Chemosphere 311 (2023) 137035

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Establishment of a human cell-based in vitro battery to assess developmental neurotoxicity hazard of chemicals

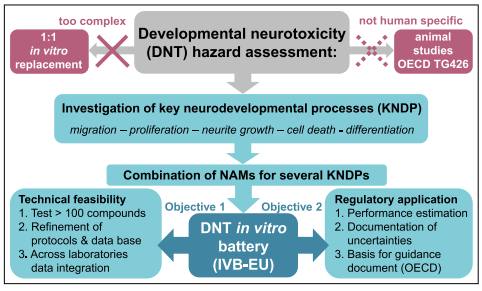
Jonathan Blum [a,1], Stefan Masjosthusmann [b,1], Kristina Bartmann [b], Farina Bendt [b], Xenia Dolde [a], Arif Dönmez [b], Nils Förster [d], Anna-Katharina Holzer [a], Ulrike Hübenthal [b], Hagen Eike Keßel [b], Sadiye Kilic [a], Jördis Klose [b], Melanie Pahl [b], Lynn-Christin Stürzl [b], Iris Mangas [e], Andrea Terron [e], Kevin M. Crofton [f], Martin Scholze [g], Axel Mosig [d], Marcel Leist [a,*], Ellen Fritsche [b,c,**]

[a] *In Vitro Toxicology and Biomedicine, Dept Inaugurated By the Doerenkamp-Zbinden Foundation, University of Konstanz, 78457, Konstanz, Germany*
[b] *IUF - Leibniz Research Institute for Environmental Medicine, 40225, Düsseldorf, Germany*
[c] *Medical Faculty, Heinrich-Heine-University, 40225, Düsseldorf, Germany*
[d] *Bioinformatics Group, Ruhr University Bochum, 44801, Bochum, Germany*
[e] *European Food Safety Authority, PREV Unit, 43126, Parma, Italy*
[f] *R3Fellows LLC, Durham, NC, USA*
[g] *Institute of Environment Health and Societies, Brunel University London, UK*

HIGHLIGHTS

- An in vitro testing battery (IVB) that allows screening of chemicals for developmental neurotoxicity (DNT) has been assembled.
- Performance estimates (>80% accuracy) have been obtained for the IVB, based on 45 negative/positive controls.
- Concentration-response data for altogether 120 compounds have been obtained for ten tests covering altogether 21 endpoints.
- Gaps of the IVB have been analyzed, and recommendations for the use of the IVB for regulatory testing have been put forward.

GRAPHICAL ABSTRACT



ARTICLE INFO

Handling Editor: Jian-Ying Hu

Keywords:
Testing battery
Stem cell
Brain development

ABSTRACT

Developmental neurotoxicity (DNT) is a major safety concern for all chemicals of the human exposome. However, DNT data from animal studies are available for only a small percentage of manufactured compounds. Test methods with a higher throughput than current regulatory guideline methods, and with improved human relevance are urgently needed. We therefore explored the feasibility of DNT hazard assessment based on new approach methods (NAMs). An in vitro battery (IVB) was assembled from ten individual NAMs that had been developed during the past years to probe effects of chemicals on various fundamental neurodevelopmental

* Corresponding author.
** Corresponding author. IUF - Leibniz Research Institute for Environmental Medicine, 40225, Düsseldorf, Germany.
E-mail addresses: jonathan.blum@uni-konstanz.de (J. Blum), marcel.leist@uni-konstanz.de (M. Leist), ellen.fritsche@iuf-duesseldorf.de (E. Fritsche).
[1] These authors contributed equally.

https://doi.org/10.1016/j.chemosphere.2022.137035
Received 8 July 2022; Received in revised form 20 October 2022; Accepted 24 October 2022
Available online 31 October 2022

In vitro testing
DNT

processes. All assays used human neural cells at different developmental stages. This allowed us to assess disturbances of: (i) proliferation of neural progenitor cells (NPC); (ii) migration of neural crest cells, radial glia cells, neurons and oligodendrocytes; (iii) differentiation of NPC into neurons and oligodendrocytes; and (iv) neurite outgrowth of peripheral and central neurons. In parallel, cytotoxicity measures were obtained. The feasibility of concentration-dependent screening and of a reliable biostatistical processing of the complex multi-dimensional data was explored with a set of 120 test compounds, containing subsets of pre-defined positive and negative DNT compounds. The battery provided alerts (hit or borderline) for 24 of 28 known toxicants (82% sensitivity), and for none of the 17 negative controls. Based on the results from this screen project, strategies were developed on how IVB data may be used in the context of risk assessment scenarios employing integrated approaches for testing and assessment (IATA).

1. Introduction

Screening of chemicals for a potential neurodevelopmental toxicity (DNT) hazard has been recognized as a pressing need by several large governmental and international organizations concerned with consumer safety. For instance, the US EPA and the European JRC took important roles in the organisation of a conference series (TestSmart) that was devoted to the development of a DNT test strategy useful in a regulatory context (Coecke et al., 2007; Lein et al., 2007; Crofton et al., 2011; Bal-Price et al., 2012). Also EFSA and the OECD embarked on similar efforts (Fritsche et al., 2017). In this context, several experimental programs were launched to probe novel approaches and to accelerate their implementation (Crofton et al., 2012; van Thriel et al., 2012; Krug et al., 2013b; Bal-Price et al., 2015; Baumann et al., 2016; Fritsche et al., 2018; Harrill et al., 2018; Behl et al., 2019; Lupu et al., 2020; Pistollato et al., 2021; Sachana et al., 2021; Vinken et al., 2021; Koch et al., 2022).

DNT is a field of toxicology concerned with effects of chemicals on the developing nervous system. Several experimental and epidemiological studies (on metals, pesticides and drugs) link compound exposure during early live phases (of the embryo, fetus or child) to functional alterations of the nervous system in adolescents or adults (Grandjean and Landrigan, 2014; Smirnova et al., 2014; Bennett et al., 2016). A particular concern is the possible role of DNT in the increased frequency of neurodevelopmental disorders, such as autism-spectrum disorders (Grandjean and Landrigan, 2006, 2014; Bellinger, 2012; Modafferi et al., 2021). The assessment is particularly challenging due to the multitude of potential toxicity manifestations (structural and functional). Moreover, there may be a time offset between toxicant exposure (before or after birth) and manifestation of effects (Grandjean et al., 2019).

The traditional methods to evaluate DNT hazard potential are based on animal studies following the OECD (OECD, 2007) or U.S. EPA (USEPA, 1998) test guidelines. To date only about 180 compounds world-wide have been tested using these guidelines (Crofton and Mundy, 2021). Several factors contribute to the limited availability of such studies: extensive time (e.g. 1–2 years) and resource requirement; limited triggered testing by chemical alerts; the need to reduce animal use; and the limited regulatory requirement for DNT testing as compared to some other test guidelines (e.g., carcinogenicity). The data available suffer from many uncertainties, and they require species extrapolation from rodents to humans. Moreover, they provide limited information on toxicity mechanisms. This can make them difficult to use in human risk assessments (Makris et al., 2009; Tsuji and Crofton, 2012; Tohyama, 2016; Paparella et al., 2020).

The strategic concepts of next generation risk assessment and of "toxicology for the 21st century" (Leist et al., 2008; Thomas et al., 2018; Pallocca et al., 2022a) suggest reductions in use of animal studies and development of new approach methods (NAMs) for toxicity assessment. The non-animal test methods should ideally be based on human-relevant test systems, reduce costs, allow a high throughput of test chemicals, and provide information on the toxicity mechanisms of toxicants. Many recent activities on scientific and regulatory levels have been undertaken to apply this strategy to the field of DNT (Sachana et al., 2019).

The establishment of DNT NAMs followed two major principles (Bal-Price et al., 2015; Aschner et al., 2017). First, a concept was developed on how complex in vivo events and their disturbances could be modeled by simplified in vitro systems. It was found that the biological process of nervous system development can be broken down to less complex key neurodevelopmental processes (KNDP). Moreover, it was assumed that the disturbance of any KNDP may lead to DNT in humans. On this basis, NAMs were developed for most of the crucial KNDP. The second principle was that the performance and robustness of the NAMs should be at a high level, so that data could be used with high confidence. The concept of test readiness was developed to provide a measure of the NAM validation status (Bal-Price et al., 2018; Krebs et al., 2019, 2020b), and several assays were deemed ready and suitable for use in chemical screening. They include: proliferation, migration and differentiation assays based on neurospheres (NPC1-5 test methods); the neurite growth assays NeuriTox and PeriTox; the neural crest migration assay (cMINC); and an assays for neural network formation and synaptogenesis (Masjosthusmann et al., 2020; Crofton and Mundy, 2021; Carstens et al., 2022). Instead of a formal OECD-type validation (e.g. skin sensitization NAMs (OECD, 2021; Strickland et al., 2022)), the concept of a fit-for-purpose biological validation based on regulatory needs has been suggested (Leist et al., 2012; Hartung et al., 2013; Judson et al., 2013; Cote et al., 2016; Griesinger et al., 2016; Bal-Price et al., 2018; Andersen et al., 2019; Masjosthusmann et al., 2020). Its application to DNT NAM involved: understanding of all technologies related to test systems and endpoint assessment; a comparison of pivotal in vitro signaling pathways to those relevant in vivo; and an assessment of the cellular presence of toxicity targets known to play a role for human DNT (Aschner et al., 2017; Bal-Price et al., 2018; Koch et al., 2022).

No individual NAM covers all key aspects of neurodevelopmental biology. Thus no single test will detect effects on all KNDP. Therefore, a battery of assays is needed, to sufficiently cover all DNT toxicants. In 2016, participants of a meeting jointly organized by the European Food Safety Autority (EFSA) and the organisation for Economic Co-operation and Development (OECD) agreed that "an in vitro testing battery (based on available DNT NAM) could be used immediately to screen and prioritize chemicals" (Fritsche et al., 2017). A test run for such a battery was planned, in order to evaluate the technical feasibility, to identify potential gaps and to provide data and experience for setting up a draft guidance on how to run battery testing, and how to interpret data therefrom (Crofton and Mundy, 2021). The purpose of this manuscript is to describe the first test run of a DNT in vitro test battery based on methods available in European laboratories (IVB-EU). Extensive raw data and method documentations can be found in a report by EFSA (Masjosthusmann et al., 2020), and the experience and learnings from the IVB-EU have led to the preparation of the draft of an OECD guidance document, which is currently (July 2022) under revision in member countries (Crofton and Mundy, 2021). However, the data from 10 assays on 120 compounds (including 28 positive and 17 negative controls) have not been made available to academia and the interested public in a peer-reviewed publication. The same applies to the preliminary performance evaluation of the IVB-EU as a whole and the considerations concerning further use. The purpose of this manuscript is to make this important information available, and to provide a basis for further developments in academia, industry and by regulatory institutions

concerned with NAM-based DNT testing.

2. Materials and methods

2.1. Chemicals

A list of screen compounds (n = 120) was assembled by a working group, using the member's experience as members/employees at the US EPA, EFSA or in OECD working groups. Compounds were selected to be chemically and biologically somewhat diverse and to reflect groups of compounds with concern for a potential DNT hazard. For instance, flame retardants and pesticides were included, as some compounds in these groups are known for biological properties of relevance to DNT. One aspect of the selection process was also to allow for diversity of effects on different fundamental neurodevelopmental processes (and respective assays), and it was important to cover the full spectrum from compounds with no or low evidence for DNT liability to compounds with rich background data to allow for a wide spread of screen results. A subset of compounds (n = 28) were included as positive controls for DNT hazard, based on human data or robust animal data (Grandjean and Landrigan, 2006, 2014; Mundy et al., 2015; Ryan et al., 2016; Aschner et al., 2017) (Fig. S1). Another subset (n = 17) were compounds considered as negative controls. They were selected for their safe use during human pregnancy or because the available extensive data on their toxicity gave no evidence (by observation or mechanism) of any effects related to DNT (at the test concentrations used) (Fig. S2). A description of chemicals, including exact chemical identity and suppliers is found in the suppl. file 2 - sheet 1.

2.2. Test methods

All test methods used for screening were selected based on their high readiness level (Bal-Price et al., 2018), as well as a very comprehensive test description compatible with the OECD Guidance Document GD211 for in vitro test method descriptions. These ToxTemp files (Krebs et al., 2019) are included in suppl. file 1. Below, only brief descriptions are given for a quick overview. Notably, most assays had at least two endpoints, and some assays were run in more than one version, e.g. measurement after 72 and 120 h.

UKN2 Assay (cMINC): The assay, is based on neural crest cells differentiated from hiPSC (Nyffeler et al., 2017). Cells were seeded into 96-well plates around a stopper. The stopper was removed after 24 h to allow migration into the cell free area. Cells were exposed to the test compound for 24 h, and then stained with calcein-AM and Hoechst H-33342. The number of migrated double positive cells was quantified independent of an observer by high content imaging and image analysis (RingAssay software; http://invitro-tox.uni-konstanz.de). The cell viability was also determined by an automated imaging algorithm. Concentration-response curves from this test were based on six test compound concentrations (plus solvent control).

UKN4 assay (NeuriTox): The assay is based on LUHMES cells that were cultured and handled as previously described (Lotharius et al., 2005; Scholz et al., 2011; Krug et al., 2013a). It assesses neurite outgrowth in central nervous system neurons (Delp et al., 2018). Cells were pre-differentiated for two days to commit them towards the neuronal fate. They were then re-seeded in 96-well plates and exposed to the chemical for 24 h. Viability and neurite area were determined by high-content imaging after staining with calcein-AM and H-33342. The neurite area was defined by a fully automated algorithm as the area of calcein-positive pixels minus the area of all cell soma (Stiegler et al., 2011). Concentration-response curves from this test were based on ten test compound concentrations (plus solvent control).

UKN5 Assay (PeriTox): The assay is based on immature sensory neurons differentiated from hiPSC as previously described (Hoelting et al., 2016; Holzer et al., 2022). The test measures neurite outgrowth in peripheral neurons. Frozen lots of peripheral neuron precursors were thawed and seeded into 96-well plates. After 1 h, the cells were exposed to test chemicals for 24 h. Testing and endpoint measurements were exactly as for the UKN4 assay (despite 6 instead of 10 compound concentrations tested).

NPC1-5 Assays: The neurosphere assays (NPC1-5) are based on primary human neural progenitor cells (hNPCs; gestational week 16–19), that are grown as floating 3D neurospheres. Their growth and viability is assessed in the 3D neurospheres (NPC1). Alternatively, spheres can be plated onto a laminin-coated matrix, where the cells start migration and differentiation to form a secondary 3D co-culture. The latter approach allows the simultaneous assessment of radial glia migration (NPC2a), neuronal differentiation (NPC3), neuronal migration (NPC2b) and neurite outgrowth (NPC4) as well as oligodendrocyte differentiation (NPC5) and their migration (NPC2c) by fully automated high content imaging. Data were obtained and analyzed from recorded microscope images by a dedicated image processing software, trained on positive and negative control images, as described earlier in detail (Forster et al., 2022; Koch et al., 2022).

For the NPC1 assay, spheres (0.3 mm) were plated in 96-well plates (U-bottom; 1 sphere/well) and directly exposed to the test compound (in proliferation medium). DNA synthesis was assessed as functional endpoint after 3 days in vitro (DIV), using a luminescence-based bromodeoxyuridine (BrdU) ELISA (Nimtz et al., 2019). Cytotoxicity was assessed as a membrane integrity assay (CytoTox-ONE Assay) measuring the LDH release into the supernatant.

For the NPC2-5 assays, spheres (0.3 mm) were plated in poly-D-lysine/laminin-coated 96-well plates (F-bottom; 1 sphere/well) and directly exposed to the test compounds (in differentiation medium). Under control conditions, NPCs migrate radially out of the attached sphere and differentiate into radial glia, neurons and oligodendrocytes. Data were obtained after 72 h and 120 h. After 72 h (3 DIV), bright field images were taken of live cell cultures, and radial glia migration (NPC2a [72 h]) was assessed using ImageJ software. The medium was partially removed (50%) and used to assess cytotoxicity (CytoTox-ONE Assay). To continue the assay, the medium was replenished and cells were allowed to further differentiate and migrate for 48 h. At 5 DIV, cells were fixated and stained for TUBB3 (neuronal marker), O4 (oligodendrocyte marker) and Hoechst H-33258 (nuclear marker). The endpoint assessment was done by high content imaging followed by different image analysis algorithms. Neuronal and oligodendrocyte differentiation (NPC3 and NPC5) was assessed as the number of all TUBB3-positive and O4-positive cells in percent of the total number of nuclei in the migration area. Neurons and oligodendrocytes were automatically recognized by a machine learning software based on convolutional neural networks (Forster et al., 2022). The high-content image analysis software Omnishpero was used to determine radial glia migration (NPC2a [120 h]), neuronal migration (NPC2b) and oligodendrocyte migration (NPC2c) as well as neuronal morphology (NPC4a: neurite length; NPC4b: neurite area) (Schmuck et al., 2017). Cytotoxicity was assessed from samples of medium removed before the fixation by the CytoTox-ONE LDH Assay. Some additional cell viability data were obtained by using a resazurin reduction assay (CellTiter-Blue Assay). Concentration-response curves from all these tests were based on seven test compound concentrations.

2.3. Screen strategy

Most of the compounds (n = 75) were provided by EPA's ToxCast chemical contractor (Evotec, South San Francisco, CA) in v-bottom 96 well plates. Separate plates were provided for different assays, and volumes shipped ranged from 50 to 300 μl as DMSO stock solutions (always 20 mM). Other compounds were obtained from commercial sources (indicated in the suppl. 2 Excel file). In some of these cases stock solution was higher than 20 mM and compounds were dissolved in water if they were highly water-soluble (e.g. valproic acid). The University of Konstanz robotics platform was used to either produce replicates of the

master plate for different screening runs and different assays (UKN assays) or to directly prepare the compound dilutions (1:3 steps) in the media in 96-well pates (NPC assays). Operators were blinded to the compound identity. For the UKN assays serial dilutions (1:3 steps) were prepared from the cloned master plates for each compound in DMSO on 96-well plates, and each of these stocks was transferred to a pre-dilution plate. On these plates compounds were diluted 1:3 in medium plus 1% DMSO to have constant levels of DMSO among all concentrations. Finally, pre-dilutions were transferred to assay plates with cells (e.g. 20 µl transfer to 180 µl cells corresponding to 1:10) in medium to a maximum DMSO level of 0.1% in each assay. Exact volumes and pre-dilutions were assay-dependent and are detailed in ToxTemps; suppl. file 1. Some compounds were tested in an adapted concentration range (e.g. it is known that valproic acid is a human teratogen and DNT toxicant at clinically used concentrations of 0.5–1 mM. Therefore, higher concentrations were also tested, and master stocks were prepared accordingly).

For some assays (e.g. UKN2), a pre-screening step was included, in which only 1–2 (highest) test compound concentrations were run. When they showed no effect, screening was ended. When there was an effect (at least 20% change of endpoint), a full concentration-response was obtained. Pre-screen and full concentration-response screen were performed three times independently for all assays. For the UKN assays this meant the use of different cell lots for each run, for the NPC assays it meant the use of cells from different donors and/or passages for each run. Each screen run contained 2–6 technical replicates (details in ToxTemps; suppl. file 1). In some cases, follow-up tests were run, when e.g. only the highest concentration showed a response. Then new stocks were produced, and the concentration range was extended to 60 or 100 µM, depending on the solubility of the compound.

