	Trichloroethylene
Butylated hydroxytoluene	Ethylene thiourea	Naloxone	Triethyllead
Carbamazepine	Flourouracil(5-)	Naltrexone	Triethyltin
Carbon monoxide	Fluoride	Nicotine	Trimethyltin
Chlordecone	Haloperiodol	Parathion	Trypan blue
Chlordiazepoxide	Halothane	PCBs	Urethane
			Vincristine

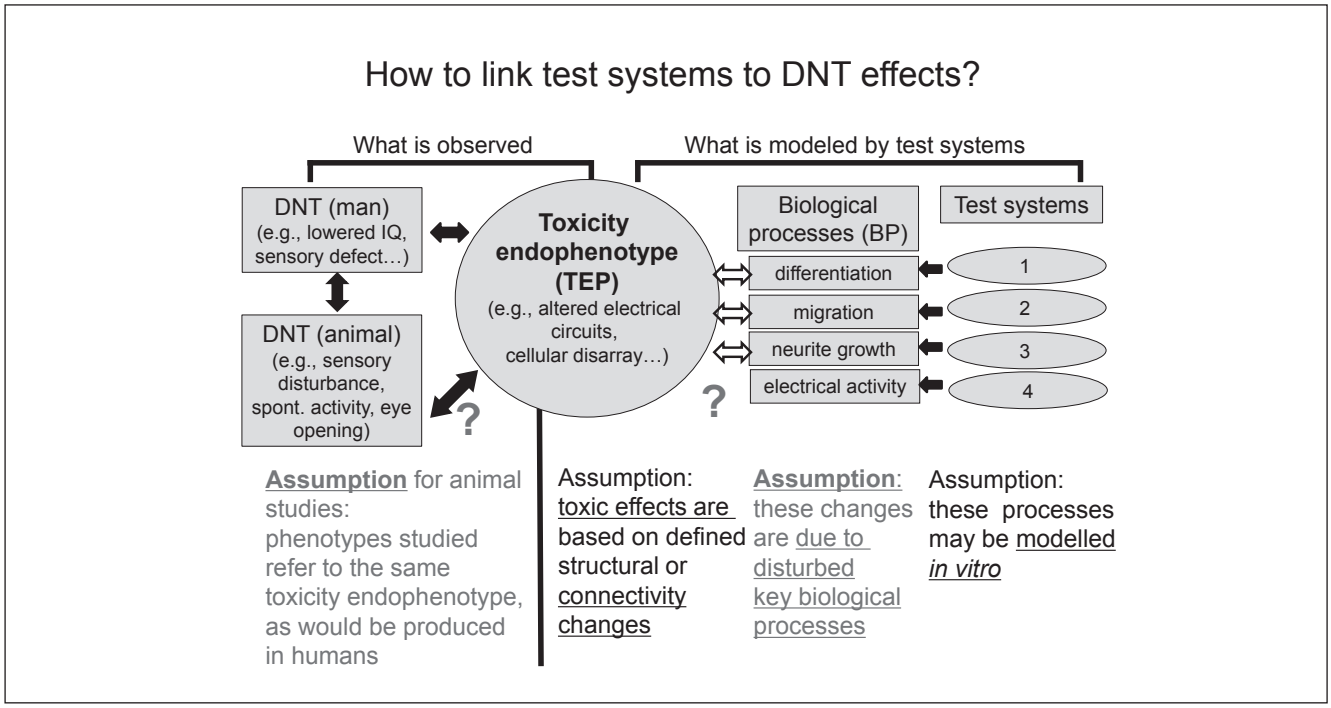

Fig. 2: Conceptual framework and basic assumptions that link *in vitro* test systems to DNT effects

For the majority of DNT issues, complete AOPs are hard to construct, as the molecular initiating events (MIE) are not clearly identified and the adverse outcome in humans is difficult to quantify and to assign to one specific agent. An alternative concept with emphasis on intermediate key events focuses on biological processes that can be tested *in vitro*. The main assumption is that the final disturbances of nervous system connectivity and function (the toxicity endophenotype (TEP)) are due to a disturbance of such processes. Further assumptions are that the human adverse outcome (e.g., lowered IQ) has a defined biological (morphological or biochemical) correlate to the TEP. When *in vitro* data are related to animal data, a further assumption is required: one has to assume that the endpoint measured in animals (e.g., altered motor activity) relates to the same TEP (e.g., basal ganglia dysfunction) as the relevant human endpoint (e.g., psychotic predisposition).

The characterization of the responses to established DNToxicants by high-content measurements (omics and high-content imaging) can detail our understanding of the respective pathways of toxicity (PoT). Adaptation to high-throughput screening (HTS) will enable the throughput necessary to validate the model and use it further for prioritization of chemicals. This will also expand our knowledge base of the PoT involved in DNT.

adverse outcome (e.g., cognitive phenotypes in humans). The biological processes link directly to a toxicity endophenotype (Kadereit et al., 2012; Balmer and Leist, 2014), i.e., a biological alteration of the nervous system, and they only require the basic assumption that nervous system development is disturbed when biological key processes are disturbed. This concept allows current, effective testing while AOPs and mechanistic knowledge are still being generated.

Consideration 6: What should we consider in order to bring forward effective and successful *in vitro* DNT testing methods?

From the previous considerations, it is obvious that there is an urgent need for cheaper, faster, and more mechanism-based approaches in DNT. The scientific community has embraced them and expects them to deliver: "*In-vitro methods have now reached a level of predictive validity that means they can be applied to neurotoxicity testing*" (Grandjean and Landrigan, 2014). Increasingly, also the regulatory community is considering them, e.g., the EFSA PPR panel (EFSA, 2013): "*In vitro assays may be regarded as complementary to animal testing because they may provide better understanding of the cellular/molecular mechanisms involved in developmental neurotoxicity*."

How we can make use of *in vitro* DNT tests to support epidemiology? There are several alternative *in vitro* models developed or under development today (Crofton et al., 2012). This includes simple two-dimensional cell lines to more complex primary cells, stem cells, 3D-cell cultures, and non-mammalian organisms. General cytotoxicity data are not good predictors of neurotoxicity, even if they are obtained from neuronal cultures (Krug et al., 2013a). DNT often manifests itself in functional disturbances that may appear hard to model *in vitro* (van Thriel et al., 2012). However, it is widely assumed (Bal-Price et al., 2012; Hogberg et al., 2009; Kadereit et al., 2012; Kuegler et al., 2010) that DNT is ultimately the consequence of the disturbance of relatively basic biological processes, such as differentiation, proliferation, migration, and neurite growth. Therefore, several *in vitro* systems have been established which test the disturbance of such biological activities by chemicals (Leist et

al., 2013; Hoelting et al., 2013; Balmer et al., 2012; Frimat et al., 2010; Harrill et al., 2011a,b; Radio et al., 2008; Zimmer et al., 2011a,b). Endpoints that have found a lot of attention are neurite outgrowth (Radio and Mundy, 2008) and electrophysiology with microelectrode arrays (van Vliet et al., 2007; Hogberg et al., 2011), as such assays using functional endpoints can specifically identify and characterize DNToxicants (Stiegler et al., 2011; Krug et al., 2013a).

Quite remarkable is the apparent predictivity of zebrafish as a model for DNT screening (de Esch et al., 2012), showing high correlation with more established animal models as well as with human data. It appears that tests can be performed efficiently, reproducibly, and reliably at an early life stage and offer the possibility to combine assays at biochemical, cellular, and molecular levels with observations at a functional and even behavioral level within an individual organism in time. Although most of the embryological processes and molecular pathways in development have been conserved between zebrafish and mammalian species, the question remains whether the model is predictive for human responses. Pharmacokinetics need to be considered, but the main challenge remains to map pathways across species to ultimately understand interspecies extrapolation.

A number of challenges exist for *in vitro* DNT testing, which shall be addressed in the following sections.

Consideration 6.1: Appropriate dosing

Just as with any other toxicological model, DNT test systems require careful consideration of dosing parameters. However, some issues are unique to this area or require particular attention. The first is the dosing resulting in a specific readout. In this context it is important to consider whether dose or concentration is the more appropriate measure. In most cases, the use of nominal concentrations is the most practical approach (Blaauboer et al., 2012). However, with highly hydrophobic compounds, such as methylmercury, the cell dose can be much higher than expected from the nominal concentration (Zimmer et al., 2011b). With compounds that are rapidly degraded or metabolized (e.g., retinoids in some systems), the opposite can be observed. The second thought should be the choice of nominal concentrations that result in specific effects, i.e., that do not cause mere cytotoxicity. In many studies, test concentrations are chosen as highest non-cytotoxic concentrations. This requires establishment of a cytotoxcity concentration/effect-relationship and then determination of an EC_{10} (Krug et al., 2013b), for example, or a benchmark concentration. A mathematical framework for incorporation of measures of variance into the determination of such cytotoxicity benchmarks is not well established, and this may lead to some of the discrepancies of test data in the literature. In transcriptomics studies, measurements too far into the cytotoxic range can yield erroneous results, but measurements at too low concentrations can dramatically reduce the power of the study (Waldmann et al., 2014). Choice of test concentrations with relation to maximal non-cytotoxic concentrations has advantages due to the simplicity of the rule, but there are also major drawbacks to be considered. One of them is that cytotoxicity can depend on small details of the experimental system (number of cells plated, medium used, time of exposure) and all these would therefore affect the final DNT test concentration. Another drawback is that meaningless high concentrations may be tested if a compound has a low cytotoxicity. In such cases, anchoring to physiological (or toxicologically expected) concentrations would make more sense. Alternatively, full concentration-responses could be tested (Hermsen et al., 2013; Schulpen et al., 2013; Waldmann et al., 2014). However, this is often prohibitively expensive for some endpoints (transcriptomics, metabolomics) if large screens are performed.

Consideration 6.2: What is the role of exposure timing?

Exposure timing should depend on the endpoint used and on the question asked, but there are also some basic rules to be followed. For instance, it is a great advantage if the endpoint parameter assessed changes monotonously during the exposure period. Otherwise data can become difficult to interpret, or they require a lot of background information for interpretation. Another fundamental consideration is that short exposures are more likely to give information on the direct signaling changes triggered by a test chemical, while long exposures are likely to yield mainly information on the adaptations of the test system to the toxic challenge (Balmer and Leist, 2014). Such system-wide descriptions may be derived from metabolomics or transcriptomics measurements (Balmer et al., 2012; Ramirez et al., 2013; Krug et al., 2013b; Kuegler et al., 2010; Zimmer et al., 2011a; Hermsen et al., 2013; Schulpen et al., 2013; Robinson et al., 2012; Bouhifd et al., 2013). For short exposures, it needs to be noted, the test system may show different responses at different times, as the cell population present in the test systems changes.

This timing of exposure and measurements is particularly important if there are distinct windows of sensitivity. For instance, the sensitivity of differentiating stem cells to mercury can differ when the exposure time covers either the period of neural stem cell formation or of neuronal maturation. The same applies to other toxicants, such as lead or cyclopamine (Zimmer et al., 2011a,b). The effects can be dramatic, as in the case of exposure to histone deacetylase inhibitors. For instance, exposure of human stem cells during the first four days of differentiation completely altered the differentiation track, while exposure from day 4-6 or from day 4-10 had no effect (Balmer et al., 2012; Balmer and Leist, 2014).

Consideration 6.3: How to deal with time-offset?

The area of developmental toxicity, and in particular DNT, poses a big challenge for epidemiology, risk assessment, and predictive testing, as the relevant effects can have a large time offset relative to the exposure. In other words, toxicity can become apparent when no toxicant is present. This difficulty is also known from the field of carcinogenesis, for example, when initial exposure can lead to cancer after 20 years. For instance, a study of Basha et al. (2005) showed a link between neonatal lead exposure with ensuing acceleration of aging as evidenced by earlier appearance of amyloid deposition, demonstrating how gestational exposure can profoundly affect late stages of life.

Several such late consequences of early life exposure have been demonstrated (Balmer and Leist, 2014). They provide proof for the Barker hypothesis[8] (originally developed for cardiovascular risk) in the field of DNT (Grandjean and Landrigan, 2014; Schug et al., 2013), i.e., this "*fetal programming hypothesis*" suggests that early lifetime exposure can lead to late disease manifestations.

The issue of nervous system plasticity is linked to the question of the "time offset." The nervous system can adapt to damage and it can show a very high plasticity that can make up for effects of damage over time. It is still unclear how these factors can be modeled appropriately, be it in rodent models or in *in vitro* models. For *in vitro* models, this means that cytotoxicity can be missed in cases when increased proliferation leads to similar final cell numbers as in control conditions. Dead cells can also be missed when phagocytosis takes place in the cultures (e.g., in primary cultures containing microglia). Phagocytosis of dead cells is an important aspect of nervous system development (Hirt et al., 2000; Hirt and Leist, 2003), and it contributes to overall plasticity.

An important mechanistic aspect deals with the memory of damage: how does the (disturbed) system "know" that it has been disturbed earlier by a chemical? For the establishment of AOP for DNT, answers to this question are essential. In some case, "memory" may not need a molecular encoding. If cell migration is disturbed during a critical phase of development, then cells end up in the wrong places and this can be sufficient to account for delayed DNT effects. In other cases, there may be molecular correlates of damage memory. For instance epigenetic modifications have been suggested to account for disturbed neurodevelopment in a stem cell based system (Balmer et al., 2012).