2.4. Data analysis

A fully automated data analysis workflow was implemented on the programming platform R (Keßel, 2022). Original code and source files are available on GitHub at (https://github.com/iuf-duesseldorf/fritsche-lab-CRStats). It included the following steps and outputs: (1) Pre-processing of data, where required by the definitions of the assay endpoints (see ToxTemps; suppl. file 1). For instance, the background signal was subtracted from all data points for the BrdU fluorescence readings. (2) Normalization of test compound data to the median of solvent controls. (3) Calculation of the median of the replicates for each experimental condition. (4) Concentration response fitting of the data for each compound. The best-fitting model (general logistic, 3-parameter log-logistic, 4-parameter log-logistic, 2-parameter exponential, 3-parameter exponential, 3-parameter Weibull, 4-parameter Weibull) was selected by the AKAIKE information criteria (Ritz et al., 2015; Jensen et al., 2020). (5) Re-normalization of the data, so that the upper asymptote of the selected curve fit was at 100% (Krebs et al., 2018; Kappenberg et al., 2020). (6) Calculation of the mean re-normalized values for each condition across independent test runs. (7) Concentration response fitting of the data for each compound. The best-fitting model (general logistic, 3-parameter log-logistic, 4-parameter log-logistic, 2-parameter exponential, 3-parameter exponential, 3-parameter Weibull, 4-parameter Weibull) was selected by the AKAIKE information criteria. (8) Determination of the benchmark concentration (BMC) as the point of the concentration-response curve that intersected with the benchmark response level (BMR). The BMR was determined and described for each assay (see ToxTemp; suppl. file 1), based on a biological and statistical rationale. It marked the extent of response considered to be statistically significant and toxicologically meaningful. It thus depended on the endpoint and on the base line noise. For most functional endpoints it was set at 75% (= 25% reduced normal function). For some assays it was set at 70% (higher baseline noise). For some viability measures it was set at 90% (a deviation of >10% was considered to potentially influence the functional endpoint). (9) After determination of the BMC, the upper (BMCU) and lower limit (BMCL) of its 95% confidence interval were calculated (Krebs et al., 2020a).

2.5. Hit definitions and prediction models

The prediction models (Worth and Balls, 2001; Leist et al., 2010; Griesinger et al., 2016; Schmidt et al., 2017; Bal-Price et al., 2018; Krebs et al., 2020b) of the NAM used in the IVB-EU had been defined during the original test setup, as documented in the literature and the ToxTemp files. A key feature of all assays was that they had a specific functional endpoint (related to a KNDP) and an endpoint characterizing compound effects on cell viability. Within each NAM, a compound was considered a specific hit (toxicant), when it affected the functional endpoint at least at one concentration that did not affect viability (Fig. S3). Notably, this does not mean that specific cytotoxicity of a given cell population (e.g. neural crest cells) would not lead to DNT. However, specific toxicity to a subpopulation can only be determined across assays, not within one assay. At present, a procedure for such a cross-IVB interpretation has not been established. Within a given assay, cytotoxicity makes the interpretation of the functional endpoint difficult. Therefore, (i) functional endpoint data were only used for concentrations that were non-cytotoxic, and (ii) specific cytotoxicity to subpopulations was not considered in this first application of the IVB-EU. For the UKN assays, specific effects were determined by the ratio of benchmark concentrations for the functional endpoint (e.g. neurite growth in UKN4) and cytotoxicity (e.g. a 4-fold offset for UKN4). For the NPC assays, specific toxicity was assumed when the 95% confidence intervals of the functional endpoint and the viability endpoint did not overlap. As the separation between “hit” and “non-hit” leads to binary data with high uncertainties at the hit/non-hit boundary (Leontaridou et al., 2017; Delp et al., 2018), we introduced a borderline category for transition compounds (e.g. when confidence intervals in NPC assays overlapped by > 10%). Thus, a given compound was classified in each assay as “no hit”, “unspecific hit”, “specific hit” or “borderline hit” (Fig. S3).

2.6. Performance parameters

A set of 45 reference compounds (28 DNT positives; 17 DNT negatives) was used for a preliminary evaluation of the IVB-EU predictivity (more may be added in the future). Various hit definitions were used (e.g. only specific hits, or specific + borderline hits). If a positive control was a hit, it was considered true positive (TP), if it was not a hit, it was considered a false negative (FN). If a negative control was a hit, it was considered a false positive (FP) and if it was not a hit, it was considered a true negative (TN). Using these four numbers (FP, FN, TP, TN), the following performance parameters were defined:

$$sensitivity\ [\%] = \frac{TP}{(TP + FN)} * 100$$

$$specificity\ [\%] = \frac{TN}{(TN + FP)} * 100$$

$$accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} * 100$$

$$balanced\ accuracy = \frac{sensitivity + specificity}{2}$$

$$positive\ predictive\ value\ (PPV) = \frac{TP}{(TP + FP)} * 100$$

$$F1\ score = \frac{2}{\frac{1}{sensitivity} + \frac{1}{PPV}} = \frac{1}{2} * (sensitivity + PPV)$$

$$Matthews\ correlation\ coefficient\ (MCC) = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

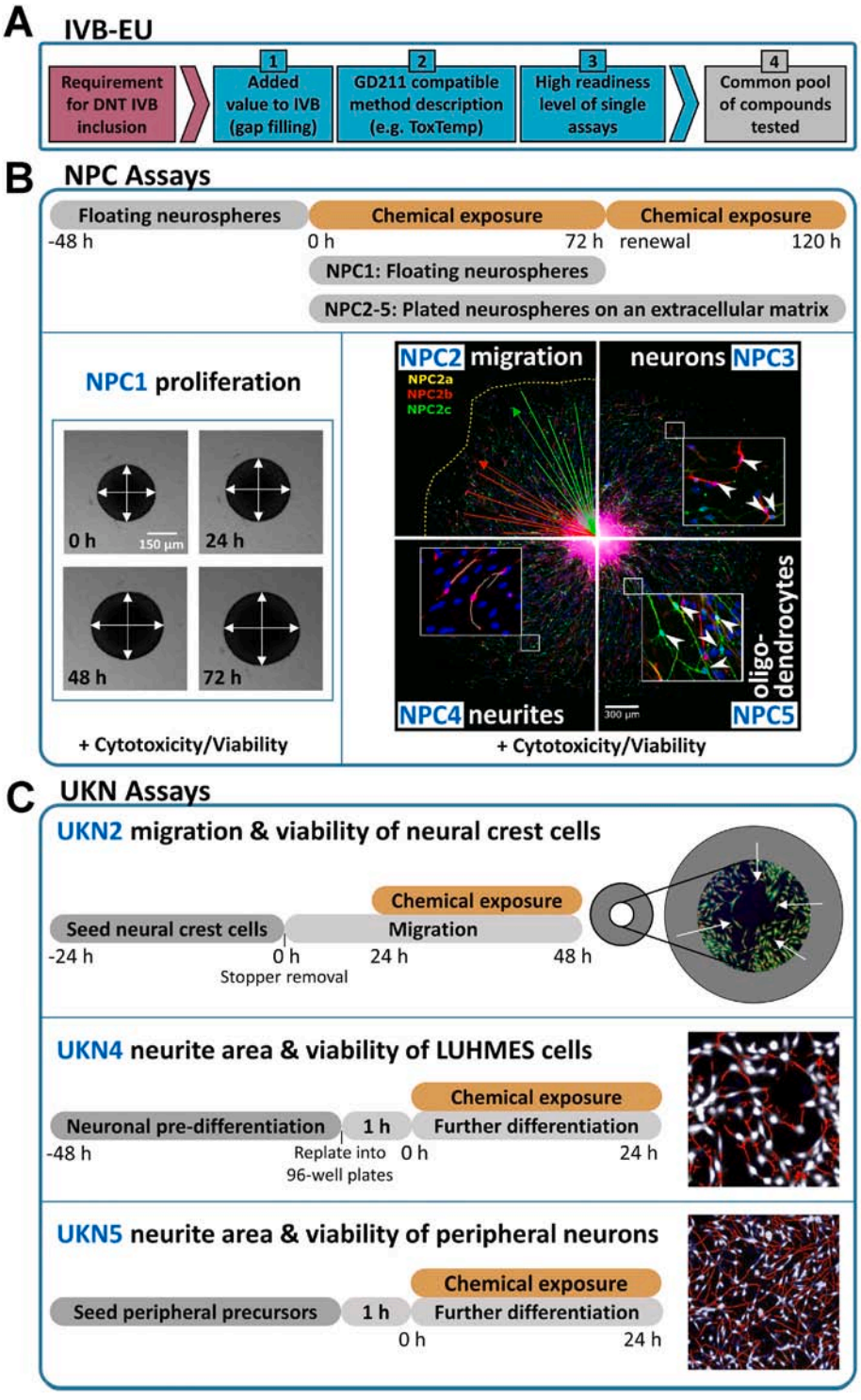


Fig. 1. Requirements and composition of the IVB-EU. (A) Criteria for assays to be included in the DNT test battery designated here IVB-EU. Criteria 1–3 were applied to this study. Criterion 4 was fulfilled in the course of this study and is suggested to be considered for future battery expansion. GD211 = OECD guidance document 211 on documentation of in vitro methods. (B) Schematic representation of the assays based on human neural progenitor cells (NPC) and their progeny. The general test system generation and exposure scheme is indicated on top. For the NPC1 test, floating neurospheres were exposed to toxicants for 72 h, and bromodeoxyuridine (BrdU) incorporation was used as endpoint for proliferation of NPC. For the NPC2-5 assays, neurospheres were plated and allowed to form secondary co-cultures of various cell types. Endpoints related to migration (NPC2), neuronal differentiation (NPC3), neurite growth (NPC4) and oligodendrocyte formation (NPC5) were assessed after 120 h by immunostaining and high content imaging. (C) Schematic representation of UKN assays. Cell types used and exposure schemes are indicated. Viability and migration of the cells in all assays were determined simultaneously by automated high content imaging after staining of the cell cultures with calcein-AM and Hoechst H-33342. The UKN2 assay evaluated the migration of neural crest cells into an empty circular area. The UKN4/UKN5 assays evaluated neural outgrowth of central nervous system and peripheral nervous system immature neurons. Detailed descriptions of NPC and UKN assays are given in the ToxTemps.

2.7. Data accessibility

The full raw data set from the IVB-EU has been entered into the ToxCast data base and is available in a machine-readable format used by many computational toxicologists after the fall 2022 ToxCast release (US EPA ORD, 2022).

3. Results and discussion

3.1. The DNT in vitro battery (IVB)

A large panel of assays with direct or indirect relevance to DNT can be found in the literature. Criteria needed to be developed to select a prototype battery of assays that was large enough for the main objective of this study, i.e. providing a basis for preparation of a general technical guidance document on battery testing for regulatory applications. At the same time, reasons of feasibility and limited resources called for keeping the number of NAMs included in the test run low. Experts with a regulatory background (from the US and Europe) were involved in the selection. The overall plan was to start testing in some European laboratories on a core battery (IVB-EU) of fully ready NAMs, and then to combine data on the same set of compounds with tests established at the US EPA. The three main selection criteria for the DNT NAMs were: (i) complementarity, (ii) documentation, and (iii) the readiness level (Fig. 1A). The first point meant that the assays were selected in a way to fill gaps of knowledge and to cover many KNDPs. It was also considered here to use assays for overlapping biological functions to learn about their orthogonality for later designs of tiered testing and sub-batteries. The second point referred to the availability of method documentations useful at a regulatory level (i.e. defined by OECD guidance document GD211) for the use of NAMs. Linked to this was the third criterion which referred to the technical performance of the NAMs, and the level of confidence into their predictivity and relevance. These issues are in some legislations referred to as validation state (Leist et al., 2012; Hartung et al., 2013; Judson et al., 2013; Cote et al., 2016; Griesinger et al., 2016; Bal-Price et al., 2018; Andersen et al., 2019; Masjosthusmann et al., 2020). In the selection of assays for the IVB-EU, we used a more flexible definition, termed "readiness" (Krebs et al., 2020b; Patterson et al., 2021). The assays used here all had undergone such an evaluation (Bal-Price et al., 2018; Klose et al., 2021a; Koch et al., 2022). An additional criterion important for development of additional assays, now recommended in the draft OECD DNT-IVB test guideline is use of a common pool of test compounds (Fig. 1A). Ten assays fulfilled all criteria, and they were considered to be suitable for forming the IVB-EU. In addition to the above points, all selected assays use human cells, cover four major KNDP, reflect seven different brain cell types and represent different neurodevelopmental stages (Fig. 1B and C; Fig. 2).

To obtain an overview of test battery relevance and predictivity, a gap analysis was performed. Comparison of the included tests with the known neurodevelopmental processes showed that some KNDP are currently not covered by the IVB-EU. These include very early developmental processes such as stem cell differentiation into neural progenitor cells and subsequent neural tube construction, as well as processes necessary for neuronal circuit building, like formation, maturation and function of neuronal networks. As such gaps may reduce the sensitivity of DNT predictions, we explored the availability of assays that fulfill the IVB-EU inclusion criteria and could become part of an expanded full battery (Fig. 2). Many assays for network formation have indeed already shown to be at high readiness, yet these are based on rat cortical cells (Carstens et al., 2022) calling for human cell-based neuronal network formation assays. The early embryonal stages of neural development may be covered by the UKN1 assay (Dreser et al., 2020; Meisig et al., 2020). Some functional endpoints related to non-neuronal cells are also desirable for the IVB, as these cells (astrocytes, microglia, myelinating oligodendrocytes, microvascular endothelial cells) do not only have support and immune function, but rather participate in multiple neurodevelopmental processes (Allen and Lyons, 2018). Several 3D systems have been described to include the necessary cell types (Brull et al., 2020; Chesnut et al., 2021; Nunes et al., 2022), but still need some development to meet basic inclusion criteria (set up of test methods, throughput, documentation) for the IVB. The same applies to dedicated assays to investigate neurotransmitter systems (e.g. glutamate and acetylcholine signaling) (Klima et al., 2021; Loser et al., 2021b). However, a large part of signaling systems is covered already by the recent development of neural network formation assays (Frank et al., 2017; Nimtz et al., 2020). An interesting endpoint to comprehensively capture neuronal differentiation is transcriptome profiling (Pallocca et al., 2016; Shinde et al., 2017; Simon et al., 2019; Dreser et al., 2020; Meisig et al., 2020; Hu et al., 2022). This was exemplified here by the UKN1 assay. Modern high throughput sequencing techniques (Simon

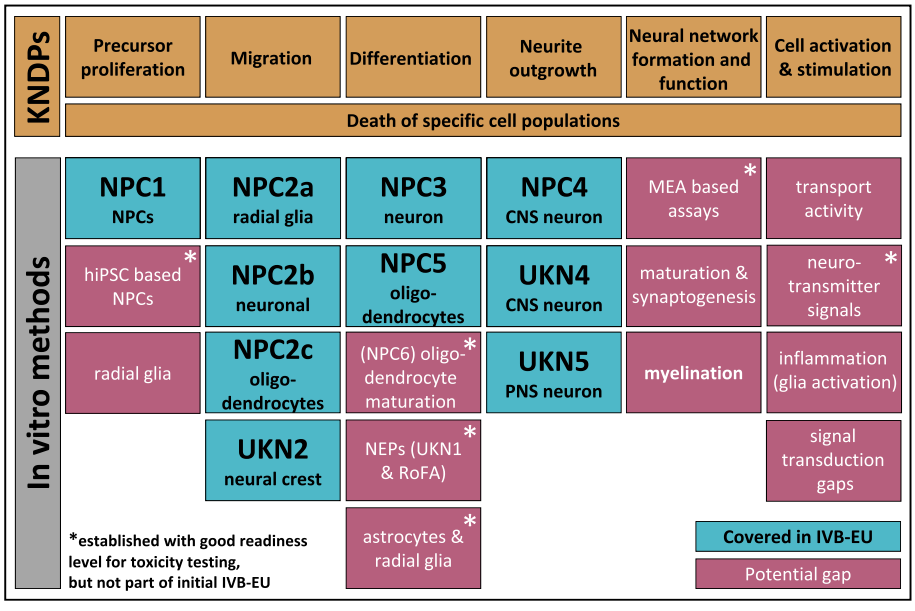


Fig. 2. Key neurodevelopmental processes (KNDP) covered by IVB-EU. Categories of KNDPs, according to Bal-Price et al., 2018 are listed on top. Specific cell death in a neurodevelopmental sub-population may either be considered a KNDP or an adverse effect. As it is measured as endpoint in all assays of other KNDP, it was considered to be broadly covered by the IVB-EU without a dedicated own assay. The lower part of the figure indicates NAM (designated here: in vitro methods) that are related to the respective KNDP on top of each column. The coverage of KNDPs by assays that are part of the current IVB-EU is shown (bold). For some KNDPs, more than one test was available. The reason was that several distinct subpopulations e. g. migrate (radial glia, neurons, oligodendrocytes and neural crest cells) or grow neurites (different types of CNS and PNS neurons). Potential gaps of the current IVB-EU are shown as assays in the non-bold in vitro method boxes. Assays that have already been established in the co-authors' labs are indicated by asterisks. They may be included in an extended version of the IVB, once they fulfill all inclusion criteria (Fig. 1). CNS: central nervous system; hiPSC: human induced pluripotent stem cells; NEP: neuroepithelial precursor; NPC: neural progenitor cell; MEA: microelectrode array; PNS: peripheral nervous system; RoFA: rosette formation assay.

et al., 2019; Jaklin et al., 2022; Spreng et al., 2022) now allow sufficient throughput for screening applications and it is likely that such assays will add additional information to the IVB in the future.

3.2. Readiness overview

The readiness of the assays of the DNT IVB was assessed on two tiers: first, the readiness of individual assays, as assessed earlier in individual publications, was an inclusion criterion (Fig. 1) of the IVB-EU. Second, the readiness of the overall battery and the performance of the assays under screening conditions was evaluated.

Concerning the first point, the underlying considerations are briefly re-iterated here, as they impinge on the interpretation and on the overall confidence into data from the NAMs of the IVB-EU. As for all toxicological assays, relevance, predictivity and reliability/robustness were considered. A major focus was put on the latter point, as suggested earlier (Leist et al., 2014; Krebs et al., 2019; Pallocca et al., 2022b). Earlier publications (summarized in Masjosthusmann et al. (2020)), and the ToxTemp (suppl. file 1) give more background information. One aspect helping to keep typical sources of variability low is that the selected IVB-EU assays all used a fully automated data capturing and evaluation procedure. However, the ultimate proof of the pudding for robustness, a blinded inter-lab comparison study, still has to be done for the assays.

When simple methods for 1:1 replacement of acute toxicity endpoints were evaluated, relevance and predictivity have been defined as separate aspects of NAMs. However, this concept has been modified for complex endpoints and batteries. In such more complex cases, the predictivity of a single NAM (for a given regulatory endpoint derived from animal studies) cannot be calculated, and the aspects of predictivity and relevance are strongly intertwined (Escher et al., 2022). In such cases, a scientific validation process is suggested that builds on two pillars: (i) comparison of the biological basis of the test system to that of the modeled human biology, and (ii) comparison of pathway modulations that lead to endpoint changes in the NAM to pathway changes known to be relevant to the respective human pathophysiology (Hartung, 2007; Leist et al., 2012; Hartung et al., 2013; Bal-Price et al., 2018; Piersma et al., 2018; Patterson et al., 2021). For the NAMs included in the IVB-EU, the test systems have been extensively documented and compared to the respective human developing nervous system counterparts. This involved the levels of cell morphology, cell function, and cell markers (see ToxTemps; suppl. file 1). Moreover, the relevant systems were profiled for their respective transcriptomes (Krug et al., 2014; Hoelting et al., 2016; Pallocca et al., 2017; Gutbier et al., 2018; Masjosthusmann et al., 2018; Klose et al., 2021a, 2021b, 2022). Also, the responses of the NAMs to modulation of signaling pathways relevant for brain development have been investigated by the use of compounds known to specifically affect signaling pathways (for overview: Klose et al. (2021b); Koch et al. (2022); Krebs et al. (2020b); Masjosthusmann et al. (2020)). A high-level summary of the responses to such "mechanistic tool compounds" is summarized in Fig. S4. One example is the Notch pathway, which determines a crucial switch between neurogenesis and oligodendrogenesis in vivo. By using the Notch pathway inhibitor DAPT, we can mimic this differentiation switch also in vivo with the NPC3/5 tests (Koch et al., 2022). Another illustrative example is the Rho pathway, which is involved in neurite growth in vivo. Activation of the RhoA kinase by narciclasine decreases neurite outgrowth in the NPC4, UKN4 and UKN5 assays. This successful characterization of neurodevelopmentally-relevant signaling in the IVB-EU assays is considered as the physiological basis and qualitative evidence for relevance and predictivity.

While the above-mentioned steps were important for the selection of NAMs and for giving confidence into their individual function within the IVB-EU, we also engaged in an effort to obtain information on the validity of the entire IVB-EU, as a battery. We considered the key parameteres robustness, predictivity and relevance (Hartung et al., 2004; Pallocca and Leist, 2022). Concerning relevance, it was mainly considered how many cell types and how many signaling pathways important for brain development were covered. A gap analysis showed that there was a need for few additional cells (e.g. microglia) and for some additional functions (e.g. neuronal network formation, astrocyte function). Moreover, more coverage of signaling (e.g. BDNF pathway and nicotinic signaling pathway) would be desirable. However, most relevant cell types were already represented, and many pathways known to be affected by toxicants were shown to be identifiable by at least one assay

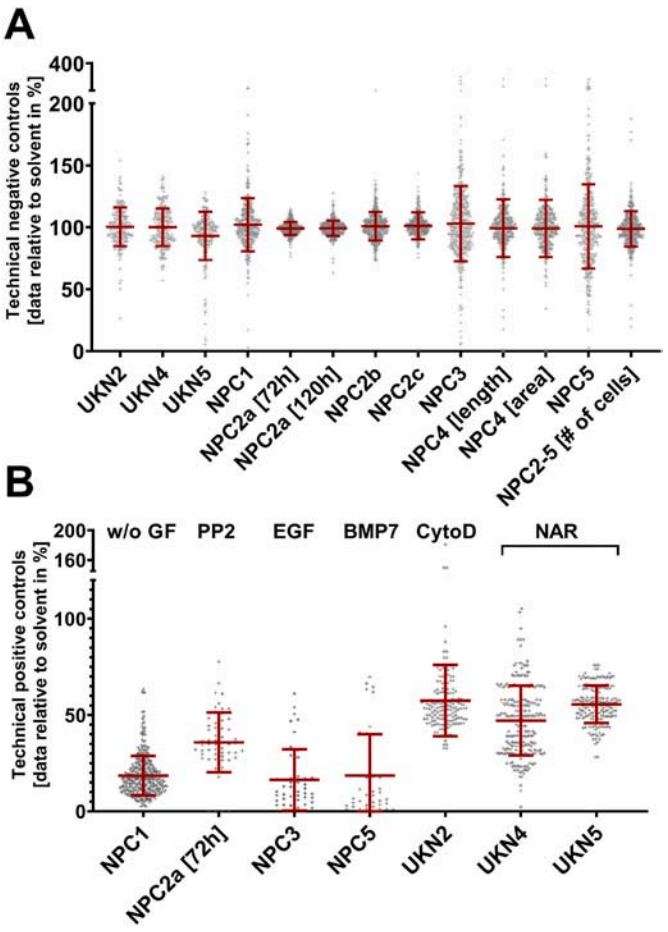


Fig. 3. Baseline noise and signal variation of acceptance controls in the IVB-EU assays. All tests were performed in a way so that each assay plate or experimental run contained wells with (i) negative controls, and at least one (ii) positive control. The reading of (ii) vs. (i) was used as acceptance criterion of the respective plate for UKN2, 4 and 5. If the positive control was not in a prespecified range, the plate data were not included in screen results and measurements were repeated. Depending on the assay, plates contained different numbers of compounds. For some tests, the different concentrations of a given compound were on different plates. Thus, some plates contained the (iii) lowest concentration of a compound, and some did not. (A) To obtain a measure of inter-plate and intra-experimental variability of the baseline signal, the lowest concentration of each test compound (iii) was compared to the solvent control (i) on each plate. Altogether >200 data points were obtained for each IVB-EU endpoint from the testing campaign. For easier overview, the means ± SD are indicated on top of the data points. (B) For each plate, the reading of the positive controls (ii) was compared to that of the negative controls (i) and normalized to negative control readings. The means ± SD of data for positive controls are given for the IVB-EU endpoints. The compounds used to set acceptance criteria were as follows: w/o GF: without growth factor (omission of normally present growth factors in the positive control well); PP-2: SRC-kinase inhibitor; EGF: epidermal growth factor; BMP7: bone morphogenetic protein 7; CytoD: cytochalasin D; NAR: narciclasine. Details on concentrations are found in the ToxTemps (suppl. file 1).

(Fig. 2; Fig. S4).

One estimate for the robustness of screening results from the test battery is the baseline noise level of the NAM. As the results of all assays are normalized to solvent control data (which are set to 100%, and therefore do not vary by default), we used a surrogate baseline data set: from each concentration-response curve of the screen compounds, we selected the lowest concentration and assumed that this was in most cases a no-effect concentration. This assumption was consistent with the average of all these data points being about 100% for all assays. With this approach it was possible to visualize the baseline noise (as standard deviation around the average signal, Fig. 3A). From such data, we also calculated the assay-specific coefficients of variation (CoVs, see ToxTemp; suppl. file 1). As a second measure of robustness, we evaluated the responses of each test to the concurrent positive technical controls, which were run along on each plate/for every experiment during the screen (Fig. 3B). The positive controls were also used to determine acceptability of the respective plates/experiments for further evaluation. The plates/experiments, for which the acceptance criteria (see ToxTemp; suppl. file 1) were not met (<10% for all tests), were discarded.

3.3. Performance analysis

The predictivity of the IVB as a whole is a key feature of its regulatory applicability. This was examined as follows: First, all of the above discussed aspects of mechanistic validation were considered: the biology and pathophysiology covered by the entirety of assays of the IVB-EU suggested a high, but not perfect, biological applicability domain. This pointed at a sufficient predictivity for many purposes.

In a second step, we evaluated the capacity of the IVB-EU to correctly identify negative and positive controls. A list of 45 such calibration compounds was assembled from various literature references (Kadereit et al., 2012; Grandjean and Landrigan, 2014; Mundy et al., 2015; Aschner et al., 2017; Paparella et al., 2020; Crofton and Mundy, 2021). The challenges and shortcomings of this approach have been widely discussed (see above references), but our compound selection appeared to be a good compromise based on the present state of knowledge (Fig. 4A and B).

Prediction models for test batteries are an active field of research, and many possibilities exist (tiered approaches, Bayesian models, Boolean rules and decision trees). The difficulty to agree on the defined approaches for the small (3 NAM) battery used to predict dermal sensitization exemplifies these difficulties (Strickland et al., 2022). Here, we used a simple Boolean rule to define a battery hit as any compound that was a hit in one of the included DNT IVB-EU NAMs. A negative was defined as a compound not being a hit in any of the assays. This rule allows for a high transparence and simplicity. For statistical reasons, this battery prediction model may be associated with a high false discovery rate (testing for multiple endpoints considered to be independent). This was considered to be acceptable for screening and prioritization use. Moreover, the use of full concentration-response curves (instead of single data points) for definition of all positive hits reduced this problem. The false discovery rate was further reduced by our use of data from three independent experiments.

The 28 positive controls were used to obtain a preliminary measure of assay sensitivity (to be refined with time and the addition of more control compounds). We used different stringencies of hit definitions to obtain an estimate of the IVB-EU performance with respect to detection of DNT toxicants. When only the specific hits (compounds causing functional impairment at non-cytotoxic concentrations) were counted, the sensitivity of the IVB-EU was 68%. When borderline hits were included, this went up to 82%. When also cytotoxic compounds were included in the "hits", a further increase was observed. However, interpretation of cytotoxic compounds is presently not part of the IVB prediction model (Fig. 4A,C).

The 17 negative controls were used to obtain data on specificity. When specific and borderline hits were counted, a value of 100% was obtained. Specificity dropped to 94%, when also cytotoxic effects were counted as "hit" (Fig. 4B,C).

Altogether, these preliminary performance estimates indicate that a balanced accuracy of about 80% or higher can be reached with the present IVB-EU. Based on the set of positive/negative control compounds, several additional performance measures were calculated (Fig. 4C) and it is particularly noteworthy that the IVB-EU had a high positive predictive value (PPV). This supports the conclusion that compounds identified as a hit should be prioritized for further evaluation of potential human hazard. Such data would also suggest that such chemicals better be excluded at early stages from further development (e.g. as a drug).

Nicotine serves as a good example for gaps in the IVB-EU, identified by the performance evaluation. It was identified as a false negative in the battery, and thus is indicative of a shortcoming with respect to sensitivity. The major action of nicotine is the stimulation of ionotropic acetylcholine receptors, and the IVB-EU does not (yet) include NAMs that would cover this biological function. This information is important when it comes to the interpretation of data from compounds that target nicotinic receptors, like neonicotinoid insecticides (Sheets et al., 2016; Loser et al., 2021a). Assays that fill these gaps are already under development (Fig. 2), and inclusion of assays based on zebra fish embryos and other model organisms (e.g. *C. elegans*) are considered an additional approach to close battery gaps (Atzei et al., 2021; Dasgupta et al., 2022).

Another limitation of the DNT IVB-EU is hard to overcome: the number of control compounds with clearly documented human effects is very limited, and also the compounds having been tested in DNT guideline studies in animals is small (Aschner et al., 2017). For this reason, performance metrics on the basis of currently-available control-compound predictivity will remain superficial. A way forward is to focus more on mechanistic validation approaches (Leist et al., 2012; Judson et al., 2013; Cote et al., 2016; Griesinger et al., 2016; Bal-Price et al., 2018; Andersen et al., 2019; Masjosthusmann et al., 2020) to gain further confidence into the predictivity of the battery for human adversities.

A final, but very important, consideration on predictivity is that this concept is highly context-dependent. In each sharply-defined use domain, it seems important to ask how far the battery is fit-for-purpose. Four issues need to be specified: (i) what regulatory problem is to be addressed (e.g. risk assessment of a new chemical, or prioritization of compounds for further testing); (ii) is there a focus on high positive predictivity or high negative predictivity; (iii) which type of chemicals is being examined (predictivity may be very high within certain compound groups, while it may be low for some compound classes); (iv) which types of biology (targets, pathways) play a role. It is likely that some adverse outcome pathways (AOP) are covered well, while others not at all. For example, acetylcholine esterase inhibitors may not be detected easily by the current IVB-EU, but this gap would be easily filled by an additional enzymatic assay (Li et al., 2017).

3.4. Compound testing and hit identification

In addition to the 45 compounds tested for the IVB-EU performance analyses, all 10 assays were challenged with additional 75 test compounds, so that the total screen comprised 120 chemicals (suppl. file 2). The result of the screen were benchmark concentrations (BMC) of effect (or no effect data within the used concentration range) for 120 compounds on ten functional and six viability endpoints, i.e. 1920 concentration response curves. A matrix including 405 BMCs for the IVB hits (with measures of uncertainty) was generated. To allow a better overview and focus, all compounds were compiled that affected at least one functional endpoint at a non-cytotoxic concentration (n = 59). To better visualize the activity profile of compounds, the endpoints for which toxicants had the highest potency (most sensitive endpoint(s)) were

A

Positive controls	specific + brdl. + cytotox	specific + brdl.	specific
Cadmium chloride	TP	TP	TP
Chlorpyrifos	TP	TP	FN
Dexamethasone	TP	TP	TP
Hexachlorophene	TP	TP	TP
Lead (II) acetate trihydrate	TP	TP	TP
Manganese (II) chloride	TP	TP	TP
Methylmercury chloride	TP	TP	TP
PBDE 47	TP	TP	TP
PBDE 99	TP	TP	FN
(±) Ketamine hydrochloride	FN	FN	FN
5,5-Diphenylhydantoin	FN	FN	FN
Acrylamide	TP	TP	TP
all-trans-Retinoic acid	TP	TP	TP
Chlorpromazine hydrochloride	TP	TP	TP
Deltamethrin	TP	TP	TP
Domoic acid	FN	FN	FN
Haloperidol	TP	TP	TP
Maneb	TP	TP	FN
Methylazoxymethanol acetate	TP	TP	TP
Nicotine	FN	FN	FN
Paraquat dichloride hydrate	TP	TP	TP
PFOA	TP	FN	FN
PFOSK	TP	TP	TP
Sodium valproate	TP	TP	TP
Tebuconazole	TP	TP	TP
Tributyltin chloride	TP	TP	TP
Trichlorfon	TP	TP	TP
Triethyl-tin bromide	TP	TP	FN

B

Negative controls			
Acetaminophen	TN	TN	TN
Amoxicillin	TN	TN	TN
Aspirin	TN	TN	TN
Buspirone	FP	TN	TN
Chlorpheniramine maleate	TN	TN	TN
D-Glucitol	TN	TN	TN
Diethylene glycol	TN	TN	TN
D-Mannitol	TN	TN	TN
Doxylamine succinate	TN	TN	TN
Famotidine	TN	TN	TN
Ibuprofen	TN	TN	TN
Metformin	TN	TN	TN
Metoprolol	TN	TN	TN
Penicillin	TN	TN	TN
Saccharin	TN	TN	TN
Sodium benzoate	TN	TN	TN
Warfarin	TN	TN	TN

C

Performance [%]			
Sensitivity	86	82	68
Specificity	94	100	100
Accuracy	89	89	80
Balanced accuracy	90	91	84
PPV	96	100	100
F1 score	91	91	84
MCC	78	80	67

(caption on next column)

Fig. 4. Performance overview of the test battery (IVB-EU). A set of predefined negative (n = 17) and positive (n = 28) control compounds was included in the set of screening compounds (n = 120). The rationale for their selection is given in Fig. S1 and S2. Note that the controls were randomly included in the overall screening workflow without being given any preferences or special treatment. This means that the standard prediction models of the assays were applied to them, so that they were classified as "no hit", "cytotoxic", "borderline (brdl)" or "specific hit" in individual NAM (see Fig. S3). A reference compound was considered to be a "positive" on the level of the overall IVB-EU, when it was an "alert" in at least one of the individual assays. The tabular display of the figure uses three definitions for an alert: anything that is not a "no hit" (first column), anything that was a specific hit or brdl (second column) or only specific hits (third column). (A) Alerts were considered true positives (TP), non-alerts were considered false negatives (FN). (B) Non-alerts were considered true negatives (TN), alerts were considered false positives (FP). (C) Performance parameters of the current DNT IVB-EU in percent. All parameters were calculated based on the TP, FN, TN, FP as indicated in (A) and (B). PPV: positive predictive value; MCC: Matthews correlation coefficient.

highlighted (Fig. 5). Compounds were considered to be about equally potent across test endpoints, when their activity did not differ by more than a factor of three. This is due to technical issues (the test concentrations were separated by a factor of three in the concentration-response curves), but also due to statistical considerations (the confidence intervals of BMCs separated by factor 3 overlapped in 85% of all cases).

Besides the 59 compounds that produced at least one specific hit (comprising 23 positive controls and 36 other compounds), there were also 61 compounds that had no specific hit in any of the 10 functional endpoints. Ten of these compounds were cytotoxic to one or more cell populations (Fig. S5A), while 51 compounds (including 16 negative controls) had no effect at all (Fig. S5B). This finding of 35 fully negatives (excluding the known negative controls) extends observations from the preliminary predictivity evaluation (using known negative control compounds) that showed that the IVB-EU, despite its large number of tests and endpoints, is not highly unspecific.

3.5. Hit patterns in the DNT IVB screen

Concerning the further analysis of battery hits, several strategies were followed. One approach was to select some individual hit compounds or groups of compounds for further toxicological evaluation. For instance, an expert group of EFSA and the OECD used IVB-EU data on deltamethrine and flufenacet for a case study within the OECD IATA program (EFSA PPR Panel, 2021). Another example is the group of flame retardants, for which the battery data were used to support a comprehensive hazard assessment (Klose et al., 2021a). Such specific toxicological follow-ups were beyond the scope of the present study. Instead, we analyzed general hit patterns of the screen to learn more about the relationship (complementarity/necessity) of the various assays and endpoints.

The first question was, how functional endpoints and specific hits related to the viability endpoints and cytotoxicity hits. To understand the overall data structure, we generated an overview, comparing for each specific hit compound the potency for the most sensitive functional endpoint in the battery (MSE) with the potencies for all cytotoxic effects across the battery test systems (cytotoxicity hits). There were 57 specific hits, plus two compounds (maneb and clorpyrifos), which were classified as borderline hits, and are being included here in the group of functional hits. Altogether 17 of the 59 compounds (29%) did not affect any of the battery's viability endpoints. For this subgroup, the functional endpoint provided a definite gain in sensitivity, compared to cytotoxicity assays. It is also very likely that the functional endpoint was directly affected by the test compounds, i.e. it was not an indirect effect of unspecific cytotoxicity.