Consideration 6.4: How to deal with cell heterogeneity?

More than any other organ the CNS is characterized by interaction between multiple cell types. The brain contains dozens of cell types, and at first thought it may seem desirable to model as much heterogeneity as possible. As in many other situations, however, the most pragmatic rule is: *as much as necessary, as little as possible*. Sometimes, cultures containing only a single cell type may be sufficient to answer a specific question. For instance, neurite outgrowth has often been measured in monocultures (Harrill et al., 2011a; Radio et al., 2008; Stiegler et al., 2011). Human LUHMES cells, for instance, have been used for this purpose in combination with high content imaging (Krug et al., 2013a). Such cells are also available in several variants expressing fluorescent proteins or reporters, which facilitate easy and fast readouts of assays (Schildknecht et al., 2013). Also, engineered murine embryonic stem cells (mESC) have been used as single cell system allowing high-throughput DNT readouts (Kern et al., 2013). More complex systems would not allow such readouts and the throughput would be compromised. Complex multicellular systems introduce additional problems and limitations for analyzing omics data sets (metabolomics, transcriptomics, etc.). Working with mixed populations demands considering the composition of the system to understand which cells were perturbed. Within a tiered testing strategy, more complex and simpler cell systems all find their space. This also implies coverage of all, or at least the most important cell types of the brain. Often there is a strong focus on neurons only. However, astrocytes and microglia are important inflammatory and immunomodulatory cells, and require some consideration (Falsig et al., 2004; Kuegler et al., 2012; Defaux et al., 2009, 2011). In addition, astrocytes provide guidance for axons and synapse formation (Aschner et al., 1999). A cell type also of high importance, but little considered for DNT assays until now, is the oligodendrocyte (Fritsche et al., 2005; Defaux et al., 2009, 2011). In essence, a compromise needs to be found between multiple cell types as possible target and to constitute brain function and simplicity of the model necessary for standardization and throughput.

Consideration 6.5: How may adversity be defined in experimental systems?

The crucial distinction to be made is between a significant effect and a toxicologically relevant effect of a substance. A living system will respond to stressors (toxicants) in order to mitigate possible hazard. Many of these responses might be exploratory, sensing the impact of the stressor or quite generally mobilizing defense pathways without necessarily indicating harm taking place. It is very tempting to consider any stress response as indicator of toxicity. This will lead, however, to an overestimation of adversity, as often no manifestation of hazard will take place in response to such minor exposure. Therefore, it is necessary to distinguish the perturbed pathways of toxicity (PoT), i.e., molecular initiating events leading to adverse outcomes, from those not leading to adversity. Thresholds of adversity need to be defined for this purpose (Boekelheide and Andersen, 2010).

Apart from the question whether an observed effect is considered to be adverse, there is a second dimension to the question of adversity – how can adverse effects be ranked and quantified? This question starts becoming complex when more than one endpoint is used in an assay. It is largely unsolved for transcriptomics studies. These mostly report qualitative changes, but only few studies have attempted to rank compounds by potency or to develop a DNT toxicity index from such data (Waldmann et al., 2014; Schulpen et al., 2014). This issue will certainly require further investigations if *in vitro* DNT test data are to support a quantitative read-across of different related compounds or if a comparison of diverse compounds is attempted to find the least toxic candidate (Krause et al., 2013).

Similarly, it has to be asked, what constitutes an adverse effect as a point of reference (Hoffmann et al., 2008) for validating DNT tests? For example, it is not clear if any alterations in MRI scanning studies on altered human brain responses represent ad-

[8] David Barker was Professor of Clinical Epidemiology at the University of Southampton, UK and Professor in the Department of Cardiovascular Medicine at the Oregon Health and Science University, US. Twenty years ago, he showed for the first time that people who had low birth weight are at greater risk of developing coronary heart disease. In 1995, the *British Medical Journal* named this the "Barker Hypothesis."

versity. The fundamental approach that all deviations from normal are adverse if not ruled out, i.e., the precautionary principle, comes with the burden of many false-positives to be replaced by substances with possibly more favorable toxicological profiles.

The more we learn with time (see, for example, carcinogenesis, where not every mouse liver tumor is considered a problem, or chronic toxicity, where phospholipidosis or initial reversible liver toxicity is not necessarily considered a problem for humans), the more our experience can help to judge which effects relate to adversity in humans. As Douglas Coupland nicely phrased it: "*Where does personality end and brain damage begin?*"

Consideration 6.6: How is DNT related to cytotoxicity and organ-specific toxicity?

Thomas A. Edison once stated "*The chief function of the body is to carry the brain around*," indicating that the brain is somehow special among organs. However, the question needs to be raised whether this holds true also for its vulnerability to toxicants. Can general cytotoxicity serve as an estimate of neurotoxicity and can neurotoxicity with some safety factors be estimated from other organ toxicities?

For *in vivo* toxicology data, the key question for interpretation of studies and extrapolation to man is: is a potential DNT observation a primary effect of chemical exposure, and thus of potentially high relevance to man, or is it a secondary effect of other toxicities that may not occur in man? A simple practical parameter is whether DNT effects are observed in the absence of maternal toxicity. If this is the case, data are "clean" and easy to interpret. If DNT is only observed at doses that lead to maternal toxicity, the situation is more complicated. Interpretation and regulation may then differ depending on the background of the study.

Some examples may illustrate the situation: (a) many forms of maternal toxicity, e.g., hepatotoxicity or pulmonary toxicity, may lead to general wasting and weight loss in the fetus, possibly associated with a misbalance of nutrients and vitamins in the blood. This would then lead to secondary developmental defects (DNT) of low relevance to man. In practice, such situations are not easy to judge as there are typically not more than three doses in a given DNT study, and it may easily occur that one is too low to show DNT effects, and the next higher one shows DNT effects and maternal toxicity. There is no way to distinguish whether the DNT effects in this case are specific, or whether they are a secondary consequence; (b) Warfarin can show DNT effects in animals. The compound is a vitamin K antagonist and affects mainly the generation of blood clotting factors in the liver, and leads to internal bleeding. This is a most likely reason for indirect DNT effects. However, there are also some vitamin K-dependent enzymes that could be directly involved in nervous system development. There are insufficient scientific data for this model toxicant to resolve the issue; (c) methylazoxymethanol (MAM) has a very simple straightforward mechanism of action. It forms electrophiles that react with DNA and block cell division. The compound is highly cytotoxic to dividing cells. One would not assume such a compound to cause specific DNT, but it does do so, and it is actually one of the best tool compounds for *in vivo* studies. This illustrates several important points. First, MAM would be maternally toxic in a standard dosing regime, and it would in standard regulatory studies most likely not be classified as a DNToxicant. The active metabolite of MAM has a very short half-life, and kills only cells dividing at exactly that time. Thus it can be used in a single dose, like a knife-sharp tool, to kill only a specific neuronal precursor subpopulation in the embryo that divides during the few hours when the compound is present. This can result in a neuropsychiatric phenotype without other types of toxicity (Penschuck et al., 2006). A simple cytotoxicant can therefore in some cases be a specific DNT compound. Another example of such compounds is ethanol, which can selectively kill neural crest cells or a subpopulation of NMDA receptor-positive central neurons, and thereby cause later alterations in neuronal connectivity and organization; (d) thyroid toxicants such as methimazole or thioureas can be highly organ-specific. They may destroy the thyroid without any significant effects on the nervous system. However, thyroid hormones play a major role in the formation of the nervous system. Thus, thyroid toxicants are amongst the best-established DNT compounds although their effects are only indirect.

These four examples give a good indication of what has to be considered to select test compounds for *in vitro* studies and to interpret data from *in vitro* testing. The indirect toxicants must necessarily be avoided when specificity or sensitivity of systems is assessed that can only show direct effects on neurons. Toxicity to the thyroid, placenta, or to functions important in feeding might adversely affect pre- and postnatal neural development. Indirect toxic effects may also be caused by interferences of chemicals with oxygen or glucose supply of the developing brain and the subsequent effects of these events on the developing neural tissue. Interferences with glucose circulation and the subsequent transport into the brain have been observed for neurotoxins like dichloroacetic acid (Moser et al., 1999) and some compartments of the brain seem to be more vulnerable for this indirect mechanism (e.g., some thalamic nuclei). Many astrocyte-specific toxicants affect their energy metabolism (fluoroacetate, 6-aminonicotinamide, fluorocitrate). This can lead to secondary neurotoxicity. Other compounds affect oligodendrocytes or blood vessels and trigger indirect neurotoxicity that may be region-specific but cannot be observed in purely neuronal *in vitro* systems. A particularly interesting example is the insecticide fipronil, which shows developmental neurotoxicity and notochord degeneration in zebrafish at concentrations around 1 μM. This apparent DNT effect is indeed due to a block of GABA or glycine receptors in the neuromuscular system. This leads to muscle cramps that are so strong that they damage the notochord. Thus, non-neural cells/organs are involved here in triggering developmental neurotoxicity by mechanical damage (Stehr et al., 2006). Substances known to act predominantly via such indirect mechanisms are not suitable for the validation of *in vitro* test systems.

For *in vitro* systems, it is highly important to be aware of cytotoxicity effects, as opposed to more specific functional effects. Therefore, general and unspecific cytotoxic compounds such as cytostatic drugs (e.g., 5-FU), detergents, or inducers of apoptosis like staurosporin should be included in the initial compound

list as controls for such effects. They act as a special type of negative controls, as they do have an effect, but not the relevant specific toxic effect, in the test system (Leist et al., 2010). Selection of this class of compounds requires great care and adaptation to the respective assay and purpose. In many cases, compounds killing cells by excessive production of reactive oxygen species or by blocking energy regeneration may be considered non-specific cytotoxicants. However, there are examples that oxidative stress controls neurodifferentiation (Yan et al., 2009) or specifically affects a (dopaminergic) subpopulation of immature human neurons (Hansson et al., 2000). Also, apparently unspecific mitochondrial toxicants show specific effects on defined neuronal subtypes (Sherer et al., 2007), and mitochondrial respiratory chain inhibitors affect, e.g., cardiac or neuronal differentiation (Krug et al., 2013a; San Martin et al., 2011).

Consideration 6.7: Immune contributions to DNT and its testing

Increasingly, inflammation and immunotoxicity as adverse outcome pathways are recognized in the etiology of ASD (Onore et al., 2012; Enstrom et al., 2010; Goines and Ashwood, 2013; Depino, 2013). The striking overlap of immune and DNToxicants suggests that environmental factors might manifest also via this axis. Indeed, several well established immunotoxicants are also suspected DNToxicants: lead, arsenic, methyl mercury, organotins (TBTO, i.e., bis(trisn-butyltin)oxide), benzene, various pesticides, ethanol, cannabinoids, cocaine, opioids, 1,1,2-trichloroethane, dexamethasone, and PCBs, among others. Furthermore, immunomodulation as a non-neuronal path has increasingly come into the foreground as an AOP of ASD. Species differences, it should be noted, are especially pronounced for immune and inflammatory mechanisms (Leist and Hartung, 2013) as they are under enormous evolutionary pressure. If a future AOP-based testing strategy for DNT is established, inflammatory and immune tests should be considered.

Consideration 6.8: Gene/environment interactions in DNT

Since many neurodevelopmental disorders, including ASD, do not have a clear or known genetic basis, they appear at least partially to be due to gene/environment interactions. Undoubtedly, there is a substantive genetic component for example to ASD etiology: gene mutations, genetic anomalies, copy number variants, and other genetic anomalies have been linked to autism (Landrigan, 2010). Autism appears to be a family of diseases with common phenotypes linked to a series of genetic anomalies, each of which is responsible for no more than 2-3% of cases and the total fraction of ASD attributable to genetic inheritance has been estimated at about 30-40% (Landrigan et al., 2012). At the same time, findings from neuropathology, brain gene expression, twin and sibling concordance/recurrence risk analyses, as well as proof-of-principal evidence from studies of now-rare teratogens, all suggest that environmental influences operating in the prenatal period also have a substantial impact on ASD risk. Exposures during early life are of concern since scientific evidence shows that drugs and environmental chemicals contribute to the increasing incidence of neurodevelopmental disorders in children (Kuehn, 2010; Sagiv et al., 2010; Grandjean and Landrigan, 2006; Landrigan, 2010).

The effects of genetic background cannot be tested in standard animal tests, which use inbred rodents ("identical twins"). A new prospect comes from the use of induced pluripotent stem cells (iPSC). Using iPSC from patients with developmental disorders, such as ASD, makes it possible to test substance sensitivity in the same model with different genetic backgrounds and answer the question whether this particular genetic makeup makes the person more sensitive to the environmental stressor. Thus, using *in vitro* models based on iPSC may determine whether gene/environmental interaction indeed take place in the development of certain disorders. Another advantage of iPSC as a model of drug sensitivity is that their differentiation *in vitro* is similar to the stages of brain development *in utero* (Nat and Dechant, 2011). Each stage of neurodevelopment is unique and displays different sensitivities to different xenobiotics. An *in vitro* model to test chemicals for toxicity during development must thus be able to screen at these different stages. Pluripotent stem cell differentiation to telencephalic neurons has been defined in three stages that are distinguished by changes in morphology and expression of transcription factors and structural genes (Liu and Zhang, 2011).

Consideration 6.9: Is there also DNT for the peripheral nervous system?

Most DNT studies focus on the central nervous system. One reason may be that the development of the peripheral nervous system is intricately linked to the development of other organs and disturbed development is often not classified as DNT. Very early during development, when the neural tube (precursor of the central nervous system) forms, some of the cells (positive for the transcription factor SOX10) on its upper part undergo an epithelial-to-mesenchymal transition. These so-called neural crest cells (nestin-positive like central neural stem cells) are the precursors of the peripheral nervous system (sympathetic, sensory, pain fibers, intestinal nervous system), but they also form some parts of the facial bone and cartilage. Cleft palate is a typical neural crest-related developmental disorder, but it is not classically defined as a DNT effect. On the other hand, neural crest cells are also involved in the proper closure of the neural tube, and a failure of this process can result in hydrocephalus or spina bifida. Also, these are not always counted as DNT effects. This separation is primarily historical, whereas there is a good biological rationale to combine all effects related to the development of the nervous system or the function of its precursors. Neural crest developmental toxicity has been investigated extensively in different model organisms, but human data are rare due to the limited availability of fetal tissue or neural crest cells. The advent of pluripotent stem cell technology (Leist et al., 2008a) has allowed the generation of neural crest cells from embryonic stem cells, for example. These have been used for establishing a functional DNT assay and for the generation of peripheral neurons that may be used for further test systems (Zimmer et al., 2012; Leist et al., 2013). From the use of primary animal cells there is strong evidence that chemicals can affect neurite growth of peripheral sensory neurons (Howard et al., 2005). Similar effects are also seen

in zebrafish (Yang et al., 2011). The latter model also showed that motor neurons were particularly affected. This cell type belongs classically to the central nervous system (with the cell bodies in the spinal cord), but they extend far into the periphery, and they cause clearly peripheral effects when adversely affected. Other peripheral neurons that may be affected in their development by chemicals are the sympathetic neurons or their ganglia organization (Kim et al., 2009).

Consideration 6.10: How to deal with species differences in model systems?

For developmental toxicities, concordance of animal species is low (in the 60% range) (Basketter et al., 2012). A primary question is how is this reflected on a cellular level? Is it necessary to model rodent cells to link to rodent *in vivo* data?

We need lists of marker genes for developmental stages and cell types for different species. Such markers have been compiled for DNT testing in murine cells (Kuegler et al., 2010; Schulpen et al., 2014) or zebrafish (Hermsen et al., 2013), for example, and they can be easily compared to the respective human markers established in a multitude of studies. Recently discovered microRNA can contribute to marker gene lists, since many of them are conserved through the phyla. The phylogenetic conservation of miRNA was described as an important feature of these small regulatory molecules (Bartel, 2004; Lee et al., 2007). For example, let-7 and lin-4 (mir-125 for mammals), the first discovered miRNAs in *C. elegans* (Lee and Ambros, 2001), are conserved in mammals, including humans, and regulate developmental timing in both *C. elegans* and humans. Mir-124 and mir-9, the two most abundant miRNA in the brain, are conserved between species and regulate neurogenesis (Shi et al., 2010). So, by using such gene lists we can at least partially overcome the problem of interspecies differences.

Few studies have explored species differences in DNT *in vitro* test systems. A prominent example is the use of neurospheres consisting of neural precursor cells. These cells can be obtained from different species, including humans, and this culture system has been used to explore species differences. It allows measurement of different endpoints (e.g., cell differentiation or cell migration). Major differences were obvious between murine and human neurospheres when the response to A_h-receptor agonists was tested (Fritsche et al., 2011). Use of *in vitro* data from human and rodent cells, in combination with the pre-existing *in vivo* data in rodents, allows improved extrapolation to man by a parallelogram approach.

Consideration 6.11: How to deal with quality control and validation for DNT screening?

The topic of quality control is not unique to DNT, but it is still worth keeping in mind when considering an *in vitro* approach to DNT. Good Cell Culture Practices (GCCP) have been developed (Hartung et al., 2002; Coecke et al., 2005; Hartung and Zurlo, 2012; Hartung, 2013) for the different aspects of quality assurance of experimental *in vitro* work, because of the limitations of these technologies (Hartung, 2007b, 2013). Proper reporting of *in vitro* work (Leist, 2010) is also required. The formal validation of *in vitro* tests (Hartung et al., 2004; Hartung, 2007a; Leist et al., 2012) has established a path for providing the evidence that non-animal tests can be used without compromising patient and consumer health. Most recently, a proposal was put forward to establish "mechanistic validation" (Hartung et al., 2013b) in order to quality assure a test not by showing correlation of results for a necessarily limited number of test agents but by demonstrating that it reflects current understanding of pathophysiology. As discussed in the next section, some tools derived from evidence-based medicine are useful here when applied as an "evidence-based toxicology" (Hartung, 2010c).

Consideration 7: How can the Tox-21c concept be applied to DNT?

Despite major advances in biotechnology, molecular biology, and information technology, the underlying conceptual framework of toxicology has not changed over the previous several decades (Leist et al. 2008b; Hartung and Leist, 2008). Today, companies and agencies still largely use animal studies to assess toxicological risk, despite their costs (in the US about $1 billion a year) and inability to test large numbers of chemicals or their combinations in mixtures. This has led to the current lack of toxicological information required to safeguard human health and enable regulatory decision-making on chemicals. Toxicologists have become increasingly aware of these limitations and have brought forward new concepts and innovative approaches to overcome them. They promise a rather revolutionary than evolutionary change (Hartung, 2008b). Novel approaches to regulatory toxicology are now often summarized as Toxicology for the 21st Century (Tox-21c). For some this term sounds comforting as it leave some 86 years to achieve the goal. However, these are, in fact, a number of quite different developments which are worth discussing in the context of DNT. The origin of this terminology was the NRC report *Toxicity Testing for the 21st Century: A Vision and a Strategy* (NRC, 2007). The first aspect to be discussed, therefore, is advancing DNT to the pathway-based approaches of Tox-21c. The report's main recommendation was to use mechanistic models (*in vitro*, lower organisms, *in silico*) and combine them with kinetic modeling and targeted *in vivo* testing. This has been strongly embraced in the adverse outcome pathway (AOP) concept furthered by OECD. Second, and more or less coinciding, the European REACH debate brought about awareness for the need of Integrated Testing Strategies (ITS) (Hartung et al., 2013b). OECD has extended this to a concept of Integrated Approaches to Testing and Assessment (IATA), which – though the exact definition is still pending – also includes exposure, kinetics, and risk assessment. ITS are a logical consequence of combining a number of pathway-based assays (Hartung, 2009a), though the European REACH discussion saw more the need to combine in ITS also existing information, *in silico* approaches, and (finally) animal studies. Similarly, the 2013 EFSA panel on DNT saw opportunities for preliminary evaluation, initial chemical prioritization identifying "the alerts" for DNT, and in information on the cellular/molecular mecha-

nisms. But, in general, the panel "*believes that* in vitro *tests currently used cannot substitute for* in vivo *DNT tests. To date, no in* vitro *test can be used to set health-based reference values*." Importantly, however, they encouraged "*the definition of clear and consistent criteria at EU level to trigger submission of mandatory DNT studies, which could include development of an integrated and cost-effective, tiered testing strategy composed of robust, reliable and validated* in vitro *assays and alternative methods complementary to* in vivo *studies*." Third, technological advances in organotypic cell cultures, stem cell-derived human tissue, and biological phenotyping with information-rich technologies (high-content and high-throughput) change the way we can approach modeling pathophysiology.

Consideration 7.1: Toward mechanism / AOP-based DNT testing

The NRC Tox-21c vision report, now principally forming the novel EPA toxicity-testing paradigm, suggests moving to a pathway-based testing strategy away from traditional animal-based methods (Hartung, 2009a,b, 2010a, 2011). Work aiming for the quality assurance and validation of novel approaches represents a key interest. Ongoing work of CAAT and the Human Toxome Project consortium (http://humantoxome.com) (Bouhifd et al., 2014), as well as with EPA ToxCast, aims to establish the identification and annotation of pathways to create a public database (Hartung and McBride, 2011; Hartung et al., 2012). At this moment, metabolomics and transcriptomics are the tools closest to pathway identification. These approaches have not been applied to DNT to a major extent. In addition, knowing the role of small, non-coding RNAs, especially miRNAs, in maintenance of crucial cellular processes and cascades (Chua et al., 2009; Rana, 2007), and the increasing number of publications elucidating the role of miRNAs in the cellular response to environmental stress, including xenobiotics (reviewed in Smirnova et al., 2012), miRNA profiling has shown itself to be a tool that should be included in the identification of PoT as well as compound-mediated aberrant physiology.

Concepts for an AOP framework first emerged at OECD in the context of ecotoxicology (Ankley et al., 2010) but quickly were combined with the Tox-21c concept to extend to all regulatory toxicology. An AOP stretches from exposure to chemical properties, molecular interactions with cells (molecular initiating events), the effects on cellular, tissue, and organism level, and, lastly, to population effects. This current understanding of the toxicological mechanism is mainly on a narrative level, referencing the scientific literature. AOP represent a structured organization of current knowledge on mechanism/mode of action in hazard manifestation in the developed template. The concept of PoT, in contrast, is to develop molecular annotations. AOP cover chemical properties to population effects while PoT only covers the cellular events. A workshop held as part of the Human Toxome Project (Kleensang et al., 2014) developed the following working definition: "*A Pathway of Toxicity is a molecular definition of the cellular processes shown to mediate adverse outcomes of toxicants*."

There is some debate whether a PoT represents a chemico-biological interaction impacting on the biological system or the perturbed normal physiology; these likely reflect the early molecular initiating events versus the homeostasis under stress, which establishes in response (Hartung et al., 2012; Kleensang et al., 2014). The differences between AOP and PoT are summarized in Table 2.

This is not meant to belittle AOP – they are the best that can be done now, i.e., compiling and evaluating our current knowledge. PoT are a hypothesis, and the first PoT has yet to be defined, validated, and agreed upon. As much as the comparison shows the shortcomings of where we are today, it shows the challenges of where we want to go. The current developments at OECD level to organize our knowledge on hazard manifestations as Adverse Outcome Pathways (AOP) from exposure, chemical properties to molecular initiating events, and further to key events in pathogenesis and finally the population effects, has not yet been adapted to DNT. About 20 test cases are in preparation under the auspices of OECD. CAAT has made a proposal to OECD for the development of such an AOP for DNT, with a decision pending. At the same time, there is a discussion to apply the concept of "mechanistic validation" (Hartung et al., 2013b) to the qualification of AOP.

Systematic reviews and meta-analyses have been key to transforming healthcare from a patchwork of inconsistently applied knowledge into a coherent and more standardized practice. CAAT hosts the secretariat of the Evidence-based Toxicology Collaboration (http://www.ebtox.com), which aims to bring systematic reviews and other tools of Evidence-based Medicine into toxicology (Hoffmann and Hartung, 2005, 2006). Although systematic literature reviews have played a role in regulatory toxicology, they have not yet been fully integrated into the standard practice of research toxicologists, and, as a result, there is often

Tab. 2: Comparison of Adverse Outcome Pathways and Pathways of Toxicity

AOP	PoT
Spans from exposure to population effects	Spans from molecular initiating events to cell and tissue effects
Narrative, low level of detail	Molecular, high level of detail
Biased by existing knowledge	Untargeted identification, causality
Not quantitative, no flux, no dynamics	Aiming for quantitative relations, fluxes
No validation yet	Mechanistic validation by Evidence-based Toxicology suggested

little consensus about molecular mechanisms in AOP/PoT. Even worse, there is no convenient, vetted resource for researchers to turn to when looking for information about a proposed pathway or mechanism, and no way to incorporate the information gleaned into a systems-level analysis.

Focusing first on the five clearly identified human DNToxicants (Grandjean and Landrigan, 2006),a preliminary literature review was carried out (see Tab. 3): lead, methylmercury, PCB, arsenic, and toluene provide a number of established molecular initiating events (e.g., interference with glutamate receptors, binding to anti-oxidative enzymes, mitochondrial accumulation), which converge into a few toxicity pathways (oxidative stress, calcium disturbance, impaired neurotransmission, impaired energy metabolism, and glia disturbance) and the associated adverse outcomes on a cellular and *in vivo* level (e.g., change in membrane biophysics, excitotoxicity, apoptosis, disturbed neurotransmission (dopaminergic, cholinergic, glutaminergic), disturbed neurite outgrowth, glial and neuronal cell loss). We envisage combining a more systematic literature review with expert discussion to further expand this map of DNT-AOPs.