As an alternative approach to understand the role of cytotoxicity, we

	Compound / Test method[1]	UKN2	UKN4	UKN5	NPC1	NPC2a [72h]	NPC2a [120h]	NPC2b	NPC2c	NPC3	NPC4 [length]	NPC4 [area]	NPC5	NPC2-5 [# of cells]
positive control compounds	Dexamethasone (DEX)				7.3					5.7				
	Tributyltin chloride (TTC)	6.9			7.2								6.8	
	Hexachlorophene (HCP)	6.4			4.9	6.9					7.1			
	Methylmercury(II) chloride (MMC)		7.0	6.4		6.2	6.3			6.3	6.7	6.3		6.1
	Chlorpromazine hydrochloride (CPH)	5.2		5.5	6.0						5.4	5.3	6.4	
	Deltamethrin (DM)	4.8				6.0		5.4			4.9		6.4	
	Cadmium chloride (CCL)	6.3			5.3	5.8	5.6							5.5
	Triethyltin bromide (TETB)			6.2										
	all-trans-Retinoic acid (RA)	5.3			5.7									
	Tebuconazole (TBN)												5.6	
	Haloperidol (HPD)		4.9	4.9		4.9							5.6	
	PBDE-47[2]	4.3											5.4	
	Trichlorfon (TCF)										5.3			4.9
	PFOSK[3]	4.5								5.0				
	Maneb (MAB)	5.0												
	PBDE 99 (BDE)		4.9											
	Paraquat dichloride hydrate (PQ)	4.0	4.0	3.8			4.9							
	Manganese(II) chloride (MC)												4.7	
	Lead(II) acetate trihydrate (LAT)	4.6												
	Chlorpyrifos (CPF)	4.4	4.0											
	Methylazoxymethanol acetate (MAM)				3.8									
	Sodium valproate (VPA)				3.3					3.3			3.3	
	Acrylamide (AAM)			2.9										
all other hits	Narciclasine (NAR)	7.3	7.9	7.7	7.8				7.2	8.3	7.9	8.0	7.8	7.6
	Rotenone (RTN)	7.5	7.0	7.2		6.8	7.0			6.3	6.3	6.2		6.7
	Glufosinate-ammonium (GFA)									7.1				
	Alpha-Endosulfan (AES)	5.1					4.8						6.4	
	Metaflumizone (MFM)												6.3	
	Endosulfan sulfate (ESS)	5.4											6.3	
	Indoxacarb (IXC)	5.0			6.3								5.8	
	3,3',5,5'-Tetrabromobisphenol A (TBBPA)	5.9											6.3	
	gamma-Cyhalothrin (CHT)												6.1	
	Emamectin benzoate (EMB)	6.0			6.0									
	2-Ethylhexyl diphenyl phosphate (EDP)	5.3			6.0								5.2	
	Beta-Cypermethrin (BCPM)												6.0	
	Malathion (MLT)												5.9	
	Endosulfan (EDS)						4.9						5.7	
	tau-Fluvalinate (TFV)												5.7	
	tert-Butylphenyl diphenyl phosphate (BDP)	5.3					5.1						5.7	
	MPP+[4]		5.7		5.6									
	Tris(1,3-dichloro-2-propyl) phosphate (TDCPP)						4.9						5.6	
	Flubendiamide (FBD)												5.5	
	Fipronil (FPN)	4.8	5.0			4.9				4.9	4.9	4.9	5.5	4.8
	Triphenyl phosphates isopropylated (TPI)	5.2											5.4	
	Tri-o-tolyl phosphate (TOTP)	5.4					5.3						5.2	4.7
	Isodecyl diphenyl phosphate (IDDP)	5.0											5.4	
	Penthiopyrad (PP)					4.8							5.3	
	Chlorpyrifos-oxon (CPFO)				5.2		5.0				4.9			
	Tris(2-butoxyethyl)phosphate (TBOEP)												5.2	
	beta-Cyfluthrin (BCFT)												5.2	
	Carbaryl (CBR)	5.0	4.8			5.1	5.2				5.1	5.0		
	Acibenzolar-S-methyl (ABM)				5.1									
	Tris(methylphenyl) phosphate (TTP)	5.1												
	Clothianidin (CTN)										5.0			
	Triphenylphosphate (TPHP)	5.0											4.9	
	Bisphenol A (BPA)	4.9											4.9	
	Thiacloprid (TC)													4.9
	1-Naphthol (NT)	4.7				4.9								
	Azinphos-methyl (APM)					4.8	4.8							

MSE: within 3-fold MSE range; outside 3-fold range

[1]all concentrations are given in -logM; [2]2,2',4,4'-Tetrabromodiphenyl ether;

[3] Heptadecafluorooctanesulfonic acid potassium salt; [4]1-Methyl-4-phenylpyridinium iodide,

Fig. 5. Hit summary of the IVB-EU screen. Overall, 120 compounds were screened in the current DNT IVB-EU. Screened substances were considered as "hits" when they were classified as a "specific hit" or a "borderline compound" in at least one assay of the battery (assays indicated on top of the columns). The upper section of the table shows all 23 hits amongst the 28 positive controls used in the screen (the remaining five positive controls were no hits). The lower section shows all additional 36 hits amongst the screened compounds. Within the groups, the compounds are ranked based on potency (indicated in units of – log [M]). The table includes all hits of the screen. For each compound, the most sensitive endpoint (MSE) is highlighted. In addition, hits of the respective chemical in other assays, which were of similar potency as in the MSE assay (within a 3-fold range), are also highlighted. The compounds that affected only viability endpoints in the IVB-EU are listed in Fig. S5A. The compounds that affected no endpoint at all are listed in Fig. S5B. Exact and complete screen data (including the uncertainties assessed as 95% confidence interval) are included in a suppl. file 2 – sheet 2 & 3.

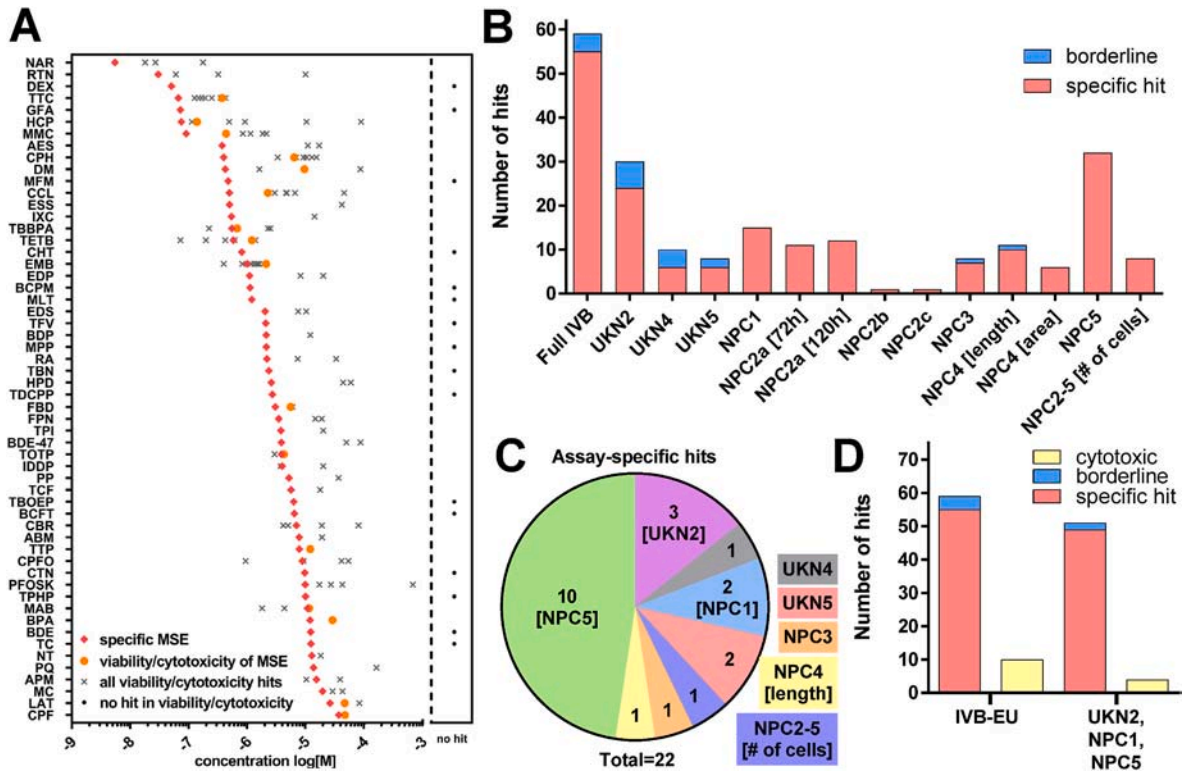


Fig. 6. Contribution of individual NAM to the overall IVB-EU. The screen was performed, hits were identified and the most sensitive endpoint (MSE) was defined for each compound as detailed in Fig. 5 (A). A potency overview of all hit compounds (see Fig. 5 for abbreviation) is displayed: The compounds are sorted according to the potency of their MSE. Note that all MSE data refer to a specific test endpoint (i.e. migration, differentiation, proliferation, neurite growth). In addition, the concentrations at which compounds were detected to be cytotoxic are indicated. Compounds that were not cytotoxic in any assay are indicated by a dot right of the dashed line. The cytotoxic concentration measured in the same assay as the MSE is given a separate symbol (filled circle) to allow an easy overview. Note that for many compounds, no cytotoxicity was measured in the assay that produced the MSE. For design reasons, three low potency compounds were not included in the figure: MAM (MSE = −3.8) orange point at x, 3 additional cytotoxic hits; VPA (MSE = −3.3) orange point at −2.7, four other cytotoxicity hits; AAM (MSE = −2.9) no other cytotoxic hit. All data are given in log(M). (B) The number of hits (out of 120 screen compounds) is indicated for each assay of the battery, and for the total IVB-EU (most leftward bar). The number of specific hits and of borderline hits can both be seen within one bar. The respective set of data for cytotoxic compounds in visualized in Fig. S7. (C) The number of compounds that were a hit in only one assay is displayed for all assays, e.g. 10 compounds were detected only in NPC5, but no other assay; one compound was detected only in UKN4 and no other assay. (D) The number of hits (separated in specific hits, borderline hits and cytotoxic-only compounds) was compared for the full IVB-EU and a hypothetical mini-battery consisting of 3 assays (UKN2, NPC1, NPC5). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

asked, how the MSE concentration related to the cytotoxic potency in the same or in any other assay. There were only five compounds (8%) for which a cytotoxic endpoint was observed at higher (≥ factor 2) potency than the functional MSE (Fig. 6A). One example is carbaryl (CBR), which specifically inhibited neurite growth in the UKN4 assay (functional endpoint). It was particularly potent as cytotoxicant for peripheral neurons and mixed NPC cultures. This may indicate that CBR exerts a cell type-specific cytotoxicity for such neural cell populations. Such viability effects may be relevant for neurodevelopment, but further investigations would be required to allow clear conclusions.

We used a comparison to published data as one preliminary approach to test whether cytotoxicity hits of the IVB-EU are specific for neurodevelopmental cell types. We hypothesized that we may see a difference between cytotoxic potencies on conventional cell lines (HepG2, HEK293, etc.) and on the test systems used here, if a compound shows a developmental-stage specific cytotoxicity. Information on unspecific toxicity (called: cytotoxicity lower bound) was obtained from the ToxCast data base (Judson et al., 2016). For the 41 compounds, for which sufficient data was available, we found that cytotoxicity hit potency in the IVB-EU was at least 10-fold below the cytotoxicity lower bound for 7 compounds; 34 compounds showed no particular sensitivity in IVB-EU test systems compared to cell lines used for ToxCast screening (Fig. S6A). This may indicate that some, but not all cytotoxicity hits may be specific for neurodevelopmental cell types. To complete this comparison, we also checked how the functional hits of the IVB-EU compared to the cytotoxicity lower bound. In general, the cytotoxicity threshold in ToxCast was often in the range of 5–20 μM. Thus, the 17 IVB screen hits with MSEs <1 μM (for which the cytotoxicity lower bound was available), seemed to separate clearly from general cytotoxicity except for TETB. The situation is complex for compounds with higher MSE potency in the IVB-EU. The data set is too small and compound behaviour is very heterogeneous. However, it is plausible, that specificity may be reduced (or lost) at higher screen concentrations (>20 μM). It has been shown that unspecific baseline toxicity increases from this threshold on, due to membrane incorporation and alterations of protein conformations (Escher et al., 2019; Lee et al., 2021, 2022). Therefore, hits in a higher concentration range (e.g. MAM, VPA, AAM) need good justifications (e.g. clinically-observed plasma levels at hit concentration levels) and/or a detailed mechanistic follow-up providing a rationale for specific functional effects in the observed concentration range (Fig. S6B).

All these potency comparisons have an important caveat: the data we obtained are based on nominal concentrations, and these might differ from the free effective concentrations in the medium, and especially at the target sites (Kisitu et al., 2020). Especially, for comparisons to assays with tumor cell lines, it needs to be considered, that such systems usually use serum supplements containing protein and lipids, while most stem cell culture media used here had a low protein and lipid content. Under

the conditions used for the IVB-EU, the free concentrations are very close to the total concentrations in medium (Krebs et al., 2020b), while this is not necessarily the case for serum-containing media.

The second question we asked was, how the hits distributed over the different assays of the battery. Altogether 67 compounds affected at least one test endpoint: 57 specific, 2 borderline, 10 cytotoxic and 51 compounds affected no endpoint at concentrations up to 20 μM (Fig. 6B, Fig. S5&S7). All cytotoxic compounds had potencies of ≥8 μM (Fig. S5A). The number of hits obtained in each assay was also compiled. For instance, the NPC5 assay (examining the KNDP oligodendrocyte differentiation) identified the highest number (n = 34) of specific hits (Fig. 6B). Moreover, 10 compounds were hits only in this assay and would have been missed as potential toxicants without the NPC5 test as part of the IVB-EU (Fig. 6C). The second highest hit rate (n = 30) was found for the UKN2 assay (represents the KNDP of neural crest cell migration). Three compounds were unique hits in this test, i.e. not identified by another endpoint. Most other assays (UKN4, UKN5, NPC1, NPC2a, NPC3 and NPC4) identified 8–15 specific hits, and each of the assay identified at least one test compound that would have been missed by the other tests of the battery (Fig. 6C). This illustrates that the cell types and endpoints assembled in the IVB-EU all differ in the pattern of toxicity pathways and targets they represent. This analysis also showed that the test methods are not redundant, even with this small number (n = 120) of screened chemicals. We anticipate that the broad coverage of cell types, developmental stages and endpoints of the IVB-EU will be even more required to ensure maximal sensitivity, when the chemical space is enlarged by broader test campaigns and a more-wide spread use of the battery.

A third question we asked dealt with resource optimization. Some assays, such as NPC2b/c (migration of neurons and oligodendrocytes) or UKN4 (neurite outgrowth) contributed relatively little to the overall hit rate, and one may consider them to be deleted from the battery or replaced. This would be a step towards a faster, more economical "mini-battery", which would be expected to have a slightly reduced sensitivity, but not greatly reduced overall performance (accuracy; Matthews coefficient). However, in case of the neurosphere assay, individual readouts are multiplexed, meaning that omission of one endpoint will not lead to saving resources, e.g. NPC2b/c are automatically assessed when NPC3/5 are evaluated. As NPC3 is multiplexed with NPC2 and 5, also this assay adds negligible extra time and costs to the overall assays

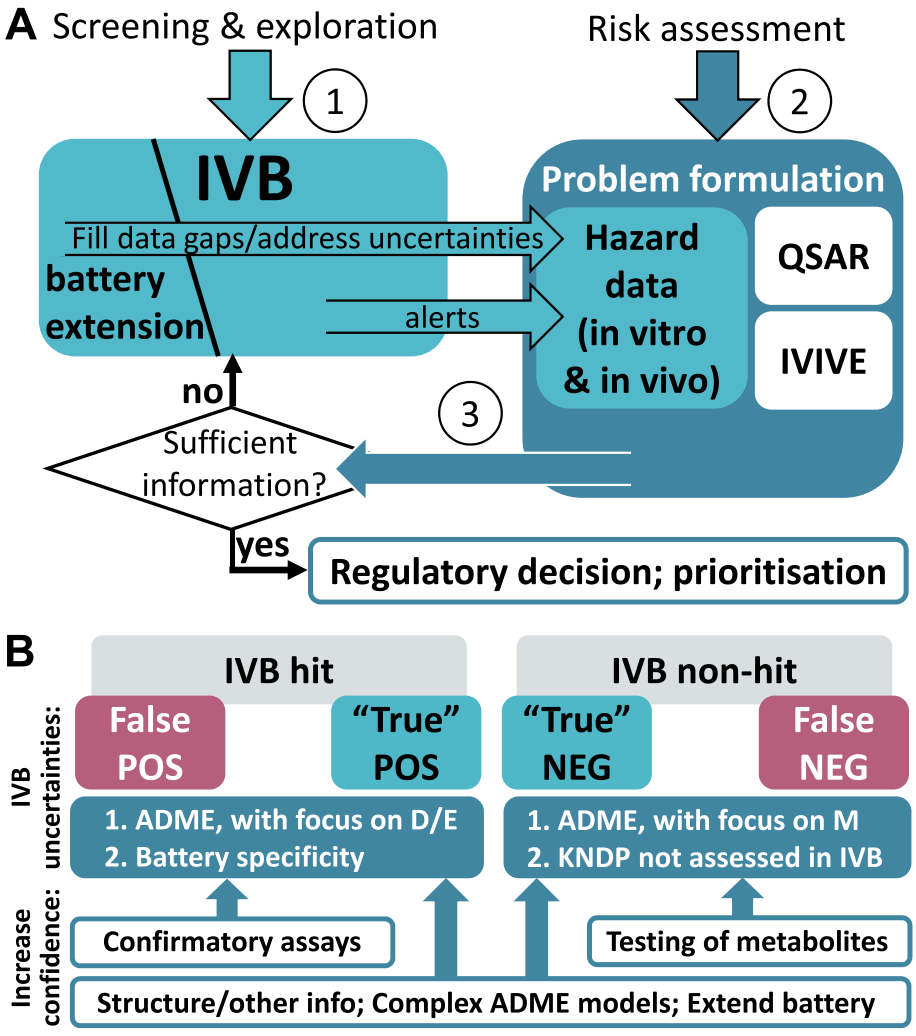


Fig. 7. Outlook on further uses and extensions of the IVB. (A) Incorporation of the IVB into an integrated approach to testing and assessment (IATA): Two different scenarios are depicted. In the first (1) the IVB will be used for screening of compound groups to generate hazard alerts (IVB hits). One way to follow up on these would be in the context of an IATA. In the second scenario (2), risk assessment of single chemicals would be performed in an IATA. This approach starts with a problem formulation (considering or not considering particular exposure situations). In this context all available data on hazard identification and characterization are collected. These may be extended via data of scenario (1). Quantitative structure activity relationships (QSAR) and in vitro-to-in vivo extrapolation (IVIVE) are shown as exemplary elements of the IATA framework. Further elements could include absorption, distribution, metabolism and excretion data (ADME) or an exposure assessment. If the hazard data of the assessed compound are considered not sufficient to derive a robust point of departure (PoD), further information could be obtained from the IVB. (3) In some cases, IVB extensions would be needed to fill data gaps and to reduce uncertainties, until sufficient information is available for regulatory action. (B) Each test method or battery has some uncertainties. The level of uncertainties that can be accepted depends on the problem formulation. For IVB hits and non-hits, one needs to consider that these may be either false positives/negatives, or compounds with a correctly identified hazard ("true" positives/negatives). One potential reason for misidentification is a lack of ADME features represented in the in vitro test systems. For example in vivo distribution and elimination (D/E) features may be misrepresented in the in vitro system. As a result, a compound never reaching the fetal brain because of the placental barrier may show effects on neurons in vitro. In contrast, some false negatives can be explained by a lack of metabolism (M) i.e. in vivo toxic metabolites which are not present in the IVB. Another reason is that a toxicant affects a key neurodevelopmental process (KNDP) that is not included in the IVB. In order to reduce the level of uncertainties and gain confidence into the results, further information can be added (low, white boxes). This includes information transfer across tested compounds (grouping and readacross (RAx)), complex ADME models, confirmatory assays (battery extension), and direct testing of potential metabolites. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

NPC2-5. Hence, a mini-battery should only omit assays that practically save resources, i.e. individual assays. If one continues this line of thought, a minimal DNT IVB may consist of NPC1 (NPC proliferation), NPC2-5 and UKN2 (NCC migration) test methods (Fig. 6D). In our screen, this mini-battery would have identified 52 compounds (88% of all specific and borderline hits) of the 59 hits covered by the whole IVB-EU. Such a reduced approach may be used e.g. for quick/inexpensive pre-screens, e.g. in situations where sensitivity is of low importance, but compounds are to be ranked according to their priority for further testing. However, one may also consider adding an assay to a mini-battery that is not yet included in the IVB-EU. The gap analysis (Fig. 2) suggested that some biological domains are still poorly covered, and that an important gap would be filled by a neural network formation assay (Carstens et al., 2022). Thus, future batteries would need to consider the assays presented here, in addition to other established and emerging DNT NAM.