Systematic literature reviews consists of five steps, which should also be performed to define and document known or pos-

Tab. 3: Molecular Initiating Events and related key events of DNT AOP
A literature review for the first five established human DNToxicants Lead (L), Methymercury (M), PCB (P), Arsenic (A), and Toluene (T)

	Molecular Initiating Event/Key events	**Cell Adverse Outcomes**	***In vivo* Correlate**	**Human (Histo-) Pathology**
Oxidative stress (clear evidence for L, M, P, A)	Inhibition delta-aminolevulinic acid dehydratase: L (1)	Oxidative stress, ROS formation, lipid peroxidation of membrane (defense AP-1, NFkB): L (1,2,5,6), M (14)	Oxidative stress, brain damage, and impaired antioxidative defense (reduced GSH, induced defense AP-1, NFkB): L (1,2,5), M (14,15), P (17), A (18)	M is GSH-bound in erythrocytes, A leading to urinary 8 OHdG and plasma lipid peroxidation (21)
	Inhibition of SOD, catalase, GSH peroxidase, GSH (via SH-binding) and Ca replacement at EF motifs and C2 domains): L (2,6), M (11,13), P (17), A (18)			
	Mitochondrial accumulation and dysfunction: L (6), M (11,13)			
Membrane effects (likely secondary to oxidative stress)	Interaction with neg-charged membrane phospholipids: L (1)	Changed membrane biophysics (leading to iron-mediated lipid peroxidation): L (1)	Changed myelin membrane fluidity: L (1)	Membrane rigidity (erythrocytes): L (1)
Ca disturbance and replacement (clear evidence for L, M, P)	Electronegativity, binding to Sulfur and Oxygen (substituting for calcium and zinc): L (1,4), M (6,14,15)	Changed calcium fluxes (also leading to ROS from mitochondria), stimulation calmodulin and cAMP phosphodiesterase: L (1), M (13), P (17)		
	Interaction with glutamate (synergy on PKC): L (1,3), M (13), P (17)			
	Inhibition of Ca-ATPase: L (6)			
	Impaired Ca channels: M (11), P (17)			
Impaired neuro-transmission (clear evidence for L, M, P, T)	Competition with Ca at NMDAR and nNOS inhibition: L (1,2), P (17)	Reduced nNOS: L (1), increased NOS: M (13)	Reduced nNOS: L (1,2), increased NO (13)	

	Molecular Initiating Event/Key events	Cell Adverse Outcomes	*In vivo* Correlate	Human (Histo-) Pathology
Impaired neuro-transmission (clear evidence for L, M, P, T)	Interaction with glutamate (excito-toxicity, inhibition of uptake): L (1,3), M (13), T (24)	Excitotoxicity: M (13)	Increased glutamate levels: M (13)	
	Inhibition of neuro-transmitter receptors (a-disintegrin and metalloprotease inhibition in membrane): M (11), T (24)	Disturbed neuro-transmission (dopaminergic, cholinergic, glutaminergic): L (1,2,4), M (16), P (17), T (24)	Disturbed neuro-transmission (dopaminergic, cholinergic, glutaminergic): L (1,2,4), M (13,16), P (17), T (24)	
	Competition with Ca at NMDAR and disturbed Ca transport as well as direct calcineurin and calmodulin stimulation: L (1,2,4,6)			
	Inhibition of voltage-dependent Ca channels: L (4,6)			
Impaired energy metabolism (clear evidence for L, M, P, A)	Interference with PKC phosphorylation: L (2,8), M (16)	Reduced cell proliferation: L (8), M (3,12,16), A (21)		
		Increased transendothelial permeability: L (2)	Edema, brain damage: L (2)	
		Neurite outgrowth impaired: M (3,16), A (19), T (24)	Altered neurogenesis and migration: T (24)	
		Impaired cytoskeleton: M (3,16), A (18,21)		
	Accumulation in mitochondria (pore opening, depo-larization, CytC and Ca release): L (6,8), M (13,14), A (21)	Impaired energy metabolism: L (6), M (3)		
		Apoptosis: L (3,8), M (3,11,14,16), P (17), A (21), T (22)	Apoptosis: M (14), T (22)	
		Reduced cell proliferation: M (3,16)		
		Excitotoxicity by glutamate: L (8)		
Glia disturbance (Clear evidence for L, M, T)	Decrease in CNPase activity: L (8)	Oligodendrocyte toxicity and delayed development: L (8), M (12)	Hypomyelination: L (8)	
	Astrocytic accumulation: L (8), M (15)	Astrocyte toxicity: L (8), T (23)	Loss of astrocytes: M (14), T (23)	

AOP span from exposure, chemico-physical properties to the initiating and adverse events depicted here and further to clinics and population effects. For the latter there is evidence for all substances: M (14,15), P (17), A (18) P (17), A (18), T (24).
Sources (mainly recent review articles):
1. Verstraeten, 2008; 2. Nava-Ruiz et al., 2012; 3. Giordano and Costa, 2012; 4. Marchetti, 2003; 5. Baranowska-Bosiacka, 2012; 6. Garza et al., 2006; 7. Senut et al., 2012; 8. Lidsky and Schneider, 2003; 9. Guzzi and La Porta, 2008; 10. Grandjean, 2007; 11. Bland and Rand, 2006; 12. Ceccatelli et al., 2013; 13. Farina et al., 2011a; 14. Farina et al., 2011b; 15. Ceccatelli et al., 2010; 16. Johansson et al., 2007; 17. Fonnum and Mariussen, 2009; 18. Vahter, 2008; 19. Wang et al., 2010; 20. de Vizcaya Ruiz et al., 2009; 21. Flora, 2011; 22. Nielsen et al., 2003; 23. Burry et al., 2003; 24. Win-Shwe and Fujimaki, 2010

tulated PoT and biomarkers from the literature in general and for known developmental neurotoxicants: framing the questions to be addressed, identifying the scope of relevant work, judging the quality of studies, determining inclusion and exclusion principles, giving a clear summary of the weight of the evidence, and interpreting the findings (Khan et al., 2003). Any annotation-based data analysis is necessarily limited to a relatively modest discovery of novel information, since it is, by definition, dependent on existing knowledge and further limited by the extent to which that knowledge is captured in the relevant databases. The resources and tools of the Evidence-based Toxicology Consortium (EBTC) are critically important here.

In conclusion, the concept of AOP, largely embraced by regulators at OECD, opens the door for tests and testing strategies for chemical safety which are based on mechanism. However, current AOP are on a narrative level of description, similar to a very structured textbook. They mainly describe linear sequences of events leading to hazard manifestation. Thus, identification and mapping of PoT within a given AOP is of high importance, since this aims for a molecular definition of mechanism and the perturbed networks. They allow modeling and intervention studies to test causality and ultimately modeling in a systems toxicology approach (Hartung et al., 2012). While there have been relatively few attempts to map PoT, the concepts and tools are emerging in the context of endocrine disruptor work of the Human Toxome consortium and DNT might benefit from this.

Consideration 7.2: Integrated Testing Strategies

We have discussed earlier the opportunities of Integrated Testing Strategies in more general terms (Hartung et al., 2013a). While single, standalone assays can rarely satisfy a regulatory information need, an integrated use of various information sources promises to approximate this task much better. This was also one of the key recommendations of the consensus process toward animal-free systemic toxicity testing, it is important to note, which started off with five white papers and a workshop (Basketter et al., 2012) endorsed by multiple stakeholders in Brussels 2012 and Washington 2013.

At this moment, there are two main concepts in the design of ITS: the favored one aims to base its components on the established mechanisms of a health effect, e.g., the AOP; and alternatively, the endpoints of current guideline studies leading to classifications can be addressed (Bremer et al., 2007). In the case of DNT, the small number of available studies leading to regulatory decisions largely rules out this possibility. The emerging AOP thus represents the prime opportunity for a rational design of an ITS with building blocks representing key events of the AOP. The emerging tools for data integration, including Bayesian networks, machine learning tools, and sensitivity analysis, should allow continuous optimization of the ITS. Given the already visible multitude of key events and targets of the DNT AOPs, we should aim for some building blocks to cover a number of them.

Consideration 7.3: The biotech revolution – organotypic cultures, stem cells, and high-content methods

Recent developments in three-dimensional (3D) cell culturing materials and techniques coupled with the advances in knowledge on stem cell differentiation allow the development of complex tissue structures in controlled conditions. There is increasing evidence that these 3D culture systems more accurately capture the complex physiology of an *in vivo* tissue or organ than two-dimensional (2D) cell monolayers. 3D cell culture clearly improves the physiological relevance of cell-based assays and the comparability between *in vitro* cultures and living organisms (Pampaloni et al., 2007). 3D cell culture models have advanced our understanding of the molecular and cellular mechanisms underlying toxicity and have great potential for use as powerful tools for assessing the impacts of exposure to chemicals (Lee et al., 2009; Cui et al., 2007; Sun et al., 2006; Dhiman et al., 2005).

Traditional 2D cell culture models are generally derived from established cell lines or primary cells freshly isolated from either animal or, if available, human tissue samples. Indeed, the differentiation of human embryonic stem cells represents a promising approach. The purity and functional properties of these cell cultures, however, remains unresolved. This is particularly a limitation for neurotoxicity studies since cell-cell interactions and biochemical signaling between neurons, astrocytes, and microglia play an important role in mechanisms of neurotoxicity (Aschner and Kimelberg, 1991; Giordano et al., 2009) or cell survival (Kirchhoff et al., 2001). In contrast, the 3D rat primary neural cell model of aggregating brain cell cultures was shown to closely reflect *in vivo* morphology and biochemical signaling (Trapp et al., 1979). Aggregating brain cell cultures are prepared from rat fetal forebrain tissue that is dissociated and spontaneously aggregated under rotation-mediated culture conditions (Honegger et al., 1979; Honegger and Monnet-Tschudi, 2001). The aggregated cultures include all the different cell types of the CNS in a 3D structure as previously demonstrated by electron microscopy (Trapp et al., 1979). Characterization of the cultures over time using biochemical assays revealed processes of neurodevelopment (Honegger and Monnet-Tschudi, 2001). Preliminary data from our laboratory confirmed neurodevelopmental processes, as significant changes in mRNA levels of specific marker genes over the time followed by the maturation of different cell types. The cell model was previously used for a variety of mechanistic neurotoxicity studies (Eskes et al., 2003; Monnet-Tschudi et al., 1996; Zurich et al., 2004, 2010). These studies demonstrated the important role of cell-cell interactions between neurons and glial cells in mechanisms of neurotoxicity. Recently, aggregated cultures have been combined with emerging advanced technologies to study neurotoxicity including multi-electrode array-based electrophysiological recordings of neuronal activity (van Vliet et al., 2007) and metabolomics (van Vliet et al., 2008). Hence the 3D primary rat aggregating brain cell cultures represent a promising model for the mechanistic study of DNT. This is in contrast to the animal model, which can only provide limited mechanistic information because of its complexity, limited

access to the cell level, and number of replicates. Moreover, the *in vitro* model is able to significantly reduce the number of animals needed for DNT assessment. A single preparation generates thousands of aggregates, while in principle only a few aggregates are needed to test a compound at a specific concentration, especially when applying sensitive techniques such as RT-PCR and mass spectrometry based metabolomics. The lowest observed effect concentrations (LOEC) can be used for pharmacokinetic modeling that introduces pharmacokinetic parameters to predict relevant tissue concentrations and the exposure scenarios that would lead to these levels. Based on these predictions it should become clear what the potential risks of drugs and chemicals are for public health.

There is an enormous potential of human iPSCs and other stem cells to enhance human risk prediction (Bremer and Hartung, 2004). The use of human cell models in toxicology is crucial for cost and throughput reasons and has the advantage of overcoming interspecies differences. The comparison of *in vitro* and human *in vivo* DNT data will provide useful information on the predictive capacity of this neuronal *in vitro* model. Such an innovative approach in line with the Tox-21c testing paradigm (NRC, 2007) is expected to provide better and more precise information for human risk assessment, and regulatory decision-making, than the current extrapolations based on high-dose animal models. The generation of human *in vitro* DNT data will provide useful information complementing studies with rat cells, allowing a comparison to animal results. One of the most promising sources to obtain human *in vitro* models is induced pluripotent stem cells (iPSCs). iPSC from different individuals (e.g., healthy donors vs. donors with ASD) can also provide a testing model with the ability to predict substance sensitivity in different genetic backgrounds (see Section 6.8). Another advantage of iPSC is that their differentiation *in vitro* is similar to the stages of brain development *in utero* (Nat and Dechant, 2011). Each stage of neurodevelopment is unique and displays different sensitivities to different xenobiotics. Thus, an *in vitro* model for developmental toxicity testing must be able to screen at these different stages.

To complement the organotypic culture methods, which are obviously more complex, laborious, and thus allow fewer replicates and throughput, high-content methods, which generate maximum information, are ideal. These information-rich methods include high content imaging (HCI), which allows evaluation of the effects of toxicants on key cellular processes of neural development and physiology, as these types of assays more directly correlate with neuropathology observed in *in vivo* DNT studies (Harrill, 2011b). A number of omics technologies represent further high-content methods enabling holistic biological phenotyping. Previous work demonstrated that gene expression can be used as a sensitive endpoint to detect chemicals that induce DNT in primary cultures of rat cerebellum granule cells at concentration levels relevant for human exposure (Hogberg et al., 2009, 2010).

More recently, miRNA profiling emerged and has already shown some utility in DNT testing. In the last decade the post-transcriptional regulation of gene expression has emerged thanks to the discovery of miRNAs, small (~22nt) non-coding regulatory RNA molecules that bind to specific binding sites in 3'UTR of target mRNA and repress their translation. About 2500 miRNAs have been identified in humans (Kozomara and Griffiths-Jones, 2011). Increasing evidence demonstrates the importance and significance of miRNA networks in coordination and fine-tuning of gene expression with high temporal and spatial specificity (reviewed in Bartel, 2004). More than 50% of all identified miRNAs are expressed in the brain. These miRNAs play a important role in brain development and morphogenesis by regulating developmental timing, cell differentiation and proliferation, and cell fate determination (reviewed in Li and Jin, 2010). There is no longer any doubt that perturbations in miRNA expression patterns have a significant impact on several disorders, including different types of cancer and neurodegenerative and neurodevelopment disorders (Alzheimer's, Parkinson's, Huntington's disease, autism) (reviewed in De Smaele, 2010). It has been shown that miRNA targets are twice as likely to be sensitive to changes in expression levels following environmental chemical exposure than those mRNAs which lack miRNA binding sites (Wu and Song, 2011). All miRNA have relatively short half lives, and are able to regulate hundreds of genes, making them candidate molecules which are able to react rapidly to environmental stress. Including miRNA profiling in transcriptomics studies for DNT may have stronger prediction than only mRNA profiling, as miRNAs mark developmental timing and cell specification (reviewed in Bartel, 2004). In addition, considering the relatively low number of known miRNAs, it can be much easier to interpret the results and make a statement about the substance effects. Exploding numbers of studies have demonstrated that miRNA are involved in numerous cellular processes and open new perspectives for analysis of pathways of toxicity. This field is rapidly progressing with the discovery of new miRNA functions in neural development and its perturbation. Significant roles of miRNA in embryonic stem cell differentiation and specification during neural development (Smirnova et al., 2005; Wulczyn et al., 2007; Rybak et al., 2008, 2009), as well as the role of miRNA in developmental neurotoxicity (for example, the effects of the antiepileptic drug and known developmental neurotoxicant valproate on neural differentiation of murine embryonic stem cells (Smirnova et al., 2014) or methyl mercury effects on human NT2 carcinoma pluripotent stem cells differentiation (Pallocca et al., 2013), have been shown.

Metabolomics studies represent another major technology for phenotyping biological responses to DNToxicants. While proteomics and transcriptomics data can inform us of the potential changes of the cellular machinery and infrastructure, direct data on altered metabolite levels or metabolite fluxes through certain pathways can only be provided by metabolomics analysis. Both intracellular and extracellular metabolomics can be used, and possibly be combined with other sets of omics data, for detailed information for identification of pathways for DNT and quantification of their activation/disturbance (Ramirez et al., 2013). A mechanistic and quantitative understanding of the links between the intracellular or extracellular concentrations of the chemical species involved in pathways of toxicity (PoT) can guide the choice of specific and sensitive early biomarkers for future

development of high-throughput screening methods. Specific alterations of neuronal metabolism that are linked to neurotoxicity can occur at lower toxicant concentrations than those required to kill cells (Krug et al., 2014). Such changes are particularly informative for PoT, as they occur when not all metabolic pathways in a cell are disturbed by the initiated death process. However, they may be missed by many assay endpoints, and only metabolomics data sets as outcome parameters can fully leverage this information.

For risk evaluation, it is important to determine a relevant point of departure, i.e., a concentration of toxicant resulting in a relevant change of the cells or their respective metabolome. The concentrations of the chemical compounds triggering specific neurotoxic changes are sometimes rather low and some of them are even close to the limit of quantification (LOQ) for most of the analytical methods used in biokinetics studies. This is an important issue to be solved for risk assessment (Blaauboer et al., 2012).

The next challenge will be integrating multi-omics technologies for DNT studies on a systems biology level. This kind of integrated approach would lead to a global assessment of adverse effects, indicating the potential of systems biology in terms of pharmacological and toxicological research. Quantitative measurement with multi-omics technologies will bridge the gap between molecular initiating events and relevant adverse outcomes. In addition, this kind of integrated approach will be a significant step towards the better understanding of the mechanisms underlying DNT, which could have profound impact on DNT chemical screening.

Conclusions

The traditional animal-based testing strategy is expensive in both time and cost, and provides only limited toxicity testing information, particularly for developmental neurotoxicity studies. Various compounds – most of them in the categories of metals, solvents, and pesticides – have been shown to have neurotoxic effects, but very few substances have been identified as developmental neurotoxicants (Grandjean and Landrigan, 2006; Breier et al., 2010; Crofton et al., 2011). The complicated structure of the central nervous system and the critical lack of knowledge of neurotoxic mechanisms are the major obstacles in this field. New directives and initiatives for developmental toxicity testing in the United States and Europe will rely increasingly on an integrated and intelligent new testing strategy (Hartung et al., 2013a) utilizing cell-based *in vitro* approaches (Krewski et al., 2010).

DNT has a strong overlap with other disciplines, such as developmental biology and neurobiology, and similar steps towards the development of new technologies and approaches are taking place. Bioinformatics plays a key role in mining the information-rich new technologies and making sense of the output by modelling. With interdisciplinary collaboration, toxicology can take advantage of such expert knowledge. The challenge and the opportunity lie in the transition from MoA models to pathway modelling, i.e., increasing the resolution of analysis, and then building ITS on this understanding of adverse outcome pathways and, ultimately, a systems integration of this mechanistic knowledge. Environmental contaminations do not present themselves in isolation but as mixtures with unknown "cocktail" effects. Traditional animal test approaches are not suitable for testing many combinations of doses and timing. New pathway-based tests, in contrast, could allow the identification of critical combinations and provide better environmental protection.

DNT affects only a certain percentage of individuals despite similar exposures, which argues for individual sensitivities. Such phenomena cannot be studied with inbred rats but only can be addressed properly if the respective pathways are understood, enabling an individualized toxicology. This will ultimately protect vulnerable subpopulations, making the use of certain consumer products more sustainable and improving resilience toward exposures, for example, by enabling avoidance of certain workplace exposures. Many exposures to DNToxicants are unavoidable. The availability of fast and inexpensive assays to monitor cleaning efforts is also critical for guiding such remediation.

In the meantime, several newly emerging technologies have demonstrated the capabilities for the development of more modern approaches for toxicology to replace the traditional "black box" animal-based paradigms by providing mechanistic details of events at the cellular and molecular levels. Such high-content methods are the logical complement to sophisticated organotypic cultures, where a maximum of information is obtained from the lower number of replicates because of duration of model preparation and technical effort for each and every parallel cell system.

Acknowledgements

We are grateful to many colleagues for insightful discussions and comments. This work was supported by grants from the Doerenkamp-Zbinden Foundation (M. L. and T. H.), the FP7 EU program ESNATS (M. L.), the NIH Transformative Research Grant "Mapping the Human Toxome by Systems Toxicology" (RO1ES020750, T. H.), and FDA grant "DNTox-21c Identification of pathways of developmental neurotoxicity for high throughput testing by metabolomics" (U01FD004230, H. H. and T. H.).

Disclosure Statement

The authors declare no conflict of interest.

References

Adinolfi, M. (1985). The development of the human blood-CSF-brain barrier. *Dev Med Child Neurol 27*, 532-537. http://dx.doi.org/10.1111/j.1469-8749.1985.tb04581.x

Al-Gadani, Y., El-Ansary, A., and Attas, O. (2009). Metabolic biomarkers related to oxidative stress and antioxidant status in Saudi autistic children. *Clin Biochem 42*, 1032-1040. http://dx.doi.org/10.1016/j.clinbiochem.2009.03.011

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. http://dx.doi.org/10.1002/etc.34

Aschner, M. and Kimelberg, H. K. (1991). The use of astrocytes in culture as model systems for evaluating neurotoxic-induced-injury. *Neurotox 12*, 505-517.

Aschner, M., Allen, J. W., Kimelberg, H. K., et al. (1999). Glial cells in neurotoxicity development. *Annual Rev Pharmacol Toxicol 39*, 151-173. http://dx.doi.org/10.1146/annurev.pharmtox.39.1.151

Austin, D. W. and Shandley, K. (2008). An investigation of porphyrinuria in Australian children with autism. *J Toxicol Env Health, Part A 71*, 1349-1351. http://dx.doi.org/10.1080/15287390802271723

Balmer, N. V., Weng, M. K., 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. http://dx.doi.org/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*, Epub ahead of print. http://dx.doi.org/10.1111/bcpt.12203

Bal-Price, A. K., Hogberg, H. T., Buzanska, L., and Coecke, S. (2010a). Relevance of in vitro neurotoxicity testing for regulatory requirements: challenges to be considered. *Neurotoxicol Teratol 32*, 36-41. http://dx.doi.org.proxy3.library.jhu.edu/10.1016/j.ntt.2008.12.003

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

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

Baranowska-Bosiacka, I., Gutowska, I., Marchlewicz, M., et al. (2012). Disrupted pro- and antioxidative balance as a mechanism of neurotoxicity induced by perinatal exposure to lead. *Brain Res 1435*, 56-71. http://dx.doi:10.1016/j.brainres.2011.11.062

Bartel, D. P. (2004). MicroRNAs. Genomics, biogenesis, mechanism, and function. *Cell 116*, 281-297. http://dx.doi.org/10.1016/S0092-8674(04)00045-5

Basha, M. R. (2005). The fetal basis of amyloidogenesis: Exposure to lead and latent overexpression of amyloid precursor protein and b-amyloid in the aging brain. *J Neurosci 25*, 823-829. http://dx.doi:10.1523/JNEUROSCI.4335-04.2005

Basketter, D. A., Clewell, H., Kimber, I., et al. (2012). A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing – t[4] report. *ALTEX 29*, 3-91. http://dx.doi.org/10.14573/altex.2012.1.003

Beronius, A., Johansson, N., Rudén, C., and Hanberg, A. (2013). The influence of study design and sex-differences on results from developmental neurotoxicity studies of bisphenol A: Implications for toxicity testing. *Toxicol 311*, 13-26. http://dx.doi:10.1016/j.tox.2013.02.012

Bjørling-Poulsen, M., Andersen, H. R., and Grandjean, P. (2008). Potential developmental neurotoxicity of pesticides used in Europe. *Environ Health Persp 7*, 50. http://dx.doi.org/10.1186/1476-069X-7-50

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

Bland, C. and Rand, M. D. (2006). Methylmercury induces activation of Notch signaling. *Neurotoxicol 27*, 982-991. http://dx.doi.org/10.1016/j.neuro.2006.04.005

Breier, J. M., Gassmann, K., Kayser, R., et al. (2010). Neural progenitor cells as models for high-throughput screens of developmental neurotoxicity: State of the science. *Neurotoxicol Teratol 32*, 4-15. http://dx.doi.org.proxy3.library.jhu.edu/10.1016/j.ntt.2009.06.005

Boas, M., Main, K. M., and Feldt-Rasmussen, U. (2009). Environmental chemicals and thyroid function: An update. *Curr Opin Endocrinol Diabetes Obes 16*, 385-391. http://dx.doi.org/10.1097/MED.0b013e3283305af7

Boekelheide, K. and Andersen, M. E. (2010). A mechanistic redefinition of adverse effects – a key step in the toxicity testing paradigm shift. *ALTEX 27*, 243-252. http://www.altex.ch/All-issues/Issue.50.html?iid=121&aid=2

Bondy, S. C. and Campbell, A. (2005). Developmental neurotoxicology. *J Neurosci Res 81*, 605-612. http://dx.doi:10.1002/jnr.20589

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

Bouhifd, M., Hogberg, H. T., Kleensang, A., et al. (2014). Mapping the human toxome by systems toxicology. *Basic Clin Pharmacol Toxicol*, Epub ahead of print. http://dx.doi.org/10.1111/bcpt.12198

Boyle, C. A., Decouflé, P., and Yeargin-Allsopp, M. (1994). Prevalence and health impact of developmental disabilities in US children. *Pediatrics 93*, 399-403.

Bremer, S. and Hartung, T. (2004). The use of embryonic stem cells for regulatory developmental toxicity testing in vitro – the current status of test development. *Curr Pharm Des 10*, 2733-2747. http://dx.doi.org/10.2174/1381612043383700

Bremer, S., Pellizzer, C., Hoffmann, S., et al. (2007). The development of new concepts for assessing reproductive toxicity applicable to large scale toxicological programs. *Curr Pharm Des 13*, 3047-3058. http://dx.doi.org/10.2174/138161207782110462

Burns, C. J., McIntosh, L. J., Mink, P. J., et al. (2013). Pesticide exposure and neurodevelopmental outcomes: review of the epidemiologic and animal studies. *J Toxicol Environ Health, Part B, Crit Revi 16*, 127-283. http://dx.doi.org/10.1080/10937404.2013.783383

Burry, M., Guizzetti, M., Oberdoerster, J., and Costa, L. G. (2003). Developmental neurotoxicity of toluene: In vivo and in vitro effects on astroglial cells. *Develop Neurosci 25*, 14-19. http://dx.doi.org/10.1159/000071463

Ceccatelli, S., Daré, E., and Moors, M. (2010). Methylmercury-

induced neurotoxicity and apoptosis. *Chem-Biol Interact 188*, 301-308. http://dx.doi.org/10.1016/j.cbi.2010.04.007

Ceccatelli, S., Bose, R., Edoff, K., et al. (2013). Long-lasting neurotoxic effects of exposure to methylmercury during development. *J Intern Med 273*, 490-497. http://dx.doi.org/10.1111/joim.12045

Chua, J. H., Armugam, A., and Jeyaseelan, K. (2009). MicroRNAs: Biogenesis, function and applications. *Curr Opin Mol Therap 11*, 189-199.

Choi, S. M., Yoo, S. D., and Lee, B. M. (2004). Toxicological characteristics of endocrine-disrupting chemicals: Developmental toxicity, carcinogenicity, and mutagenicity. *J Toxicol Environ Health B Crit Rev 7*, 1-24. http://dx.doi.org/10.1080/10937400490253229

Coecke, S., Balls, M., Bowe, G., et al. (2005). Guidance on good cell culture practice. *Altern Lab Anim 33*, 261-287.