4. Conclusions and outlook

We have demonstrated here how NAMs with endpoints related to KNDP can be selected and assembled to an in vitro battery to screen for DNT hazard of chemicals. The technical feasibility and the implementation of solid reporting standards have been demonstrated by the use of 120 test compounds in a battery test-run that produced close to 2000 BMCs. These were used to provide battery performance estimates and to classify test compounds as specific hits, cytotoxicants or non-hits. The pattern of results was used to discuss the contribution of the assays and their endpoints to the overall IVB-EU and to define gaps still to be filled.

Pivotal questions for the future are (i) how battery hits would be further used and (ii) how the IVB-EU (or its future expanded version = IVB) could be implemented in a regulatory context (Fig. 7A and B). We anticipate that the first application of the IVB will be for screening of data-poor compounds to explore their DNT liabilities. As the overwhelming majority of chemicals lacks data on DNT hazard, compounds of particular concern (because of high exposure or structural alerts) may be screened first. The IVB would produce alerts for further testing. The underlying toxicological rationale is that disturbance of any KNDP covered by the IVB has the potential to lead to DNT. In a regulatory environment, the IVB data would provide a hazard characterization, and could be used as point-of-departure for further steps. In this context, physiology-based kinetic modelling (PBK) followed by in vitro-to-in vivo extrapolations (IVIVE) could be applied to convert the BMCs to estimated adverse doses (AEDs). These would be used to perform a risk assessment.

With growing experience and confidence into the IVB, its output could become a pivotal element of DNT risk assessment. Such a development is supported by the guidance document on the generation and use of the NAM-based DNT data (Crofton and Mundy, 2021). In a risk assessment situation with a defined problem formulation (e.g. for pesticide marketing re-approval in the EU, or during registration of a chemical in Japan) the compound to be evaluated would be run through the battery to provide hazard data. These might be clear and unambiguous. Or they may need to be complemented by additional rounds of testing in battery extensions. Together with the use of ADME data or other information (such as QSAR) and an IVIVE procedure, sufficient information for risk assessment would be generated (Fig. 7A).

One important aspect of using the battery data as hazard characterization is the interpretation and follow-up of hits. It is at present unclear, whether the number of positive battery endpoints correlates with the strength of DNT hazard. Hence, in the hazard characterization scenario one would be equally concerned if a compound produced one or several hits. However, the BMCs producing the hits have to be considered as multiple hits in the same order of magnitude suggest a higher concern than hits that only produce one low BMC. In the screening and prioritization scenario concern could be based on a combination of BMC magnitude and number of hits similar to the approach practiced in Klose et al. (2021a) in the flame retardant case study. However, singleton-hit chemicals can be of high concern as exemplified by the illustrative example lead, which is one of the best-proven human DNT toxicants and only affected one functional endpoint of the IVB-EU.

For each battery hit, there is always the uncertainty, that it is either a true positive, i.e. that the battery results reflect real DNT hazard for humans, or that it is a false positive (FP). A reasons for the latter scenario may be toxicokinetic (ADME) properties. E.g. a compound may never reach the foetal or child brain because of barrier functions, but there is no such barrier in vitro. Some FP will also arise from test classification uncertainties (alpha error) and the IVB false discovery rate (FDR) due to the combination of a large number of assays. Fortunately, there are also ways to build confidence into the hit pattern and to reduce the uncertainty of a hit being a FP. The assays and their prediction models can be trimmed for high specificity (multiple test runs, full concentration-response curves, conservative thresholds for hit definition). Another powerful approach is to functionally group hit compounds and to use information on one compound to read across to others. This way, consistency and plausibility can be established and/or strengthened.

For some applications, also non-hits play an important role, e.g. for providing confidence to consumers on the safety of food constituents or contaminants. Non-hits may either be true (no hazard) or be false negatives (FN), i.e. have non-discovered toxic properties. The main sources of uncertainty on negatives are the gaps in the battery (KNDP or specific signaling pathway not covered) and toxicokinetic aspects. For instance, a tested parent compound may not be toxic, but a metabolite generated only in vivo may be a DNT toxicant. Fortunately, there are also strategies available to increase confidence in negative hits. If this is of particular importance, the sensitivity of assays can be increased by running a higher number of replicates. Also, a less conservative prediction model may be applied. This strategy is demonstrated here by the introduction of a borderline category, to capture toxic compounds that would otherwise have dropped out of the hit definition. Another major approach is the extension of the battery, e.g. by combination with the US EPA assays (Carstens et al., 2022). Last, but not least, grouping, and other information from data bases and the literature could be used for further evaluation of negative hits and decisions on potential extended testing (Fig. 7A).

Author contributions

All authors read, commented, and approved the manuscript. Jonathan Blum: study conception, investigation, data analysis, supervision, figure design, writing of article. Stefan Masjosthusmann: study conception, data analysis, supervision, figure design, writing of article. Kristina Bartmann: investigation. Farina Bendt: investigation. Xenia Dolde: investigation, data analysis. Anna-Katharina Holzer: investigation, data analysis. Ulrike Hübenthal: investigation. Sadiye Kilic: investigation. Jördis Klose: investigation. Melanie Pahl: investigation. Lynn-Christin Stürzl: investigation. Arif Dönmez: software development, data analysis. Nils Förster: software development, data analysis. Hagen Eike Keβel: software development, data analysis. Martin Scholze: software development, data analysis. Axel Mosig: software development, supervision. Iris Mangas: editing of article. Andrea Terron: editing of article. Kevin Crofton: editing of article. Marcel Leist: study conception, supervision, funding acquisition, project administration, figure design, writing of article. Ellen Fritsche: study conception, supervision, funding acquisition, project administration, figure design, writing of article.

Funding

This work was supported by the European Food Safety Authority (EFSA-Q-2018-00308), the Danish Environmental Protection Agency (EPA), Denmark, under the grant number MST-667-00205, the State

Ministry of Baden-Wuerttemberg, Germany, for Economic Affairs, Labour and Tourism (NAM-Accept), the project CERST (Center for Alternatives to Animal Testing) of the Ministry for culture and science of the State of North-Rhine Westphalia, Germany (file number 233–1.08.03.03- 121972/131–1.08.03.03–121972), the European Chemical Industry Council Long-Range Research Initiative (Cefic LRI) under the project name AIMT11 and the BMBF (NeuroTool). It has also received funding from the European Union's Horizon 2020 research and innovation program under grant agreements No. 964537 (RISK-HUNT3R), No. 964518 (ToxFree), No. 101057014 (PARC) and No. 825759 (ENDpoiNTs).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ellen Fritsche, Kristina Bartmann, Arif Dönmez and Axel Mosig are shareholders of the company DNTOX that provides DNT-IVB assay services. The authors declare no potential conflicts of interest with respect to the research in this article. All other authors have no conflict of interest to declare.

Data availability

Data will be made available on request.

Acknowledgements

The authors are very grateful to Tanja Waldmann as well as Tim Shafer, Katie Paul Friedman and Kelly Carstens (all involved in compound selection, study design, data interpretation and transfer of data to the ToxCast data base). We would also like to thank R. Göttler for the experimental and J. Kapr for layout support. Furthermore, we thank T. Mayer, S. Müller, and the screening centre of the University of Konstanz for the experimental and technical support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2022.137035.

Abbreviations

AOP – adverse outcome pathway
BMC – benchmark concentration
BMCL – lower limit of 95% confidence interval of BMC
BMCU – upper limit of 95% confidence interval of BMC
DIV – days *in vitro*
DNT – developmental neurotoxicity
EFSA – European Food Safety Authority
FDR – false discovery rate
hNPC – human neural progenitor cell
hiPSC – human induced pluripotent stem cell
IVB – *in vitro* battery
IVB-EU – DNT IVB based on methods available in European laboratories
IVIVE – *in vitro* to *in vivo* extrapolation
KNDP – key neurodevelopmental process
MSE – most sensitive endpoint
NAM – new approach methods
PPV – positive predictive value
IATA – integrated approaches for testing and assessment
OECD – Organisation for Economic Co-operation and Development
TN – true negative
TP – true positive
UKN – University of Konstanz
US EPA – United States Environmental Protection Agency

References

Allen, N.J., Lyons, D.A., 2018. Glia as architects of central nervous system formation and function. Science 362, 181–185.
Andersen, M.E., McMullen, P.D., Phillips, M.B., Yoon, M., Pendse, S.N., Clewell, H.J., Hartman, J.K., Moreau, M., Becker, R.A., Clewell, R.A., 2019. Developing context appropriate toxicity testing approaches using new alternative methods (NAMs). ALTEX 36, 523–534.
Aschner, M., Ceccatelli, S., Daneshian, M., Fritsche, E., Hasiwa, N., Hartung, T., Hogberg, H.T., Leist, M., Li, A., Mundi, W.R., Padilla, S., Piersma, A.H., Bal-Price, A., Seiler, A., Westerink, R.H., Zimmer, B., Lein, P.J., 2017. Reference compounds for alternative test methods to indicate developmental neurotoxicity (DNT) potential of chemicals: example lists and criteria for their selection and use. ALTEX 34, 49–74.
Atzei, A., Jense, I., Zwart, E.P., Legradi, J., Venhuis, B.J., van der Ven, L.T.M., Heusinkveld, H.J., Hessel, E.V.S., 2021. Developmental neurotoxicity of environmentally relevant pharmaceuticals and mixtures thereof in a zebrafish embryo behavioural test. Int. J. Environ. Res. Publ. Health 18.
Bal-Price, A., Crofton, K.M., Leist, M., Allen, S., Arand, M., Buetler, T., Delrue, N., FitzGerald, R.E., Hartung, T., Heinonen, T., Hogberg, H., Bennekou, S.H., Lichtensteiger, W., Oggier, D., Paparella, M., Axelstad, M., Piersma, A., Rached, E., Schilter, B., Schmuck, G., Stoppini, L., Tongiorgi, E., Tiramani, M., Monnet-Tschudi, F., Wilks, M.F., Ylikomi, T., Fritsche, E., 2015. International STakeholder NETwork (ISTNET): creating a developmental neurotoxicity (DNT) testing road map for regulatory purposes. Arch. Toxicol. 89, 269–287.
Bal-Price, A., Hogberg, H.T., Crofton, K.M., Daneshian, M., FitzGerald, R.E., Fritsche, E., Heinonen, T., Hougaard Bennekou, S., Klima, S., Piersma, A.H., Sachana, M., Shafer, T.J., Terron, A., Monnet-Tschudi, F., Viviani, B., Waldmann, T., Westerink, R.H.S., Wilks, M.F., Witters, H., Zurich, M.G., Leist, M., 2018. Recommendation on test readiness criteria for new approach methods in toxicology: exemplified for developmental neurotoxicity. ALTEX 35, 306–352.
Bal-Price, A.K., Coecke, S., Costa, L., Crofton, K.M., Fritsche, E., Goldberg, A., Grandjean, P., Lein, P.J., Li, A., Lucchini, R., Mundy, W.R., Padilla, S., Persico, A.M., Seiler, A.E., Kreysa, J., 2012. Advancing the science of developmental neurotoxicity (DNT): testing for better safety evaluation. ALTEX 29, 202–215.
Baumann, J., Gassmann, K., Masjosthusmann, S., DeBoer, D., Bendt, F., Giersiefer, S., Fritsche, E., 2016. Comparative human and rat neurospheres reveal species differences in chemical effects on neurodevelopmental key events. Arch. Toxicol. 90, 1415–1427.
Behl, M., Ryan, K., Hsieh, J.H., Parham, F., Shapiro, A.J., Collins, B.J., Sipes, N.S., Birnbaum, L.S., Bucher, J.R., Foster, P.M.D., Walker, N.J., Paules, R.S., Tice, R.R., 2019. Screening for developmental neurotoxicity at the national toxicology program: the future is here. Toxicol. Sci. 167, 6–14.
Bellinger, D.C., 2012. A strategy for comparing the contributions of environmental chemicals and other risk factors to neurodevelopment of children. Environ. Health Perspect. 120, 501–507.
Bennett, D., Bellinger, D.C., Birnbaum, L.S., Bradman, A., Chen, A., Cory-Slechta, D.A., Engel, S.M., Fallin, M.D., Halladay, A., Hauser, R., Hertz-Picciotto, I., Kwiatkowski, C.F., Lanphear, B.P., Marquez, E., Marty, M., McPartland, J., Newschaffer, C.J., Payne-Sturges, D., Patisaul, H.B., Perera, F.P., Ritz, B., Sass, J., Schantz, S.L., Webster, T.F., Whyatt, R.M., Woodruff, T.J., Zoeller, R.T., Anderko, L., Campbell, C., Conry, J.A., DeNicola, N., Gould, R.M., Hirtz, D., Huffling, K., Landrigan, P.J., Lavin, A., Miller, M., Mitchell, M.A., Rubin, L., Schettler, T., Tran, H. L., Acosta, A., Brody, C., Miller, E., Miller, P., Swanson, M., Witherspoon, N.O., American College of, O., , Gynecologists, Child Neurology, S., Endocrine, S., International Neurotoxicology, A., International Society for Children's, H., the, E., International Society for Environmental, E., National, 2016. Council of asian pacific islander, P., national hispanic medical, A., national medical, A.,. In: Project TENDR: Targeting Environmental Neuro-Developmental Risks the TENDR Consensus Statement. Environ Health Perspect, vol. 124, pp. A118–A122.
Brull, M., Spreng, A.S., Gutbier, S., Loser, D., Krebs, A., Reich, M., Kraushaar, U., Britschgi, M., Patsch, C., Leist, M., 2020. Incorporation of stem cell-derived astrocytes into neuronal organoids to allow neuro-glial interactions in toxicological studies. ALTEX 37, 409–428.
Carstens, K.E., Carpenter, A.F., Martin, M.M., Harrill, J.A., Shafer, T.J., Paul Friedman, K., 2022. Integrating data from in vitro new approach methodologies for developmental neurotoxicity. Toxicol. Sci. 187, 62–79.
Chesnut, M., Paschoud, H., Repond, C., Smirnova, L., Hartung, T., Zurich, M.G., Hogberg, H.T., Pamies, D., 2021. Human IPSC-derived model to study myelin disruption. Int. J. Mol. Sci. 22.
Coecke, S., Goldberg, A.M., Allen, S., Buzanska, L., Calamandrei, G., Crofton, K., Hareng, L., Hartung, T., Knaut, H., Honegger, P., Jacobs, M., Lein, P., Li, A., Mundy, W., Owen, D., Schneider, S., Silbergeld, E., Reum, T., Trnovec, T., Monnet-Tschudi, F., Bal-Price, A., 2007. Workgroup report: incorporating in vitro alternative methods for developmental neurotoxicity into international hazard and risk assessment strategies. Environ. Health Perspect. 115, 924–931.
Cote, I., Andersen, M.E., Ankley, G.T., Barone, S., Birnbaum, L.S., Boekelheide, K., Bois, F.Y., Burgoon, L.D., Chiu, W.A., Crawford-Brown, D., Crofton, K.M., DeVito, M., Devlin, R.B., Edwards, S.W., Guyton, K.Z., Hattis, D., Judson, R.S., Knight, D., Krewski, D., Lambert, J., Maull, E.A., Mendrick, D., Paoli, G.M., Patel, C. J., Perkins, E.J., Poje, G., Portier, C.J., Rusyn, I., Schulte, P.A., Simeonov, A., Smith, M.T., Thayer, K.A., Thomas, R.S., Thomas, R., Tice, R.R., Vandenberg, J.J., Villeneuve, D.L., Wesselkamper, S., Whelan, M., Whittaker, C., White, R., Xia, M., Yauk, C., Zeise, L., Zhao, J., DeWoskin, R.S., 2016. The next generation of risk