Coecke, S., Eskes, C., Gartlon, J., et al. (2006). The value of alternative testing for neurotoxicity in the context of regulatory needs. *Environ Toxicol Pharmacol 21*, 153-167. http://dx.doi.org/10.1016/j.etap.2005.07.006

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 Persp 115*, 924-931. http://dx.doi.org/10.1289/ehp.9427

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. http://dx.doi.org/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. http://dx.doi.org/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 52*, 140-146. http://dx.doi.org/10.1111/j.1741-4520.2012.00377.x

Cui, Z., Xu, X., Trainor, N., et al. (2007). Application of multiple parallel perfused microbioreactors and three-dimensional stem cell culture for toxicity testing. *Toxicol In Vitro 21*, 1318-1324. http://dx.doi.org/10.1016/j.tiv.2007.05.015

de Esch, C., Slieker, R., Wolterbeek, A., et al. (2012). Zebrafish as potential model for developmental neurotoxicity testing. *Neurotoxicol Teratol 34*, 545-553. http://dx.doi.org/10.1016/j.ntt.2012.08.006

De Smaele, E., Ferretti, E., and Gulino, A. (2010). MicroRNAs as biomarkers for CNS cancer and other disorders. *Brain Res 1338*, 100-111. http://dx.doi.org/10.1016/j.brainres.2010.03.103

de Vizcaya Ruiz, A., Barbier, O., Ruiz-Ramos, R., and Cebrian, M. E. (2009). Biomarkers of oxidative stress and damage in human populations exposed to arsenic. *Mutation Res 674*, 85-92. http://dx.doi.org/10.1016/j.mrgentox.2008.09.020

Decouflé, P., Boyle, C. A., Paulozzi, L. J., and Lary, J. M. (2001). Increased risk for developmental disabilities in children who have major birth defects: a population-based study. *Pediatrics 108*, 728-734. http://dx.doi.org/10.1542/peds.108.3.728

Defaux, A., Zurich, M.-G., Braissant, O., et al. (2009). Effects of the PPAR-beta agonist GW501516 in an in vitro model of brain inflammation and antibody-induced demyelination. *J Neuroinflamm 6*, 15. http://dx.doi.org/10.1186/1742-2094-6-15

Defaux, A., Zurich, M.-G., Honegger, P., and Monnet-Tschudi, F. (2011). Minocycline promotes remyelination in aggregating rat brain cell cultures after interferon-γ plus lipopolysaccharide-induced demyelination. *Neurosci 187*, 84-92. http://dx.doi.org/10.1016/j.neuroscience.2011.04.053

Depino, A. M. (2013). Peripheral and central inflammation in autism spectrum disorders. *Mol Cell Neurosci 53*, 69-76. http://dx.doi.org/10.1016/j.mcn.2012.10.003

Dhiman, H. K., Ray, A. R., and Panda, A. K. (2005). Three-dimensional chitosan scaffold-based MCF-7 cell culture for the determination of the cytotoxicity of tamoxifen. *Biomaterials 26*, 979-986. http://dx.doi.org/10.1016/j.biomaterials.2004.04.012

DiPietro, N. A. (2010). Methods in epidemiology: observational study designs. *Pharmacotherapy 30*, 973-984. http://dx.doi.org/10.1592/phco.30.10.973

EFSA (2013). Developmental neurotoxicity potential of acetamiprid and imidacloprid. *EFSA Journal 11*, 3471-3524. http://www.efsa.europa.eu/en/efsajournal/doc/3471.pdf

Ema, M., Fukui, Y., Aoyama, H., et al. (2007). Comments from the developmental neurotoxicology committee of the Japanese teratology society on the OECD guideline for the testing of chemicals, proposal for a new guideline 426, developmental neurotoxicity study, draft document (October 2006 version), and on the draft document of the retrospective performance assessment of the draft test guideline 426 on developmental neurotoxicity. *Congenit Anomal 47*, 74-76. http://dx.doi.org/10.1111/j.1741-4520.2007.00148.x

Enstrom, A. M., Onore, C. E., Van de Water, J. A., and Ashwood, P. (2010). Differential monocyte responses to TLR ligands in children with autism spectrum disorders. *Brain Behavior Immun 24*, 64-71. http://dx.doi.org/10.1016/j.bbi.2009.08.001

Eskes, C., Juillerat-Jeanneret, L., Leuba, G., et al. (2003). Involvement of microglia-neuron interactions in the tumor necrosis factor-alpha release, microglial activation, and neurodegeneration induced by trimethyltin. *J Neurosci Res 71*, 583-590. http://dx.doi.org/10.1002/jnr.10508

European Commission (2006). Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02006R1907-20130701&qid=1395826642786

Falsig, J., Latta, M., and Leist, M. (2004). Defined inflammatory states in astrocyte cultures: Correlation with susceptibility towards CD95-driven apoptosis. *J Neurochem 88*, 181-193. http://dx.doi.org/10.1111/j.1471-4159.2004.02144.x

Farina, M., Rocha, J. B. T., and Aschner, M. (2011a). Mecha-

nisms of methylmercury-induced neurotoxicity: Evidence from experimental studies. *Life Sci 89*, 555-563. http://dx.doi.org/10.1016/j.lfs.2011.05.019

Farina, M., Aschner, M., and Rocha, J. B. T. (2011b). Oxidative stress in MeHg-induced neurotoxicity. *Toxicol Appl Pharmacol 256*, 405-417. http://dx.doi.org/10.1016/j.taap.2011.05.001

Flora, S. J. (2011). Arsenic-induced oxidative stress and its reversibility. *Free Radical Biol Med 51*, 257-281. http://dx.doi.org/10.1016/j.freeradbiomed.2011.04.008

Fowke, J. H. (2009). Issues in the design of molecular and genetic epidemiologic studies. *J Prevent Med Public Health 42*, 343-348. http://dx.doi.org/10.3961/jpmph.2009.42.6.343

Frimat, J. P., Sisnaiske, J., Subbiah, S., et al. (2010). The network formation assay: a spatially standardized neurite outgrowth analytical display for neurotoxicity screening. *Lab Chip 10*, 701-709. http://dx.doi.org/10.1039/b922193j

Fritsche, E., Cline, J. E., Nguyen, N.-H., et al. (2005). Polychlorinated biphenyls disturb differentiation of normal human neural progenitor cells: clue for involvement of thyroid hormone receptors. *Environ Health Persp 113*, 871-876. http://dx.doi.org/10.1289/ehp.7793

Fritsche, E., Gassmann, K., and Schreiber, T. (2011). Neurospheres as a model for developmental neurotoxicity testing. *Meth Molec Biol 758*, 99-114. http://dx.doi.org/10.1007/978-1-61779-170-3_7

Fonnum, F. and Mariussen, E. (2009). Mechanisms involved in the neurotoxic effects of environmental toxicants such as polychlorinated biphenyls and brominated flame retardants. *J Neurochem 111*, 1327-1347. http://dx.doi.org/10.1111/j.1471-4159.2009.06427.x

Fung, M., Thornton, A., Mybeck, K., et al. (2001). Evaluation of the characteristics of safety withdrawal of prescription drugs from worldwide pharmaceutical markets – 1960 to 1999*. *Drug Information Journal 35*, 293-317.

Ganz, M. L. (2007). The lifetime distribution of the incremental societal costs of autism. *Arch Pediatr Adolesc Med 161*, 343-349. http://dx.doi.org/10.1001/archpedi.161.4.343

Garza, A., Vega, R., and Soto, E. (2006). Cellular mechanisms of lead neurotoxicity. *Med Sci Monitor 12*, RA57-65.

Giordano, G., Kavanagh, T. J., and Costa, L. G. (2009). Mouse cerebellar astrocytes protect cerebellar granule neurons against toxicity of the polybrominated diphenyl ether (PBDE) mixture DE-71. *Neurotoxicol 30*, 326-329. http://dx.doi.org/10.1016/j.neuro.2008.12.009

Giordano, G. and Costa, L. G. (2012). Developmental neurotoxicity: some old and new issues. *ISRN Toxicol 814795*. http://dx.doi.org/10.5402/2012/814795

Goines, P. E. and Ashwood, P. (2013). Cytokine dysregulation in autism spectrum disorders (ASD): Possible role of the environment. *Neurotoxicol Teratol 36*, 67-81. http://dx.doi.org/10.1016/j.ntt.2012.07.006

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

Grandjean, P. (2007). Methylmercury toxicity and functional programming. *Reprod Toxicol 23*, 414-420. http://dx.doi.org/10.1016/j.reprotox.2007.03.002

Grandjean, P. and Landrigan, P. J. (2014). Neurobehavioural effects of developmental toxicity. *Lancet Neurol 13*, 330-338. http://dx.doi.org/10.1016/S1474-4422(13)70278-3

Gustafson, P. and McCandless, L. C. (2010). Probabilistic approaches to better quantifying the results of epidemiologic studies. *Int J Environ Res Public Health 7*, 1520-1539. http://dx.doi.org/10.3390/ijerph7041520

Guzzi, G. and La Porta, C. A. M. (2008). Molecular mechanisms triggered by mercury. *Toxicol 244*, 1-12. http://dx.doi.org/10.1016/j.tox.2007.11.002

Hammock, E., Veenstra-Vanderweele, J., Yan, Z., et al. (2012). Examining autism spectrum disorders by biomarkers: Example from the oxytocin and serotonin systems. *J Am Acad Child Adolesc Psych 51*, 712-721. http://dx.doi.org/10.1016/j.jaac.2012.04.010

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. http://dx.doi.org/10.1006/exnr.2000.7406

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

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

Hartung, T., Balls, M., Bardouille, C., et al. (2002). Report of ECVAM task force on good cell culture practice (GCCP). *Altern Lab Anim 30*, 407-414

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. (2007a). Food for thought on ... validation. *ALTEX 24*, 67-73. http://www.altex.ch/All-issues/Issue.50.html?iid=86&aid=3

Hartung, T. (2007b). Food for thought on ... cell culture. *ALTEX 24*, 3-9. http://www.altex.ch/All-issues/Issue.50.html?iid=87&aid=1

Hartung, T. (2008a). Food for thought on ... animal tests. *ALTEX 25*, 3-10. http://www.altex.ch/All-issues/Issue.50.html?iid=98&aid=1

Hartung, T. (2008b). Towards a new toxicology – evolution or revolution? *Altern Lab Anim 36*, 633-639.

Hartung, T. and Leist, M. (2008). Food for thought ... on the evolution of toxicology and the phasing out of animal testing. *ALTEX 25*, 91-102. http://www.altex.ch/All-issues/Issue.50.html?iid=99&aid=1

Hartung, T. (2009a). A toxicology for the 21st century: Mapping the road ahead. *Toxicol Sci 109*, 18-23. http://dx.doi.org/10.1093/toxsci/kfp059

Hartung, T. (2009b). Toxicology for the twenty-first century.

Nature 460, 208-212. http://dx.doi.org/10.1038/460208a

Hartung, T. (2010a). Lessons learned from alternative methods and their validation for a new toxicology in the 21st century. *J Toxicol Env Health 13*, 277-290. http://dx.doi.org/10.1080/10937404.2010.483945

Hartung, T. (2010b). Food for thought ... on alternative methods for chemical safety testing. *ALTEX 27*, 3-14. http://www.altex.ch/All-issues/Issue.50.html?iid=112&aid=1

Hartung, T. (2010c). Evidence based-toxicology – the toolbox of validation for the 21st century? *ALTEX 27*, 241-251. http://www.altex.ch/All-issues/Issue.50.html?iid=121&aid=3

Hartung, T. (2011). From alternative methods to a new toxicology. *Eur J Pharmaceutics Biopharmaceutics 77*, 338-349. http://dx.doi.org/10.1016/j.ejpb.2010.12.027

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

Hartung, T. and Zurlo, J. (2012). Alternative approaches for medical countermeasures to biological and chemical terrorism and warfare. *ALTEX 29*, 251-260. http://dx.doi.org/10.14573/altex.2012.3.251

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

Hartung, T. (2013). Look back in anger – what clinical studies tell us about preclinical work. *ALTEX 30*, 275-291. http://dx.doi.org/10.14573/altex.2013.3.275

Hartung, T., Luechtefeld, T., Maertens, A., and Kleensang, A. (2013a). Integrated testing strategies for safety assessments. *ALTEX 30*, 3-18. http://dx.doi.org/10.14573/altex.2013.1.003

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

Hermsen, S. A., Pronk, T. E., van den Brandhof, E. J., et al. (2013). Transcriptomic analysis in the developing zebrafish embryo after compound exposure: individual gene expression and pathway regulation. *Toxicol Appl Pharmacol 272*, 161-171. http://dx.doi.org/10.1016/j.taap.2013.05.037

Hirt, U. A., Gantner, F., and Leist, M. (2000). Phagocytosis of nonapoptotic cells dying by caspase-independent mechanisms. *J Immunol 164*, 6520-6529.

Hirt, U. A. and Leist, M. (2003). Rapid, noninflammatory and PS-dependent phagocytic clearance of necrotic cells. *Cell Death Differ 10*, 1156-1164. http://dx.doi.org/10.1038/sj.cdd.4401286

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. http://dx.doi.org/10.1007/s00204-012-0984-2

Hoffmann, S. and Hartung, T. (2005). Diagnosis: Toxic! – Trying to apply approaches of clinical diagnostics and prevalence in toxicology considerations. *Toxicol Sci 85*, 422-428. http://dx.doi.org/10.1093/toxsci/kfi099

Hoffmann, S. and Hartung, T. (2006). Towards an evidence-based toxicology. *Human Exp Toxicol 25*, 497-513. http://dx.doi.org/10.1191/0960327106het648oa

Hoffmann, S., Edler, L., Gardner, I., et al. (2008). Points of reference in validation – the report and recommendations of ECVAM Workshop. *Altern Lab Anim 36*, 343-352.

Hogberg, H. T., Kinsner-Ovaskainen, A., Hartung, T., et al. (2009). Gene expression as a sensitive endpoint to evaluate cell differentiation and maturation of the developing central nervous system in primary cultures of rat cerebellar granule cells (CGCs) exposed to pesticides. *Toxicol Appl Pharmacol 235*, 268-286. http://dx.doi.org/10.1016/j.taap.2008.12.014

Hogberg, H. T., Kinsner-Ovaskainen, A., Coecke, S., et al. (2010). mRNA expression is a relevant tool to identify developmental neurotoxicants using an in vitro approach. *Toxicol Sci 113*, 95-115. http://dx.doi.org/10.1093/toxsci/kfp175

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. *Neurotoxicol 32*, 158-168. http://dx.doi.org/10.1016/j.neuro.2010.10.007

Holson, R. R., Freshwater, L., Maurissen, J. P., et al. (2008). Statistical issues and techniques appropriate for developmental neurotoxicity testing: a report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. *Neurotoxicol Teratol 30*, 326-348. http://dx.doi.org/10.1016/j.ntt.2007.06.001

Honegger, P., Lenoir, D., and Favrod, P. (1979). Growth and differentiation of aggregating fetal brain cells in a serum-free defined medium. *Nature 282*, 305-308. http://dx.doi.org/10.1038/282305a0

Honegger, P. and Monnet-Tschudi, F. (2001). Aggregating neural cell cultures. In S. Fedoroff (ed.), *Protocols for neural cell cultures* (199-218). Totowa, N.J., USA: Humana Press. http://dx.doi.org/10.1385/1-59259-207-4:199

Howard, A. S., Bucelli, R., Jett, D. A., et al. (2005). Chlorpyrifos exerts opposing effects on axonal and dendritic growth in primary neuronal cultures. *Toxicol Appl Pharmacol 207*, 112-124. http://dx.doi.org/10.1016/j.taap.2004.12.008

Hu, V. W. (2013). From genes to environment: using integrative genomics to build a "systems-level" understanding of autism spectrum disorders. *Child Develop 84*, 89-103. http://dx.doi.org/10.1111/j.1467-8624.2012.01759.x

IACC – Interagency Autism Coordinating Committee (2011). IACC strategic plan for autism spectrum disorder research. Retrieved from the Department of Health and Human Services Interagency Autism Coordinating Committee website http://iacc.hhs.gov/strategic-plan/2012/index.shtml (accessed 21.08.2012).

Jager, T., Vermeire, T. G., Rikken, M. G., and van der Poel, P. (2001). Opportunities for a probabilistic risk assessment of chemicals in the European Union. *Chemosphere 43*, 257-264. http://dx.doi.org/10.1016/S0045-6535(00)00087-4

James, S. J., Cutler, P., Melnyk, S., and Jernigan, S. (2004). Metabolic biomarkers of increased oxidative stress and impaired methylation capacity in children with autism. *Am J Clin Nutr 80*, 1611-1617.

Johansson, C., Castoldi, A. F., Onishchenko, N., et al. (2007). Neurobehavioural and molecular changes induced by meth-

ylmercury exposure during development. *Neurotox Res 11*, 241-260.

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

Kaufmann, W. (2003). Current status of developmental neurotoxicity: An industry perspective. *Toxicol Lett 140-141*, 161-169. http://dx.doi.org/10.1016/S0378-4274(02)00503-9

Kern, I., Xu, R., Julien, S., et al. (2013). Embryonic stem cell-based screen for small molecules: Cluster analysis reveals four response patterns in developing neural cells. *Curr Med Chem 20*, 710-723. http://dx.doi.org/10.2174/092986713804999330

Khan, K. S., Kunz, R., Kleijnen, J., et al. (2003). Five steps to conducting a systematic review. *J R Soc Med 96*, 118-121. http://dx.doi.org/10.1258/jrsm.96.3.118

Kim, W. Y., Gonsiorek, E. A., Barnhart, C., et al. (2009). Statins decrease dendritic arborization in rat sympathetic neurons by blocking RhoA activation. *J Neurochem 108*, 1057-1071. http://dx.doi.org/10.1111/j.1471-4159.2008.05854.x

Kirchhoff, F., Dringen, R., and Giaume, C. (2001). Pathways of neuron-astrocyte interactions and their possible role in neuroprotection. *Eur Arch Psychiatry Clin Neurosci 251*, 159-169. http://dx.doi.org/10.1007/s004060170036

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

Kozomara, A. and Griffiths-Jones, S. (2011). miRBase: Integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res 39*, D152-157. http://dx.doi.org/10.1093/nar/gkq1027

Krause, K. H., van Thriel, C., De Sousa, P. A., et al. (2013). Monocrotophos in Gandaman village: India school lunch deaths and need for improved toxicity testing. *Arch Toxicol 87*, 1877-1881. http://dx.doi.org/10.1007/s00204-013-1113-6

Krewski, D., Acosta, D., Jr, Andersen, M., et al. (2010). Toxicity testing in the 21st century: A vision and a strategy. *J Toxicol Environ Health B Cit Rev 13*, 51-138.

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

Krug, A. K., Kolde, R., Gaspar, J. A., et al. (2013b). Human embryonic stem cell-derived test systems for developmental neurotoxicity: A transcriptomics approach. *Arch Toxicol 87*, 123-143. http://dx.doi.org/10.1007/s00204-012-0967-3

Krug, A. K., Gutbier, S., Zhao, L., et al. (2014). Transcriptional and metabolic adaptation of human neurons to the mitochondrial toxicant MPP+. *Cell Death Disease*, in press.

Kuegler, P. B., Zimmer, B., Waldmann, T., et al. (2010). Markers of murine embryonic and neural stem cells, neurons and astrocytes: Reference points for developmental neurotoxicity testing. *ALTEX 27*, 17-42. http://www.altex.ch/All-issues/Issue.50.html?iid=112&aid=2

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. http://dx.doi.org/10.1002/glia.21257

Kuehn, B. M. (2010). Increased risk of ADHD associated with early exposure to pesticides, PCBs. *JAMA 304*, 27-28. http://dx.doi.org/10.1001/jama.2010.860

Kuwagata, M. (2012). Current problems of in vivo developmental neurotoxicity tests and a new in vivo approach focusing on each step of the developing central nervous system. *Congenital Anomalies 52*, 129-139. http://dx.doi.org/10.1111/j.1741-4520.2012.00376.x

Landrigan, P. J., Kimmel, C. A., Correa, A., and Eskenazi, B. (2004). Children's health and the environment: Public health issues and challenges for risk assessment. *Environ Health Persp 112*, 257-265. http://dx.doi.org/10.1289/ehp.6115

Landrigan, P. J. (2010). What causes autism? Exploring the environmental contribution. *Curr Opin Pediat 22*, 219-225. http://dx.doi.org/10.1097/MOP.0b013e328336eb9a

Landrigan, P. J., Lambertini, L., and Birnbaum, L. S. (2012). A research strategy to discover the environmental causes of autism and neurodevelopmental disabilities. *Environ Health Persp 120*, 258-260. http://dx.doi.org/10.1289/ehp.1104285

Lee, C.-T., Risom, T., and Strauss, W. M. (2007). Evolutionary conservation of microRNA regulatory circuits: An examination of microRNA gene complexity and conserved microRNA-target interactions through metazoan phylogeny. *DNA Cell Biol 26*, 209-218. http://dx.doi.org/10.1089/dna.2006.0545

Lee, J., Lilly, G. D., Doty, R. C., et al. (2009). In vitro toxicity testing of nanoparticles in 3D cell culture. *Small 5*, 1213-1221.

Lee, R. C. and Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. *Science 294*, 862-864. http://dx.doi.org/10.1126/science.1065329

Lein, P., Locke, P., and Goldberg, A. (2007). Meeting report: Alternatives for developmental neurotoxicity testing. *Environ Health Persp 115*, 764-768. http://dx.doi.org/10.1289/ehp.9841

Leist, M., Bremer, S., Brundin, P., et al. (2008a). The biological and ethical basis of the use of human embryonic stem cells for in vitro test systems or cell therapy. *ALTEX 25*, 163-190. http://www.altex.ch/All-issues/Issue.50.html?iid=101&aid=3

Leist, M., Hartung, T., and Nicotera, P. (2008b). The dawning of a new age of toxicology. *ALTEX 25*, 103-114. http://www.altex.ch/All-issues/Issue.50.html?iid=99&aid=3

Leist, M., Efremova, L., and Karreman, C. (2010). Food for thought ... considerations and guidelines for basic test method descriptions in toxicology. *ALTEX 27*, 309-317. http://www.altex.ch/All-issues/Issue.50.html?iid=121&aid=7

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

Leist, M. and Hartung, T. (2013). Inflammatory findings on species extrapolations: Humans are definitely no 70-kg mice. *Arch Toxicol 87*, 563-567. http://dx.doi.org/10.1007/s00204-013-1038-0

Leist, M., Ringwald, A., Kolde, R., et al. (2013). Test systems of developmental toxicity: State-of-the art and future perspec-

tives. *Arch Toxicol 87*, 2037-2042. http://dx.doi.org/10.1007/s00204-013-1154-x

Li, X. and Jin, P. (2010). Roles of small regulatory RNAs in determining neuronal identity. *Nature 11*, 329-338.

Lidsky, T. I. and Schneider, J. S. (2003). Lead neurotoxicity in children: basic mechanisms and clinical correlates. *Brain 126*, 5-19. http://dx.doi.org/10.1093/brain/awg014

Liu, H. and Zhang, S. C. (2011). Specification of neuronal and glial subtypes from human pluripotent stem cells. *Cell Mol Life Sci 68*, 3995-4008. http://dx.doi.org/10.1007/s00018-011-0770-y

Makris, S. L., Raffaele, K., Allen, S., et al. (2009). A retrospective performance assessment of the developmental neurotoxicity study in support of OECD test guideline 426. *Environ Health Persp 117*, 17-25.

Marchetti, C. (2003). Molecular targets of lead in brain neurotoxicity. *Neurotox Res 5*, 221-236.

Masuo, Y. and Ishido, M. (2011). Neurotoxicity of endocrine disruptors: Possible involvement in brain development and neurodegeneration. *J Toxicol Environ Health B Crit Rev 14*, 346-369. http://dx.doi.org/10.1080/10937404.2011.578557

Ming, X., Stein, T. P., Brimacombe, M., et al. (2005). Increased excretion of a lipid peroxidation biomarker in autism. *Prostagl Leukotr Ess Fatty Acids 73*, 379-384. http://dx.doi.org/10.1016/j.plefa.2005.06.002

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. http://dx.doi.org/10.1016/S0006-8993(96)00895-5

Moser, V. C., Phillips, P. M., McDaniel, K. L., and MacPhail, R. C. (1999). Behavioral evaluation of the neurotoxicity produced by dichloroacetic acid in rats. *Neurotoxicol Teratol 21*, 719-731. http://dx.doi.org/10.1016/S0892-0362(99)00029-X

Nat, R. and Dechant, G. (2011). Milestones of directed differentiation of mouse and human embryonic stem cells into telencephalic neurons based on neural development in vivo. *Stem Cells Develop 20*, 947-958. http://dx.doi.org/10.1089/scd.2010.0417

Nava-Ruiz, C., Méndez-Armenta, M., and Ríos, C. (2012). Lead neurotoxicity: Effects on brain nitric oxide synthase. *J Mol Histol 43*, 553-563. http://dx.doi.org/10.1007/s10735-012-9414-2

Nielsen, B. S., Lam, H. R., and Ladefoged, O. (2003). Developmental neurotoxicity of toluene in rats as measured by L-ornithine decarboxylase in cerebellum. *Pharmacol Toxicol 92*, 51-54.

NRC – National Research Council (2000). National Research Council: *Scientific frontiers in developmental toxicology and risk assessment*. 1st edition. Washington DC, USA: National Academy Press.

NRC (2007). *Toxicity Testing in the 21st Century: a Vision and a Strategy*. 1st edition. Washington DC, USA: National Academy Press. http://www.nap.edu/catalog.php?record_id=11970

NRC (2009). *Science and Decisions: Advancing Risk Assessment*. Washington, DC, USA: National Academies Press.