assessment multi-year study-highlights of findings, applications to risk assessment, and future directions. Environ. Health Perspect. 124, 1671–1682.
Crofton, K.M., Mundy, W.R., 2021. External scientific report on the interpretation of data from the developmental neurotoxicity in vitro testing assays for use in integrated approaches for testing and assessment. EFSA Support. Pub. 18, 6924E.
Crofton, K.M., Mundy, W.R., Lein, P.J., Bal-Price, A., Coecke, S., Seiler, A.E., Knaut, H., Buzanska, L., Goldberg, A., 2011. Developmental neurotoxicity testing: recommendations for developing alternative methods for the screening and prioritization of chemicals. ALTEX 28, 9–15.
Crofton, K.M., Mundy, W.R., Shafer, T.J., 2012. Developmental neurotoxicity testing: a path forward. Congenital. Anom. 52, 140–146.
Dasgupta, S., Simonich, M.T., Tanguay, R.L., 2022. Zebrafish behavioral assays in toxicology. Methods Mol. Biol. 2474, 109–122.
Delp, J., Gutbier, S., Klima, S., Hoelting, L., Pinto-Gil, K., Hsieh, J.H., Aichem, M., Klein, K., Schreiber, F., Tice, R.R., Pastor, M., Behl, M., Leist, M., 2018. A high-throughput approach to identify specific neurotoxicants/developmental toxicants in human neuronal cell function assays. ALTEX 35, 235–253.
Dreser, N., Madjar, K., Holzer, A.K., Kapitza, M., Scholz, C., Kranaster, P., Gutbier, S., Klima, S., Kolb, D., Dietz, C., Trefzer, T., Meisig, J., van Thriel, C., Henry, M., Berthold, M.R., Bluthgen, N., Sachinidis, A., Rahnenfuhrer, J., Hengstler, J.G., Waldmann, T., Leist, M., 2020. Development of a neural rosette formation assay (RoFA) to identify neurodevelopmental toxicants and to characterize their transcriptome disturbances. Arch. Toxicol. 94, 151–171.
Escher, B.I., Glauch, L., Konig, M., Mayer, P., Schlichting, R., 2019. Baseline toxicity and volatility cutoff in reporter gene assays used for high-throughput screening. Chem. Res. Toxicol. 32, 1646–1655.
Escher, S.E., Partosch, F., Konzok, S., Jennings, P., Luijten, M., Kienhuis, A., de Leeuw, V., Reuss, R., Lindemann, K.-M., Bennekou, S.H., 2022. Development of a roadmap for action on new approach methodologies in risk assessment. EFSA Support. Pub. 19, 7341E.
Forster, N., Butke, J., Kessel, H.E., Bendt, F., Pahl, M., Li, L., Fan, X., Leung, P.C., Klose, J., Masjosthusmann, S., Fritsche, E., Mosig, A., 2022. Reliable identification and quantification of neural cells in microscopic images of neurospheres. Cytometry 101, 411–422.
Frank, C.L., Brown, J.P., Wallace, K., Mundy, W.R., Shafer, T.J., 2017. From the cover: developmental neurotoxicants disrupt activity in cortical networks on microelectrode arrays: results of screening 86 compounds during neural network formation. Toxicol. Sci. 160, 121–135.
Fritsche, E., Crofton, K.M., Hernandez, A.F., Hougaard Bennekou, S., Leist, M., Bal-Price, A., Reaves, E., Wilks, M.F., Terron, A., Solecki, R., Sachana, M., Gourmelon, A., 2017. OECD/EFSA workshop on developmental neurotoxicity (DNT): the use of non-animal test methods for regulatory purposes. ALTEX 34, 311–315.
Fritsche, E., Grandjean, P., Crofton, K.M., Aschner, M., Goldberg, A., Heinonen, T., Hessel, E.V.S., Hogberg, H.T., Bennekou, S.H., Lein, P.J., Leist, M., Mundy, W.R., Paparella, M., Piersma, A.H., Sachana, M., Schmuck, G., Solecki, R., Terron, A., Monnet-Tschudi, F., Wilks, M.F., Witters, H., Zurich, M.G., Bal-Price, A., 2018. Consensus statement on the need for innovation, transition and implementation of developmental neurotoxicity (DNT) testing for regulatory purposes. Toxicol. Appl. Pharmacol. 354, 3–6.
Grandjean, P., Abdennebi-Najar, L., Barouki, R., Cranor, C.F., Etzel, R.A., Gee, D., Heindel, J.J., Hougaard, K.S., Hunt, P., Nawrot, T.S., Prins, G.S., Ritz, B., Soffritti, M., Sunyer, J., Weihe, P., 2019. Timescales of developmental toxicity impacting on research and needs for intervention. Basic Clin. Pharmacol. Toxicol. 125 (Suppl. 3), 70–80.
Grandjean, P., Landrigan, P.J., 2006. Developmental neurotoxicity of industrial chemicals. Lancet 368, 2167–2178.
Grandjean, P., Landrigan, P.J., 2014. Neurobehavioural effects of developmental toxicity. Lancet Neurol. 13, 330–338.
Griesinger, C., Desprez, B., Coecke, S., Casey, W., Zuang, V., 2016. Validation of alternative in vitro methods to animal testing: concepts, challenges, processes and tools. Adv. Exp. Med. Biol. 856, 65–132.
Gutbier, S., May, P., Berthelot, S., Krishna, A., Trefzer, T., Behbehani, M., Efremova, L., Delp, J., Gstraunthaler, G., Waldmann, T., Leist, M., 2018. Major changes of cell function and toxicant sensitivity in cultured cells undergoing mild, quasi-natural genetic drift. Arch. Toxicol. 92, 3487–3503.
Harrill, J.A., Freudenrich, T., Wallace, K., Ball, K., Shafer, T.J., Mundy, W.R., 2018. Testing for developmental neurotoxicity using a battery of in vitro assays for key cellular events in neurodevelopment. Toxicol. Appl. Pharmacol. 354, 24–39.
Hartung, T., 2007. Food for thought ... on validation. ALTEX 24, 67–80.
Hartung, T., Bremer, S., Casati, S., Coecke, S., Corvi, R., Fortaner, S., Gribaldo, L., Halder, M., Hoffmann, S., Roi, A.J., Prieto, P., Sabbioni, E., Scott, L., Worth, A., Zuang, V., 2004. A modular approach to the ECVAM principles on test validity. Altern Lab Anim 32, 467–472.
Hartung, T., Hoffmann, S., Stephens, M., 2013. Mechanistic validation. ALTEX 30, 119–130.
Hoelting, L., Klima, S., Karreman, C., Grinberg, M., Meisig, J., Henry, M., Rotshteyn, T., Rahnenfuhrer, J., Bluthgen, N., Sachinidis, A., Waldmann, T., Leist, M., 2016. Stem cell-derived immature human dorsal root ganglia neurons to identify peripheral neurotoxicants. Stem Cells Transl Med 5, 476–487.
Holzer, A.K., Suciu, I., Karreman, C., Goj, T., Leist, M., 2022. Specific attenuation of purinergic signaling during bortezomib-induced peripheral neuropathy in vitro. Int. J. Mol. Sci. 23.
Hu, W., Liu, C.W., Jimenez, J.A., McCoy, E.S., Hsiao, Y.C., Lin, W., Engel, S.M., Lu, K., Zylka, M.J., 2022. Detection of azoxystrobin fungicide and metabolite azoxystrobin-acid in pregnant women and children, estimation of daily intake, and evaluation of placental and lactational transfer in mice. Environ. Health Perspect. 130, 27013.
Jaklin, M., Zhang, J.D., Schafer, N., Clemann, N., Barrow, P., Kung, E., Sach-Peltason, L., McGinnis, C., Leist, M., Kustermann, S., 2022. Optimization of the TeraTox assay for preclinical teratogenicity assessment. Toxicol. Sci. 188, 17–33.
Jensen, S.M., Kluxen, F.M., Streibig, J.C., Cedergreen, N., Ritz, C., 2020. bmd: an R package for benchmark dose estimation. PeerJ 8, e10557.
Judson, R., Houck, K., Martin, M., Richard, A.M., Knudsen, T.B., Shah, I., Little, S., Wambaugh, J., Woodrow Setzer, R., Kothiya, P., Phuong, J., Filer, D., Smith, D., Reif, D., Rotroff, D., Kleinstreuer, N., Sipes, N., Xia, M., Huang, R., Crofton, K., Thomas, R.S., 2016. Editor's highlight: analysis of the effects of cell stress and cytotoxicity on in vitro assay activity across a diverse chemical and assay space. Toxicol. Sci. 152, 323–339.
Judson, R., Kavlock, R., Martin, M., Reif, D., Houck, K., Knudsen, T., Richard, A., Tice, R. R., Whelan, M., Xia, M., Huang, R., Austin, C., Daston, G., Hartung, T., Fowle 3rd, J. R., Wooge, W., Tong, W., Dix, D., 2013. Perspectives on validation of high-throughput assays supporting 21st century toxicity testing. ALTEX 30, 51–56.
Kadereit, S., Zimmer, B., van Thriel, C., Hengstler, J.G., Leist, M., 2012. Compound selection for in vitro modeling of developmental neurotoxicity. Front Biosci (Landmark Ed) 17, 2442–2460.
Kappenberg, F., Brecklinghaus, T., Albrecht, W., Blum, J., van der Wurp, C., Leist, M., Hengstler, J.G., Rahnenfuhrer, J., 2020. Handling deviating control values in concentration-response curves. Arch. Toxicol. 94, 3787–3798.
Keßel, 2022. Biostatistics and its impact on hazard characterization using in vitro developmental neurotoxicity assays. ALTEX. https://doi.org/10.1101/2022.10.18.512648. Submitted for publication.
Kisitu, J., Hollert, H., Fisher, C., Leist, M., 2020. Chemical concentrations in cell culture compartments (C5) - free concentrations. ALTEX 37, 693–708.
Klima, S., Brull, M., Spreng, A.S., Suciu, I., Falt, T., Schwamborn, J.C., Waldmann, T., Karreman, C., Leist, M., 2021. A human stem cell-derived test system for agents modifying neuronal N-methyl-D-aspartate-type glutamate receptor Ca(2+)-signalling. Arch. Toxicol. 95, 1703–1722.
Klose, J., Li, L., Pahl, M., Bendt, F., Hubenthal, U., Jungst, C., Petzsch, P., Schauss, A., Kohrer, K., Leung, P.C., Wang, C.C., Koch, K., Tigges, J., Fan, X., Fritsche, E., 2022. Application of the adverse outcome pathway concept for investigating developmental neurotoxicity potential of Chinese herbal medicines by using human neural progenitor cells in vitro. Cell Biol. Toxicol.
Klose, J., Pahl, M., Bartmann, K., Bendt, F., Blum, J., Dolde, X., Forster, N., Holzer, A.K., Hubenthal, U., Kessel, H.E., Koch, K., Masjosthusmann, S., Schneider, S., Sturzl, L.C., Woeste, S., Rossi, A., Covaci, A., Behl, M., Leist, M., Tigges, J., Fritsche, E., 2021a. Neurodevelopmental toxicity assessment of flame retardants using a human DNT in vitro testing battery. Cell Biol. Toxicol.
Klose, J., Tigges, J., Masjosthusmann, S., Schmuck, K., Bendt, F., Hubenthal, U., Petzsch, P., Kohrer, K., Koch, K., Fritsche, E., 2021b. TBBPA targets converging key events of human oligodendrocyte development resulting in two novel AOPs. ALTEX 38, 215–234.
Koch, K., Bartmann, K., Hartmann, J., Kapr, J., Klose, J., Kuchovska, E., Pahl, M., Schluppmann, K., Zuhr, E., Fritsche, E., 2022. Scientific validation of human neurosphere assays for developmental neurotoxicity evaluation. Front Toxicol 4, 816370.
Krebs, A., Nyffeler, J., Karreman, C., Schmidt, B.Z., Kappenberg, F., Mellert, J., Pallocca, G., Pastor, M., Rahnenfuhrer, J., Leist, M., 2020a. Determination of benchmark concentrations and their statistical uncertainty for cytotoxicity test data and functional in vitro assays. ALTEX 37, 155–163.
Krebs, A., Nyffeler, J., Rahnenfuhrer, J., Leist, M., 2018. Normalization of data for viability and relative cell function curves. ALTEX 35, 268–271.
Krebs, A., van Vugt-Lussenburg, B.M.A., Waldmann, T., Albrecht, W., Boei, J., Ter Braak, B., Brajnik, M., Braunbeck, T., Brecklinghaus, T., Busquet, F., Dinnyes, A., Dokler, J., Dolde, X., Exner, T.E., Fisher, C., Fluri, D., Forsby, A., Hengstler, J.G., Holzer, A.K., Janstova, Z., Jennings, P., Kisitu, J., Kobolak, J., Kumar, M., Limonciel, A., Lundqvist, J., Mihalik, B., Moritz, W., Pallocca, G., Ulloa, A.P.C., Pastor, M., Rovida, C., Sarkans, U., Schimming, J.P., Schmidt, B.Z., Stober, R., Strassfeld, T., van de Water, B., Wilmes, A., van der Burg, B., Verfaillie, C.M., von Hellfeld, R., Vrieling, H., Vrijenhoek, N.G., Leist, M., 2020b. The EU-ToxRisk method documentation, data processing and chemical testing pipeline for the regulatory use of new approach methods. Arch. Toxicol. 94, 2435–2461.
Krebs, A., Waldmann, T., Wilks, M.F., Van Vugt-Lussenburg, B.M.A., Van der Burg, B., Terron, A., Steger-Hartmann, T., Ruegg, J., Rovida, C., Pedersen, E., Pallocca, G., Luijten, M., Leite, S.B., Kustermann, S., Kamp, H., Hoeng, J., Hewitt, P., Herzler, M., Hengstler, J.G., Heinonen, T., Hartung, T., Hardy, B., Gantner, F., Fritsche, E., Fant, K., Ezendam, J., Exner, T., Dunkern, T., Dietrich, D.R., Coecke, S., Busquet, F., Braeuning, A., Bondarenko, O., Bennekou, S.H., Beilmann, M., Leist, M., 2019. Template for the description of cell-based toxicological test methods to allow evaluation and regulatory use of the data. ALTEX 36, 682–699.
Krug, A.K., Balmer, N.V., Matt, F., Schonenberger, F., Merhof, D., Leist, M., 2013a. Evaluation of a human neurite growth assay as specific screen for developmental neurotoxicants. Arch. Toxicol. 87, 2215–2231.
Krug, A.K., Gutbier, S., Zhao, L., Poltl, D., Kullmann, C., Ivanova, V., Forster, S., Jagtap, S., Meiser, J., Leparc, G., Schildknecht, S., Adam, M., Hiller, K., Farhan, H., Brunner, T., Hartung, T., Sachinidis, A., Leist, M., 2014. Transcriptional and metabolic adaptation of human neurons to the mitochondrial toxicant MPP(+). Cell Death Dis. 5, e1222.
Krug, A.K., Kolde, R., Gaspar, J.A., Rempel, E., Balmer, N.V., Meganathan, K., Vojnits, K., Baquie, M., Waldmann, T., Ensenat-Waser, R., Jagtap, S., Evans, R.M., Julien, S., Peterson, H., Zagoura, D., Kadereit, S., Gerhard, D., Sotiriadou, I., Heke, M., Natarajan, K., Henry, M., Winkler, J., Marchan, R., Stoppini, L., Bosgra, S.,

Westerhout, J., Verwei, M., Vilo, J., Kortenkamp, A., Hescheler, J., Hothorn, L., Bremer, S., van Thriel, C., Krause, K.H., Hengstler, J.G., Rahnenfuhrer, J., Leist, M., Sachinidis, A., 2013b. Human embryonic stem cell-derived test systems for developmental neurotoxicity: a transcriptomics approach. Arch. Toxicol. 87, 123–143.

Lee, J., Braun, G., Henneberger, L., Konig, M., Schlichting, R., Scholz, S., Escher, B.I., 2021. Critical membrane concentration and mass-balance model to identify baseline cytotoxicity of hydrophobic and ionizable organic chemicals in mammalian cell lines. Chem. Res. Toxicol. 34, 2100–2109.

Lee, J., Escher, B.I., Scholz, S., Schlichting, R., 2022. Inhibition of neurite outgrowth and enhanced effects compared to baseline toxicity in SH-SY5Y cells. Arch. Toxicol. 96, 1039–1053.

Lein, P., Locke, P., Goldberg, A., 2007. Meeting report: alternatives for developmental neurotoxicity testing. Environ. Health Perspect. 115, 764–768.

Leist, M., Efremova, L., Karreman, C., 2010. Food for thought ... considerations and guidelines for basic test method descriptions in toxicology. ALTEX 27, 309–317.

Leist, M., Hartung, T., Nicotera, P., 2008. The dawning of a new age of toxicology. ALTEX 25, 103–114.

Leist, M., Hasiwa, N., Daneshian, M., Hartung, T., 2012. Validation and quality control of replacement alternatives – current status and future challenges. Toxicology Research 1, 8–22.

Leist, M., Hasiwa, N., Rovida, C., Daneshian, M., Basketter, D., Kimber, I., Clewell, H., Gocht, T., Goldberg, A., Busquet, F., Rossi, A.M., Schwarz, M., Stephens, M., Taalman, R., Knudsen, T.B., McKim, J., Harris, G., Pamies, D., Hartung, T., 2014. Consensus report on the future of animal-free systemic toxicity testing. ALTEX 31, 341–356.

Leontaridou, M., Urbisch, D., Kolle, S.N., Ott, K., Mulliner, D.S., Gabbert, S., Landsiedel, R., 2017. The borderline range of toxicological methods: quantification and implications for evaluating precision. ALTEX 34, 525–538.

Li, S., Huang, R., Solomon, S., Liu, Y., Zhao, B., Santillo, M.F., Xia, M., 2017. Identification of acetylcholinesterase inhibitors using homogenous cell-based assays in quantitative high-throughput screening platforms. Biotechnol. J. 12.

Loser, D., Hinojosa, M.G., Blum, J., Schaefer, J., Brull, M., Johansson, Y., Suciu, I., Grillberger, K., Danker, T., Moller, C., Gardner, I., Ecker, G.F., Bennekou, S.H., Forsby, A., Kraushaar, U., Leist, M., 2021a. Functional alterations by a subgroup of neonicotinoid pesticides in human dopaminergic neurons. Arch. Toxicol. 95, 2081–2107.

Loser, D., Schaefer, J., Danker, T., Moller, C., Brull, M., Suciu, I., Uckert, A.K., Klima, S., Leist, M., Kraushaar, U., 2021b. Human neuronal signaling and communication assays to assess functional neurotoxicity. Arch. Toxicol. 95, 229–252.

Lotharius, J., Falsig, J., van Beek, J., Payne, S., Dringen, R., Brundin, P., Leist, M., 2005. Progressive degeneration of human mesencephalic neuron-derived cells triggered by dopamine-dependent oxidative stress is dependent on the mixed-lineage kinase pathway. J. Neurosci. 25, 6329–6342.

Lupu, D., Andersson, P., Bornehag, C.G., Demeneix, B., Fritsche, E., Gennings, C., Lichtensteiger, W., Leist, M., Leonards, P.E.G., Ponsonby, A.L., Scholze, M., Testa, G., Tresguerres, J.A.F., Westerink, R.H.S., Zalc, B., Ruegg, J., 2020. The ENDpoiNTs project: novel testing strategies for endocrine disruptors linked to developmental neurotoxicity. Int. J. Mol. Sci. 21.

Makris, S.L., Raffaele, K., Allen, S., Bowers, W.J., Hass, U., Alleva, E., Calamandrei, G., Sheets, L., Amcoff, P., Delrue, N., Crofton, K.M., 2009. A retrospective performance assessment of the developmental neurotoxicity study in support of OECD test guideline 426. Environ. Health Perspect. 117, 17–25.

Masjosthusmann, S., Becker, D., Petzuch, B., Klose, J., Siebert, C., Deenen, R., Barenys, M., Baumann, J., Dach, K., Tigges, J., Hubenthal, U., Kohrer, K., Fritsche, E., 2018. A transcriptome comparison of time-matched developing human, mouse and rat neural progenitor cells reveals human uniqueness. Toxicol. Appl. Pharmacol. 354, 40–55.

Masjosthusmann, S., Blum, J., Bartmann, K., Dolde, X., Holzer, A.-K., Stürzl, L.-C., Keßel, E.H., Förster, N., Dönmez, A., Klose, J., Pahl, M., Waldmann, T., Bendt, F., Kisitu, J., Suciu, I., Hübenthal, U., Mosig, A., Leist, M., Fritsche, E., 2020. Establishment of an a priori protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of developmental neurotoxicity. EFSA Support. Pub. 17, 1938E.

Meisig, J., Dreser, N., Kapitza, M., Henry, M., Rotshteyn, T., Rahnenfuhrer, J., Hengstler, J.G., Sachinidis, A., Waldmann, T., Leist, M., Bluthgen, N., 2020. Kinetic modeling of stem cell transcriptome dynamics to identify regulatory modules of normal and disturbed neuroectodermal differentiation. Nucleic Acids Res. 48, 12577–12592.

Modafferi, S., Zhong, X., Kleensang, A., Murata, Y., Fagiani, F., Pamies, D., Hogberg, H. T., Calabrese, V., Lachman, H., Hartung, T., Smirnova, L., 2021. Gene-environment interactions in developmental neurotoxicity: a case study of synergy between chlorpyrifos and CHD8 knockout in human BrainSpheres. Environ. Health Perspect. 129, 77001.

Mundy, W.R., Padilla, S., Breier, J.M., Crofton, K.M., Gilbert, M.E., Herr, D.W., Jensen, K. F., Radio, N.M., Raffaele, K.C., Schumacher, K., Shafer, T.J., Cowden, J., 2015. Expanding the test set: chemicals with potential to disrupt mammalian brain development. Neurotoxicol. Teratol. 52, 25–35.

Nimtz, L., Hartmann, J., Tigges, J., Masjosthusmann, S., Schmuck, M., Kessel, E., Theiss, S., Kohrer, K., Petzsch, P., Adjaye, J., Wigmann, C., Wieczorek, D., Hildebrandt, B., Bendt, F., Hubenthal, U., Brockerhoff, G., Fritsche, E., 2020. Characterization and application of electrically active neuronal networks established from human induced pluripotent stem cell-derived neural progenitor cells for neurotoxicity evaluation. Stem Cell Res. 45, 101761.

Nimtz, L., Klose, J., Masjosthusmann, S., Barenys, M., Fritsche, E., 2019. The neurosphere assay as an in vitro method for developmental neurotoxicity (DNT) evaluation. In: Aschner, M., Costa, L. (Eds.), Cell Culture Techniques. Springer New York, New York, NY, pp. 141–168.

Nunes, C., Singh, P., Mazidi, Z., Murphy, C., Bourguignon, A., Wellens, S., Chandrasekaran, V., Ghosh, S., Zana, M., Pamies, D., Thomas, A., Verfaillie, C., Culot, M., Dinnyes, A., Hardy, B., Wilmes, A., Jennings, P., Grillari, R., Grillari, J., Zurich, M.G., Exner, T., 2022. An in vitro strategy using multiple human induced pluripotent stem cell-derived models to assess the toxicity of chemicals: a case study on paraquat. Toxicol. Vitro 81, 105333.

Nyffeler, J., Karreman, C., Leisner, H., Kim, Y.J., Lee, G., Waldmann, T., Leist, M., 2017. Design of a high-throughput human neural crest cell migration assay to indicate potential developmental toxicants. ALTEX 34, 75–94.

OECD, 2007. Test No. 426. Developmental Neurotoxicity Study.

OECD, 2021. Guideline No. 497: Defined Approaches on Skin Sensitisation.

Pallocca, G., Grinberg, M., Henry, M., Frickey, T., Hengstler, J.G., Waldmann, T., Sachinidis, A., Rahnenfuhrer, J., Leist, M., 2016. Identification of transcriptome signatures and biomarkers specific for potential developmental toxicants inhibiting human neural crest cell migration. Arch. Toxicol. 90, 159–180.

Pallocca, G., Leist, M., 2022. On the usefulness of animals as a model system (part II): considering benefits within distinct use domains. ALTEX 39, 531–539.

Pallocca, G., Mone, M.J., Kamp, H., Luijten, M., Van de Water, B., Leist, M., 2022a. Next-generation Risk Assessment of Chemicals - Rolling Out a Human-Centric Testing Strategy to Drive 3R Implementation: the RISK-Hunt3r Project Perspective. ALTEX.

Pallocca, G., Nyffeler, J., Dolde, X., Grinberg, M., Gstraunthaler, G., Waldmann, T., Rahnenfuhrer, J., Sachinidis, A., Leist, M., 2017. Impairment of human neural crest cell migration by prolonged exposure to interferon-beta. Arch. Toxicol. 91, 3385–3402.

Pallocca, G., Rovida, C., Leist, M., 2022b. On the usefulness of animals as a model system (part I): overview of criteria and focus on robustness. ALTEX 39, 347–353.

Paparella, M., Bennekou, S.H., Bal-Price, A., 2020. An analysis of the limitations and uncertainties of in vivo developmental neurotoxicity testing and assessment to identify the potential for alternative approaches. Reprod. Toxicol. 96, 327–336.

Patterson, E.A., Whelan, M.P., Worth, A.P., 2021. The role of validation in establishing the scientific credibility of predictive toxicology approaches intended for regulatory application. Comput Toxicol 17, 100144.

Piersma, A.H., van Benthem, J., Ezendam, J., Kienhuis, A.S., 2018. Validation redefined. Toxicol. Vitro 46, 163–165.

Pistollato, F., Carpi, D., Mendoza-de Gyves, E., Paini, A., Bopp, S.K., Worth, A., Bal-Price, A., 2021. Combining in vitro assays and mathematical modelling to study developmental neurotoxicity induced by chemical mixtures. Reprod. Toxicol. 105, 101–119.

Products, E.Panel o.P.P., Residues, t., Hernández-Jerez, A., Adriaanse, P., Aldrich, A., Berny, P., Coja, T., Duquesne, S., Focks, A., Marinovich, M., Millet, M., Pelkonen, O., Pieper, S., Tiktak, A., Topping, C., Widenfalk, A., Wilks, M., Wolterink, G., Crofton, K., Hougaard Bennekou, S., Paparella, M., Tzoulaki, I., 2021. Development of Integrated Approaches to Testing and Assessment (IATA) case studies on developmental neurotoxicity (DNT) risk assessment. EFSA J. 19, e06599.

Ritz, C., Baty, F., Streibig, J.C., Gerhard, D., 2015. Dose-response analysis using R. PLoS One 10, e0146021.

Ryan, K.R., Sirenko, O., Parham, F., Hsieh, J.H., Cromwell, E.F., Tice, R.R., Behl, M., 2016. Neurite outgrowth in human induced pluripotent stem cell-derived neurons as a high-throughput screen for developmental neurotoxicity or neurotoxicity. Neurotoxicology 53, 271–281.

Sachana, M., Bal-Price, A., Crofton, K.M., Bennekou, S.H., Shafer, T.J., Behl, M., Terron, A., 2019. International regulatory and scientific effort for improved developmental neurotoxicity testing. Toxicol. Sci. 167, 45–57.

Sachana, M., Willett, C., Pistollato, F., Bal-Price, A., 2021. The potential of mechanistic information organised within the AOP framework to increase regulatory uptake of the developmental neurotoxicity (DNT) in vitro battery of assays. Reprod. Toxicol. 103, 159–170.

Schmidt, B.Z., Lehmann, M., Gutbier, S., Nembo, E., Noel, S., Smirnova, L., Forsby, A., Hescheler, J., Avci, H.X., Hartung, T., Leist, M., Kobolak, J., Dinnyes, A., 2017. In vitro acute and developmental neurotoxicity screening: an overview of cellular platforms and high-throughput technical possibilities. Arch. Toxicol. 91, 1–33.

Schmuck, M.R., Temme, T., Dach, K., de Boer, D., Barenys, M., Bendt, F., Mosig, A., Fritsche, E., 2017. Omnisphero: a high-content image analysis (HCA) approach for phenotypic developmental neurotoxicity (DNT) screenings of organoid neurosphere cultures in vitro. Arch. Toxicol. 91, 2017–2028.

Scholz, D., Poltl, D., Genewsky, A., Weng, M., Waldmann, T., Schildknecht, S., Leist, M., 2011. Rapid, complete and large-scale generation of post-mitotic neurons from the human LUHMES cell line. J. Neurochem. 119, 957–971.

Sheets, L.P., Li, A.A., Minnema, D.J., Collier, R.H., Creek, M.R., Peffer, R.C., 2016. A critical review of neonicotinoid insecticides for developmental neurotoxicity. Crit. Rev. Toxicol. 46, 153–190.

Shinde, V., Hoelting, L., Srinivasan, S.P., Meisig, J., Meganathan, K., Jagtap, S., Grinberg, M., Liebing, J., Bluethgen, N., Rahnenfuhrer, J., Rempel, E., Stoeber, R., Schildknecht, S., Forster, S., Godoy, P., van Thriel, C., Gaspar, J.A., Hescheler, J., Waldmann, T., Hengstler, J.G., Leist, M., Sachinidis, A., 2017. Definition of transcriptome-based indices for quantitative characterization of chemically disturbed stem cell development: introduction of the STOP-Toxukn and STOP-Toxukk tests. Arch. Toxicol. 91, 839–864.

Simon, J.M., Paranjape, S.R., Wolter, J.M., Salazar, G., Zylka, M.J., 2019. High-throughput screening and classification of chemicals and their effects on neuronal gene expression using RASL-seq. Sci. Rep. 9, 4529.

Smirnova, L., Hogberg, H.T., Leist, M., Hartung, T., 2014. Developmental neurotoxicity - challenges in the 21st century and in vitro opportunities. ALTEX 31, 129–156.

Spreng, A.S., Brull, M., Leisner, H., Suciu, I., Leist, M., 2022. Distinct and dynamic transcriptome adaptations of iPSC-generated astrocytes after cytokine stimulation. In: Cells 11.
Stiegler, N.V., Krug, A.K., Matt, F., 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.
Strickland, J., Truax, J., Corvaro, M., Settivari, R., Henriquez, J., McFadden, J., Gulledge, T., Johnson, V., Gehen, S., Germolec, D., Allen, D.G., Kleinstreuer, N., 2022. Application of defined approaches for skin sensitization to agrochemical products. Front Toxicol 4, 852856.
Thomas, R.S., Paules, R.S., Simeonov, A., Fitzpatrick, S.C., Crofton, K.M., Casey, W.M., Mendrick, D.L., 2018. The US Federal Tox21 Program: a strategic and operational plan for continued leadership. ALTEX 35, 163–168.
Tohyama, C., 2016. Developmental neurotoxicity test guidelines: problems and perspectives. J. Toxicol. Sci. 41, SP69–SP79.
Tsuji, R., Crofton, K.M., 2012. Developmental neurotoxicity guideline study: issues with methodology, evaluation and regulation. Congenital. Anom. 52, 122–128.
Us Epa Ord, C.f.C.T., 2022. ToxCast Database: Invitrodb Version 3.5. The United States Environmental Protection Agency's Center for Computational Toxicology and Exposure.
USEPA, 1998. Health Effects Test Guidelines OCSPP 870.6300 Developmental Neurotoxicity Study. Washington, DC.
van Thriel, C., Westerink, R.H., Beste, C., Bale, A.S., Lein, P.J., Leist, M., 2012. Translating neurobehavioural endpoints of developmental neurotoxicity tests into in vitro assays and readouts. Neurotoxicology 33, 911–924.
Vinken, M., Benfenati, E., Busquet, F., Castell, J., Clevert, D.A., de Kok, T.M., Dirven, H., Fritsche, E., Geris, L., Gozalbes, R., Hartung, T., Jennen, D., Jover, R., Kandarova, H., Kramer, N., Krul, C., Luechtefeld, T., Masereeuw, R., Roggen, E., Schaller, S., Vanhaecke, T., Yang, C., Piersma, A.H., 2021. Safer chemicals using less animals: kick-off of the European ONTOX project. Toxicology 458, 152846.
Worth, A.P., Balls, M., 2001. The importance of the prediction model in the validation of alternative tests. Altern Lab Anim 29, 135–144.

Supplementary information for

Establishment of a human cell-based in vitro battery to assess developmental neurotoxicity hazard of chemicals

Jonathan Blum[1,#], Stefan Masjosthusmann[2,#], Kristina Bartmann[2], Farina Bendt[2], Xenia Dolde[1], Arif Dönmez[2], Nils Förster[4], Anna-Katharina Holzer[1], Ulrike Hübenthal[2], Hagen Eike Keßel[2], Sadiye Kilic[1], Jördis Klose[2], Melanie Pahl[2], Lynn-Christin Stürzl[2], Iris Mangas[5], Andrea Terron[5], Martin Scholze[6], Axel Mosig[4], Marcel Leist[1,°], Ellen Fritsche[2,3,°]

[#]these authors contributed equally; [°]these authors contributed equally

Table of Contents

	Positive controls (28)	Group	Reference	Comments
1	**5,5-Diphenylhydantoin**	pharmaceutical drug	[1]	Anti-seizure medication.
2	**Acrylamide**	industrial chemical	[2], [6]	Generated also in processed food; known neurotoxicant.
3	**all-trans Retinoic acid**	signalling molecule	[1]	Morphogen involved in brain development. Human evidence. Vitamin A metabolite.
4	**Cadmium chloride**	heavy metal	[1], [3]	DNT evidence based on animal and human data.
5	**Chlorpromazine**	pharmaceutical drug	[1]	First generation neuroleptic; multiple receptor inhibitions
6	**Chlorpyrifos**	pesticide	[1], [8]	Inhibitor of AChE. Evidence for DNT in humans.
7	**Deltamethrin**	pesticide	[4], [6]	DNT evidence (including epidemiological human studies) summarized by EFSA panel 2021.
8	**Dexamethasone**	pharmaceutical drug	[1]	Glucocorticoid
9	**Domoic acid**	environmental	[1], [7]	Causes shellfish poisoning. DNT evidence from animal studies.
10	**Haloperidol**	pharmaceutical drug	[1]	First generation neuroleptic; multiple receptor inhibitions
11	**Hexachlorophene**	industrial/pesticide	[1]	Human evidence. Disinfectant.
12	**(±) Ketamine**	pharmaceutical drug	[1]	NMDA receptor antagonist.
13	**Lead (II) acetate**	heavy metal	[1], [5]	Human evidence.
14	**Maneb**	pesticide	[1]	Thiourea fungicide containing manganese.
15	**Manganese (II) chloride**	metal	[1], [8]	Human evidence.
16	**Methylazoxymethanol**	environmental	[1]	Targets neuroblasts in CNS. Used in animal models to induce disease phenotypes.
17	**Methylmercury chloride**	metals	[1], [5]	Human evidence.
18	**Nicotine**	environmental	[1], [9]	Agonist of nAChRs. Human evidence.
19	**Paraquat dichloride hydrate**	pesticide	[1]	Herbicide. Linked to development of Parkinson's disease.
20	**PBDE 47**	industrial chemical	[1], [8]	Bromoaromatic flame retardant.
21	**PBDE 99**	industrial chemical	[1], [8]	Bromoaromatic flame retardant.
22	**PFOA**	industrial chemical	[1]	Used as industrial surfactant. Perfluorinated carboxylic acid.
23	**PFOSK**	industrial chemical	[1]	Used as industrial surfactant. Perfluorinated sulfonic acid.
24	**Sodium valproate**	pharmaceutical drug	[1]	Human evidence. Used to treat epilepy and bipolar disorders.
25	**Tebuconazole**	pesticide	[6]	Triazole fungicide.
26	**Tributyltin chloride**	industrial chemical	[6]	Organotin compound. Inhibitor of mitochondrial ATP synthase.
27	**Trichlorfon**	pesticide	[6]	Inhibitor of AChE. Prodrug, which is activated non-enzymatically into dichlorvos (DDVP).
28	**Triethyl-tin bromide**	organotin	[1]	Neurotoxic organotin. Toxic to myelin.

Fig. S1: Commented list of positive controls used in the IVB-EU

A rough functional grouping of the 28 chemicals used as positive controls is provided. Details on the compounds (CAS-number, full name, abbreviation, etc.) are provided in the suppl. Excel sheet. The reference column indicates the source of information used for classification of the compounds as positive controls. [1] Aschner et al. (2017); [2] Chain (2015); [3] Chandravanshi et al. (2021); [4] EFSA PPR Panel (2021); [5] Grandjean and Landrigan (2014); [6] Mundy et al. (2015); [7] Costa et al. (2010); [8] Grandjean and Landrigan (2006); [9] LeSage et al. (2006). Full citations are found in the references chapter of this suppl. document.

	Negative controls (17)	Group	Comments
1	**Acetaminophen***	drug	Pain medication during pregnancy.
2	**Amoxicillin***	antibiotic	Used to treat infections during pregnancy.
3	**Aspirin**	drug	Prescribed against pre-eclampsia.
4	**Buspirone**	drug	Anxiolytic
5	**Chlorpheniramine**	drug	Antihistamine
6	**D-Glucitol***	sugar derivative	Converted in body to fructose.
7	**Diethylene glycol***	solvent	Metabolite ethylen glycol is toxic at high conc.
8	**D-Mannitol***	sugar derivative	Sweetener
9	**Doxylamine**	drug	Antihistamine
10	**Famotidine**	drug	Histamine H2-receptor antagonist (anti-ulcer).
11	**Ibuprofen***	drug	Pain medication; COX-inhibitor.
12	**Metformin**	drug	Type 2 diabetes medication.
13	**Metoprolol***	drug	β-receptor blocker
14	**Penicillin**	antibiotic	Used to treat bacterial infections.
15	**Saccharin***	food additive	Sweetener
16	**Sodium benzoate**	food additive	Antioxidant
17	**Warfarin***	drug	Reproductive toxicant, but not DNT.

Fig. S2: Commented list of negative controls used in the IVB-EU

A rough functional grouping of the 17 chemicals used as negative controls is provided. Details on the compounds (CAS-number, full name, abbreviation, etc.) are provided in the suppl. Excel sheet. Note that the negative classification refers not only to the compounds as such, but to the compounds used in a concentration range of up to 20 µM. In this range, literature data, and often clinical use suggest the absence of effects or of mechanisms relevant to DNT. *: suggested as negative control in Aschner et al. (2017).

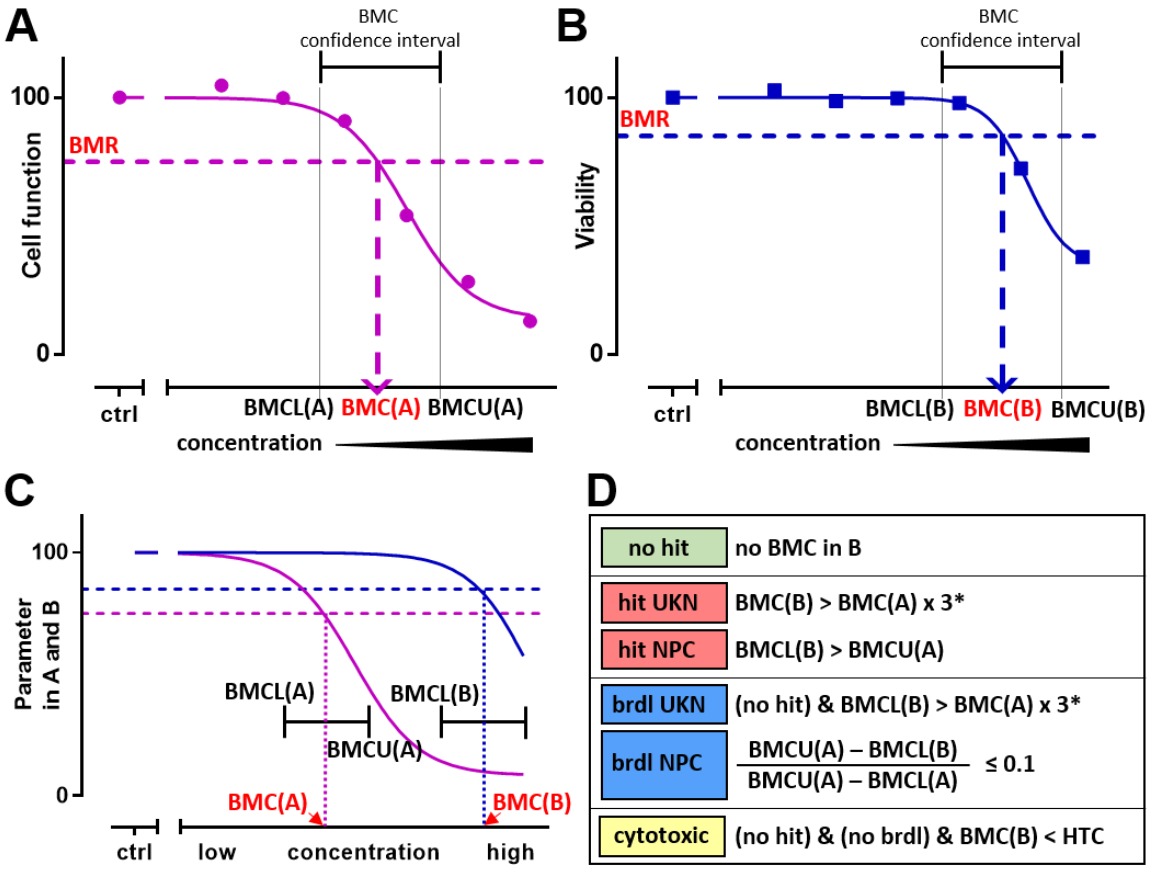


Fig. S3: Classifications of test compounds as hits and alerts

For each compound and each assay, two sets of concentration-response data were produced, one for the main functional endpoint of the NAM (e.g. migration, proliferation, neurite growth or differentiation) and one for the viability of the test system used. Summary data (e.g. the benchmark concentration (BMC) and its confidence interval) were produced from both data sets and used for classification of compounds. **(A)** Example of a data set on a functional endpoint. In each assay, a benchmark response (BMR) was defined (see ToxTemps annexes) as threshold between effect and no effect. The intersection of the BMR with the concentration-response curve defined the BMC. The uncertainty of the BMC was expressed by a confidence interval with the BMCL as lower limit and the BMCU as upper limit,. BMC (A), BMCL(A) and BMCU(A) are the specific values of the example curve A. (**B**) Example of a data set on a viability endpoint. In each assay, a benchmark response (BMR) was defined as threshold between effect and no effect. Note that BMRs are assay-specific. BMC(B), BMCL(B) and BMCU(B) are the specific values of the example curve B. (**C**) An example is given for a data set for a compound that would be considered a screen hit: the BMC(A) and BMC(B) are separated by a large extent. For compounds with less separation, a borderline classification would result. Cytotoxic compounds would show no separation. Inactive compounds would have no responses. (**D**) Quantitative classification scheme according to the principle described qualitatively in (C): specific hits (hits), borderline hits (brdl) and cytotoxic hits (cytotoxic) would all be considered as "alerts". They can be grouped in different ways for hit definitions and statistics (Fig. 4). The definitions are given for all assays (UKN = UKN2, UKN4, UKN5; NPC = NPC1-5) according to the respective assays' ToxTemp description. *UKN assays are defined by ratios between summary data for functional endpoint and viability. A ratio of three is indicated here exemplarily and applies to the UKN5 test. Other ratios are part of the prediction models of UKN2 and UKN4.