Newschaffer, C. J., Croen, L. A., Fallin, M. D., et al. (2012). Infant siblings and the investigation of autism risk factors. *J Neurodevelop Disorders 4*, 7. http://dx.doi.org/10.1186/1866-1955-4-7

OECD (2007). Test Guideline 426. OECD Guideline for Testing of Chemicals. Developmental Neurotoxicity Study. http://www.oecd.org/document/55/0,3343,en_2649_34377_2349687_1_1_1_1,00.html

OECD (2011). Test Guideline 443. Extended One-Generation Reproductive Toxicity Study. http://www.oecd-ilibrary.org/environment/test-no-443-extended-one-generation-reproductive-toxicity-study_9789264185371-en

Onore, C., Careaga, M., and Ashwood, P. (2012). The role of immune dysfunction in the pathophysiology of autism. *Brain Behavior Immun 26*, 383-392. http://dx.doi.org/10.1016/j.bbi.2011.08.007

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. http://dx.doi.org/10.1007/s10565-013-9250-5

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

Pastural, E., Ritchie, S., Lu, Y., et al. (2009). Novel plasma phospholipid biomarkers of autism: Mitochondrial dysfunction as a putative causative mechanism. *Prostagl Leukotr Ess Fatty Acids 81*, 253-264. http://dx.doi.org/10.1016/j.plefa.2009.06.003

Penschuck, S., Flagstad, P., Didriksen, M., et al. (2006). Decrease in parvalbumin-expressing neurons in the hippocampus and increased phencyclidine-induced locomotor activity in the rat methylazoxymethanol (MAM) model of schizophrenia. *Eur J Neurosci 23*, 279-284. http://dx.doi.org/10.1111/j.1460-9568.2005.04536.x

Radio, N. M. and Mundy, W. R. (2008). Developmental neurotoxicity testing in vitro: Models for assessing chemical effects on neurite outgrowth. *Neurotoxicol 29*, 361-376. http://dx.doi.org/10.1016/j.neuro.2008.02.011

Radio, N. M., Breier, J. M., Shafer, T. J., and Mundy, W. R. (2008). Assessment of chemical effects on neurite outgrowth in PC12 cells using high content screening. *Toxicol Sci 105*, 106-118. http://dx.doi.org/10.1093/toxsci/kfn114

Raffaele, K. C., Fisher, J. E., Hancock, S., et al. (2008). Determining normal variability in a developmental neurotoxicity test. A report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. *Neurotoxicol Teratol 30*, 288-325. http://dx.doi.org/10.1016/j.ntt.2007.12.009

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. http://dx.doi.org/doi:10.1016/j.ntt.2010.04.053

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

Rana, T. M. (2007). Illuminating the silence: Understanding the structure and function of small RNAs. *Nature Rev Molecular Cell Biol 8*, 23-36. http://dx.doi.org/10.1038/nrm2085

Rappaport, S. M. (2011). Implications of the exposome for exposure science. *J Expos Sci Environ Epidemiol 21*, 5-9. http://dx.doi.org/10.1038/jes.2010.50

Ratajczak, H. V. (2011). Theoretical aspects of autism: biomarkers – a review. *J Immunotoxicol 8*, 80-94. http://dx.doi.org/10.3109/1547691X.2010.538749

Rice, D. and Barone, S., Jr (2000). Critical periods of vulnerability for the developing nervous system: Evidence from humans and animal models. *Environ Health Persp 108, Suppl 3*, 511-533.

Rice, D. C. (2005). Assessing the effects of environmental toxicant exposure in developmental epidemiological studies: Issues for risk assessment. *Neurotoxicol 26*, 483-489. http://dx.doi.org/10.1016/j.neuro.2004.12.009

Robinson, J. F., Verhoef, A., van Beelen, V. A., et al. (2012). Dose-response analysis of phthalate effects on gene expression in rat whole embryo culture. *Toxicol Appl Pharmacol 264*, 32-41. http://dx.doi.org/10.1016/j.taap.2012.07.015

Rodier, P. M. (1994). Vulnerable periods and processes during central nervous system development. *Environ Health Persp 102, Suppl 2*, 121-124.

Rodier, P. M. (1995). Developing brain as a target of toxicity. *Environ Health Persp 103, Suppl 6*, 73-76.

Rovida, C., Longo, F., and Rabbit, R. R. (2011). How are reproductive toxicity and developmental toxicity addressed in REACH dossiers? *ALTEX 28*, 273-294. http://dx.doi.org/10.14573/altex.2011.4.273

Rybak, A., Fuchs, H., Smirnova, L., et al. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nature Cell Biol 10*, 987-993. http://dx.doi.org/10.1038/ncb1759

Rybak, A., Fuchs, H., Hadian, K., et al. (2009). The let-7 target gene mouse lin-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. *Nature Cell Biol 11*, 1411-1420. http://dx.doi.org/10.1038/ncb1987

Sagiv, S. K., Thurston, S. W., Bellinger, D. C., et al. (2010). Prenatal organochlorine exposure and behaviors associated with attention deficit hyperactivity disorder in school-aged children. *Am J Epidemiol 171*, 593-601. http://dx.doi:10.1093/aje/kwp427

San Martin, N., Cervera, A. M., Cordova, C., et al. (2011). Mitochondria determine the differentiation potential of cardiac mesoangioblasts. *Stem Cells 29*, 1064-1074. http://dx.doi.org/10.1002/stem.654

Schettler, T. (2001). Toxic threats to neurologic development of children. *Environ Health Persp 109, Suppl 6*, 813-816.

Schildknecht, S., Karreman, C., Pöltl, D., et al. (2013). Generation of genetically-modified human differentiated cells for toxicological tests and the study of neurodegenerative diseases. *ALTEX 30*, 427-444. http://dx.doi.org/10.14573/altex.2013.4.427

Schmid, C. and Rotenberg, J. S. (2005). Neurodevelopmental toxicology. *Neurologic Clinics 23*, 321-336. http://dx.doi.org/10.1016/j.ncl.2004.12.010

Schug, T. T., Barouki, R., Gluckman, P. D., et al. (2013). PPTOXIII: Environmental stressors in the developmental origins of disease – evidence and mechanisms. *Toxicol Sci 131*, 343-350. http://dx.doi.org/10.1093/toxsci/kfs267

Schulpen, S. H., Robinson, J. F., Pennings, J. L., et al. (2013). Dose response analysis of monophthalates in the murine embryonic stem cell test assessed by cardiomyocyte differentiation and gene expression. *Reprod Toxicol 35*, 81-88. http://dx.doi.org/10.1016/j.reprotox.2012.07.002

Schulpen, S. H., Pennings, J. L., Tonk, E. C., and Piersma, A. H. (2014). A statistical approach towards the derivation of predictive gene sets for potency ranking of chemicals in the mouse embryonic stem cell test. *Toxicol Lett 225*, 342-349. http://dx.doi.org/10.1016/j.toxlet.2014.01.017

Selenica, M. L., Jensen, H. S., Larsen, A. K., et al. (2007). Efficacy of small-molecule glycogen synthase kinase-3 inhibitors in the postnatal rat model of tau hyperphosphorylation. *Br J Pharmacol 152*, 959-979. http://dx.doi.org/10.1038/sj.bjp.0707471

Senut, M.-C., Cingolani, P., Sen, A., et al. (2012). Epigenetics of early-life lead exposure and effects on brain development. *Epigenomics 4*, 665-674. http://dx.doi:10.2217/epi.12.58

Sherer, T. B., Richardson, J. R., Testa, C. M., et al. (2007). Mechanism of toxicity of pesticides acting at complex I: Relevance to environmental etiologies of Parkinson's disease. *J Neurochem 100*, 1469-1479.

Shi, Y., Zhao, X., Hsieh, J., et al. (2010). MicroRNA regulation of neural stem cells and neurogenesis. *J Neurosci 30*, 14931-14936. http://dx.doi.org/10.1523/JNEUROSCI.4280-10.2010

Sim, M. (2002). Case studies in the use of toxicological measures in epidemiological studies. *Toxicol 181-182*, 405-409. http://dx.doi.org/10.1016/S0300-483X(02)00478-X

Smirnova, L., Gräfe, A., Seiler, A., et al. (2005). Regulation of miRNA expression during neural cell specification. *Eur J Neurosci 21*, 1469-1477. http://dx.doi.org/10.1111/j.1460-9568.2005.03978.x

Smirnova, L., Sittka, A., and Luch, A. (2012). On the role of low-dose effects and epigenetics in toxicology. *Conservation Genetics 101*, 499-550.

Smirnova, L., Block, K., Hunecke, D., et al. (2014) miRNA profiling in neural differentiated mouse embryonic stem cells after exposure to developmental neurotoxicants. *PLOS One*, revised.

Stehr, C. M., Linbo, T. L., Incardona, J. P., and Scholz, N. L. (2006). The developmental neurotoxicity of fipronil: Notochord degeneration and locomotor defects in zebrafish embryos and larvae. *Toxicol Sci 92*, 270-278. http://dx.doi.org/10.1093/toxsci/kfj185

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. http://dx.doi.org/10.1093/toxsci/kfr034

Sullivan, M. G. (2005). Autism increase not a result of reclassification. *Clin Psychiat News*, p.68

Sun, T., Jackson, S., Haycock, J. W., et al. (2006). Culture of skin cells in 3D rather than 2D improves their ability to survive exposure to cytotoxic agents. *J Biotechnol 122*, 372-381. http://dx.doi.org/10.1016/j.jbiotec.2005.12.021

Tilson, H. A. (2000). Neurotoxicology risk assessment guidelines: Developmental neurotoxicology. *Neurotoxicol 21*, 189-194.

Trapp, B. D., Honegger, P., Richelson, E., and Webster, H. D. (1979). Morphological differentiation of mechanically dissociated fetal rat brain in aggregating cell cultures. *Brain Res 160*, 117-30. http://dx.doi.org/10.1016/0006-8993(79)90605-X

Tyl, R. W., Crofton, K., Moretto, A., et al. (2008). Identification and interpretation of developmental neurotoxicity effects: A report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. *Neurotoxicol Teratol 30*, 349-381. http://dx.doi.org/10.1016/j.ntt.2007.07.008

US EPA (1998). Health effects guidelines OPPTS 870.6300. http://www.epa.gov/nscep/index.html

Vahter, M. (2008). Health effects of early life exposure to arsenic. *Basic Clin Pharmacol Toxicol 102*, 204-211. http://dx.doi.org/10.1111/j.1742-7843.2007.00168.x

van Thriel, C., Westerink, R. H., Beste, C., et al. (2012). Translating neurobehavioural endpoints of developmental neurotoxicity tests into in vitro assays and readouts. *Neurotoxicol 33*, 911-924. http://dx.doi.org/10.1016/j.neuro.2011.10.002

van Vliet, E., Stoppini, L., Balestrino, M., et al. (2007). Electrophysiological recording of re-aggregating brain cell cultures on multi-electrode arrays to detect acute neurotoxic effects. *Neurotoxicol 28*, 1136-1146. http://dx.doi.org/10.1016/j.neuro.2007.06.004

van Vliet, E., Morath, S., Eskes, C., et al. (2008). A novel in vitro metabolomics approach for neurotoxicity testing, proof of principle for methyl mercury chloride and caffeine. *Neurotoxicol 29*, 1-12. http://dx.doi.org/10.1016/j.neuro.2007.09.007

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. http://dx.doi.org/10.1021/tx400402j

Walsh, P., Elsabbagh, M., Bolton, P., and Singh, I. (2011). In search of biomarkers for autism: Scientific, social and ethical challenges. *Nature Rev Neurosci 12*, 603-612. http://dx.doi.org/10.1038/nrn3113

Wang, X., Meng, D., Chang, Q., et al. (2010). Arsenic inhibits neurite outgrowth by inhibiting the LKB1-AMPK signaling pathway. *Environ Health Persp 118*, 627-634. http://dx.doi.org/10.1289/ehp.0901510

Weiss, B. (2011). Endocrine disruptors as a threat to neurological function. *J Neurol Sci 305*, 11-21. http://dx.doi.org/10.1016/j.jns.2011.03.014

Wild, C. P. (2011). Future research perspectives on environment and health: The requirement for a more expansive concept of translational cancer research. *Environ Health Persp 10*, S15-S15. http://dx.doi.org/10.1186/1476-069X-10-S1-S15

Win-Shwe, T.-T. and Fujimaki, H. (2010). Neurotoxicity of toluene. *Toxicol Lett 198*, 93-99. http://dx.doi.org/10.1016/j.toxlet.2010.06.022

Wu, X. and Song, Y. (2011). Preferential regulation of miRNA targets by environmental chemicals in the human genome. *BMC Genomics 12*, 244. http://dx.doi.org/10.1186/1471-2164-12-244

Wulczyn, F. G., Smirnova, L., Rybak, A., et al. 2007. Post-transcriptional regulation of the let-7 microRNA during neural cell specification. *FASEB J 21*, 415-426. http://dx.doi.org/10.1096/fj.06-6130com

Yan, Y., Sabharwal., P., Rao, M., and Sockanathan, S. (2009). The antioxidant enzyme Prdx1 controls neuronal differentiation by thiol-redox-dependent activation of GDE2. *Cell 138*, 1209-1221. http://dx.doi.org/10.1016/j.cell.2009.06.042

Yang, D., Lauridsen, H., Buels, K., et al. (2011). Chlorpyrifos-oxon disrupts zebrafish axonal growth and motor behavior. *Toxicol Sci 121*, 146-159. http://dx.doi.org/10.1093/toxsci/kfr028

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. http://dx.doi.org/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 121*, 357-367. http://dx.doi.org/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 Persp 120*, 1116-1122. http://dx.doi.org/10.1289/ehp.1104489

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

Zurich, M. G., Honegger, P., Schilter, B., et al. (2004). Involvement of glial cells in the neurotoxicity of parathion and chlorpyrifos. *Toxicol Appl Pharmacol 201*, 97-104. http://dx.doi.org/10.1016/j.taap.2004.05.003

Correspondence to
Thomas Hartung, MD PhD
Center for Alternatives to Animal Testing
Johns Hopkins Bloomberg School of Public Health
615 North Wolfe Street
Baltimore, MD 21205, USA
e-mail: thartung@jhsph.edu