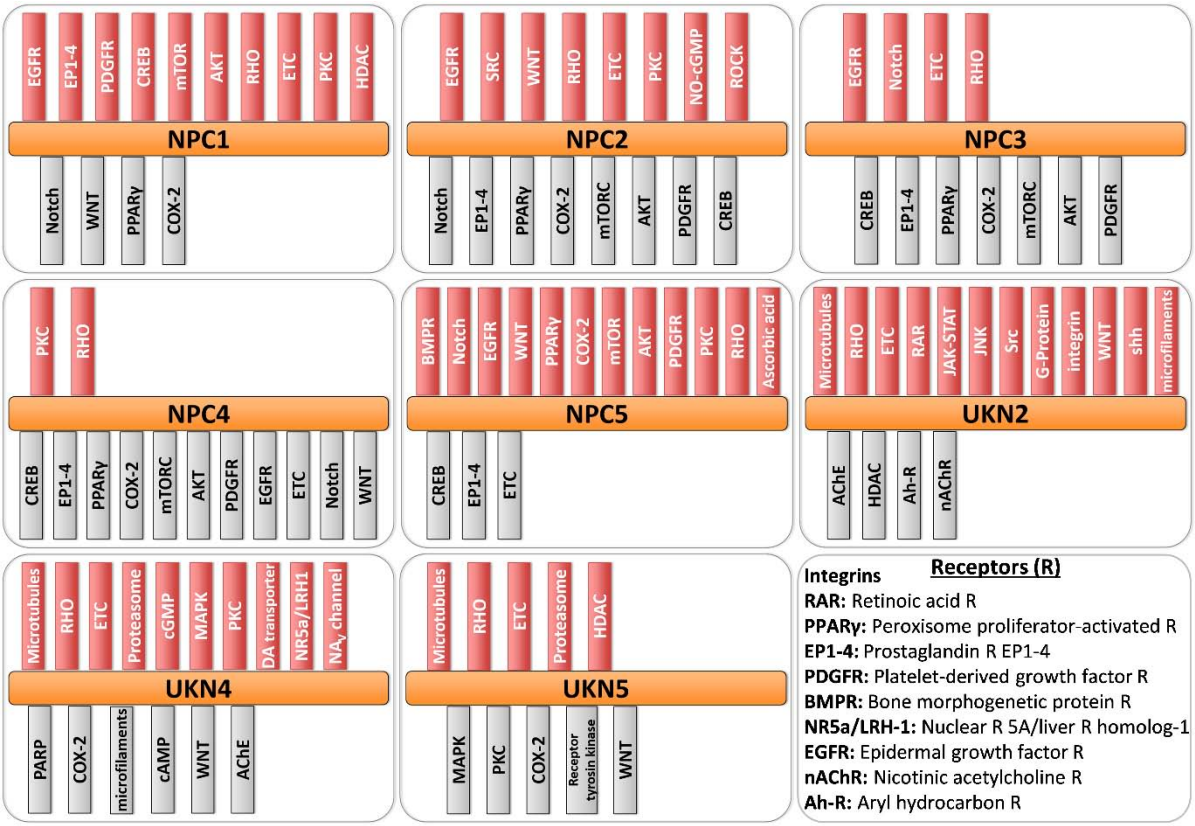


Fig. S4: Overview of biological pathways known to contribute to the readouts of NAM used in the IVB-EU

During the setup and readiness evaluation of the assays, pathway specific "tool compounds" were tested and evaluated for their effect on the test endpoint. These compounds are mostly pharmacological inhibitors or activators of enzymes/receptors/transporters with high specificity for their target. When they effected the test endpoint, it was assumed that the pathway or biochemical mechanism affected by these compounds played a role in the test system, so that it affected the overall readout. For instance, if modulators of the Rho/ROCK signaling cascade affected a test endpoint, it was concluded that toxicants that regulate this pathway would also be detected (displayed by bars on top of the assay name). If modulators of a pathway/target did not affect a test endpoint, it was concluded that a toxicant affecting the respective target or pathway would not be detected by the test (bars below the assay name). In many cases, pathway modulation was only tested in one direction (e.g. only inhibitors of the electron transport chain or only activators of the wnt pathway). This leaves open whether opposite modulator would also have an effect. For such details, original publications, (Koch et al., 2022) and (Masjosthusmann et al., 2020) give more details. Abbreviations of receptors are given in the figure. Notch: notch signalling pathway; COX-2: cyclooxygenase-2; CREB: cAMP response element-binding protein; mTOR: mammalian target of rapamycin; AKT: protein kinase B; ETC: electron transport chain; PKC: protein kinase C; HDAC: histone deacetylase; SRC: proto-oncogene tryosin-protein kinase Src; NO-cGMP: nitric oxide-cGMP sensitive kinase; ROCK: Rho-associated protein kinase; JNK: c-Jun N-terminal kinases; shh: sonic hedgehoc protein; AChE: acetylcholinesterase; PARP: Poly (ADP-ribose) polymerase; MAPK; mitogen-activated protein kinase; DA: dopamine; NA_V: voltage gated sodium channel; WNT: wnt signaling; JAK-STAT: JAK-STAT signaling pathway; cGMP: cGMP-related signal transduction; cAMP: cAMP- related signal transduction

A

Compound	Viability assays					
	UKN2	UKN4	UKN5	NPC1	NPC2a [72 h]	NPC2-5 [120 h]
Buspirone		BMR nr*				
Ethylene Thiourea				5.1		
Flufenacet	4.1					
Glycerol		BMR nr*				
Malaoxon	4.1					
Mancozeb			BMR nr*			
Omethoate			4.1			
Parathion-methyl						5.0
Perfluorooctanoic acid		4.3	3.1			
Tri-allate		4.7				

B

	compound	category
1	(-)-Nicotine	positive control
2	(+-)-Ketamine hydrochloride	positive control
3	5,5-Diphenylhydantoin	positive control
4	Acephate	screening compound
5	Acetaminophen	negative control
6	Acetamiprid	screening compound
7	Aldicarb	screening compound
8	Amoxicillin	negative control
9	Aspirin	negative control
10	Bis-(2-butoxyethyl)phosphate	screening compound
11	Boscalid	screening compound
12	Captopril	screening compound
13	Chlorpheniramine maleate	negative control
14	Chlorpyrifos-methyl	screening compound
15	Cymoxanil	screening compound
16	Cypermethrin	screening compound
17	D-Glucitol	negative control
18	Diazinon	screening compound
19	Diethylene glycol	negative control
20	Dimethoate	screening compound
21	Dinotefuran	screening compound
22	Disulfoton	screening compound
23	D-Mannitol	negative control
24	Domoic acid	positive control
25	Doxylamine succinate	negative control
26	Etofenprox	screening compound

	compound	category
27	Famotidine	negative control
28	Fenamidone	screening compound
29	Ibuprofen	negative control
30	Imidacloprid	screening compound
31	Mepiquat chloride	screening compound
32	Metformin	negative control
33	Methamidophos	screening compound
34	Methimazole	screening compound
35	Metoprolol	negative control
36	Octamethylcyclotetrasiloxane	screening compound
37	Penicillin VK	negative control
38	Pymetrozine	screening compound
39	Saccharin	negative control
40	Sodium benzoate	negative control
41	Sodium chlorite	screening compound
42	Sodium L-glutamate hydrate	screening compound
43	Sodium perchlorate	screening compound
44	Spirodiclofen	screening compound
45	Tembotrione	screening compound
46	Terbutaline hemisulfate	screening compound
47	Thiamethoxam	screening compound
48	Topramezone	screening compound
49	Tris(2-Chloroisopropyl)phosphate	screening compound
50	Tris(chloroethyl)phosphate	screening compound
51	Warfarin	negative control

Fig. S5: List of compounds that had only cytotoxic or no effects

Based on the screen results, all 61 compounds were selected that produced no specific hit on any of the assays (no functional endpoint affected at non-cytotoxic concentrations). (**A**) Ten compounds were found to be cytotoxic in at least one assay. The table lists the viability assessment belonging to the mentioned functional assays (e.g. UKN2 means the viability assay run within the functional testing of UKN2 and thus assessing effects on neural crest cells as test system). The cytotoxic potency is given in units of –log(M). BMR nr*: the concentration-response curve did not cross the BMR (defined in this assay at 75% for hit classification). But compounds reduced viability by more than 10%, which is defined in this assay as cytotoxicity alert. Therefore, compounds have no BMC value according to the classification scheme, but are still cytotoxic according to the assays own prediction model (at the highest screen concentration). (**B**) List of all 51 compounds that were neither cytotoxic nor produced any other alert across the IVB-EU.

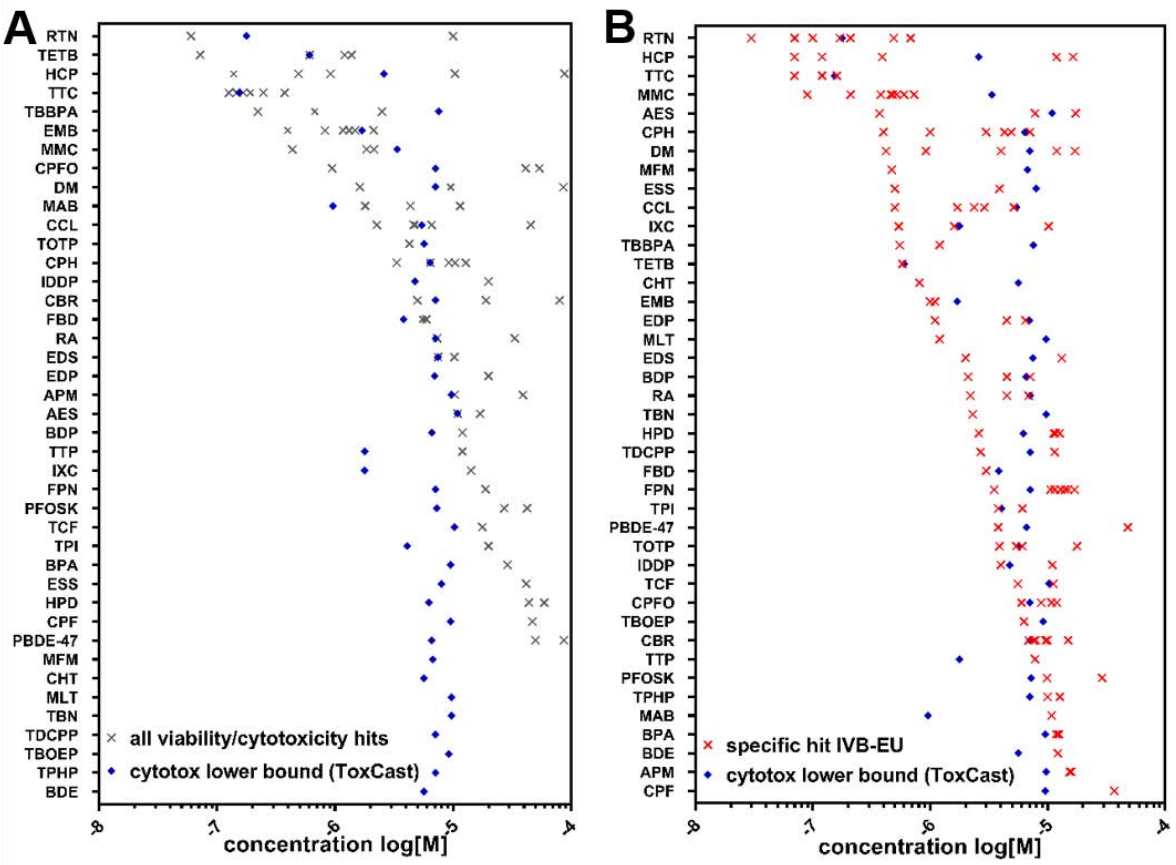


Fig. S6: Screen hits of IVB-EU in comparison to ToxCast cytotoxicity assays

Screen hits (see Fig. 5 for all screen hits and abbreviations) are compared to the cytotox lower bound (CBL) across all ToxCast cytotoxicity assays extracted from the EPA ToxCast Screening Library (https://comptox.epa.gov/dashboard/chemical-lists/toxcast). The CBL is calculated as 3-times the median absolut deviation below the median of all hits across the set of ToxCast cytotoxicity assays for each compounds with at least 2 cytotoxicity hits (Judson et al., 2016). Compound with less than 2 hits are left out. (**A**) Cytotoxic concentrations of all screen hits across the IVB-EU are compared to the CBL. Compounds are sorted by their cytotoxicity potency in the IVB-EU. The lower eight compounds have no cytotoxicity hit in the IVB-EU. (**B**) Screen hits of specific test endpoints (i.e. migration, differentiation, proliferation, neurite outgrowth) across the IVB-EU are compared to the CLB. Compounds are sorted according to the potency of their most sensitive endpoint. All data is given in log(M).

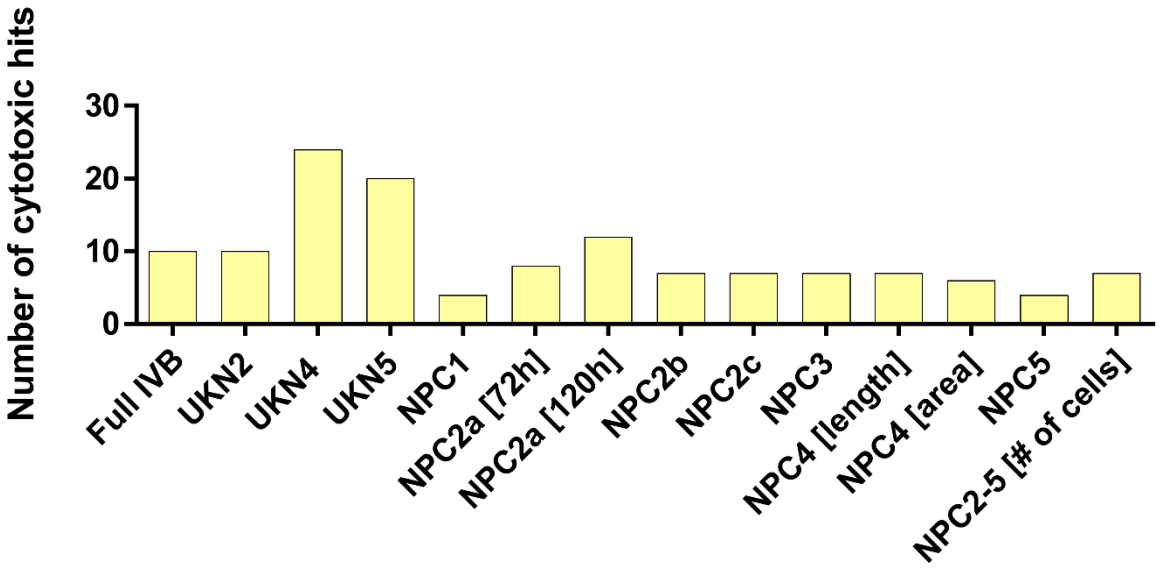


Fig. S7: Numbers of compounds detected by each assay of the IVB-EU as being cytotoxic

The screen was performed and hits were identified as detailed in Fig. S3. The number of cytotoxic hits (out of 120 screen compounds) is indicated for each assay of the battery, and for the total IVB-EU (most leftward bar). The number of specific hits and of borderline hits can be seen in Fig. 6.

References

Aschner, M., Ceccatelli, S., Daneshian, M., Fritsche, E., Hasiwa, N., Hartung, T., Hogberg, H.T., Leist, M., Li, A., Mundi, W.R., Padilla, S., Piersma, A.H., Bal-Price, A., Seiler, A., Westerink, R.H., Zimmer, B., Lein, P.J., 2017. Reference compounds for alternative test methods to indicate developmental neurotoxicity (DNT) potential of chemicals: example lists and criteria for their selection and use. ALTEX 34, 49-74.

Chain, E.P.o.C.i.t.F., 2015. Scientific Opinion on acrylamide in food. EFSA Journal 13, 4104.

Chandravanshi, L., Shiv, K., Kumar, S., 2021. Developmental toxicity of cadmium in infants and children: a review. Environ Anal Health Toxicol 36, e2021003-2021000.

Costa, L.G., Giordano, G., Faustman, E.M., 2010. Domoic acid as a developmental neurotoxin. Neurotoxicology 31, 409-423.

Grandjean, P., Landrigan, P.J., 2006. Developmental neurotoxicity of industrial chemicals. Lancet 368, 2167-2178.

Grandjean, P., Landrigan, P.J., 2014. Neurobehavioural effects of developmental toxicity. Lancet Neurol 13, 330-338.

Judson, R., Houck, K., Martin, M., Richard, A.M., Knudsen, T.B., Shah, I., Little, S., Wambaugh, J., Woodrow Setzer, R., Kothiya, P., Phuong, J., Filer, D., Smith, D., Reif, D., Rotroff, D., Kleinstreuer, N., Sipes, N., Xia, M., Huang, R., Crofton, K., Thomas, R.S., 2016. Editor's Highlight: Analysis of the Effects of Cell Stress and Cytotoxicity on In Vitro Assay Activity Across a Diverse Chemical and Assay Space. Toxicol Sci 152, 323-339.

Koch, K., Bartmann, K., Hartmann, J., Kapr, J., Klose, J., Kuchovska, E., Pahl, M., Schluppmann, K., Zuhr, E., Fritsche, E., 2022. Scientific Validation of Human Neurosphere Assays for Developmental Neurotoxicity Evaluation. Front Toxicol 4, 816370.

LeSage, M.G., Gustaf, E., Dufek, M.B., Pentel, P.R., 2006. Effects of maternal intravenous nicotine administration on locomotor behavior in pre-weanling rats. Pharmacol Biochem Behav 85, 575-583.

Masjosthusmann, S., Blum, J., Bartmann, K., Dolde, X., Holzer, A.-K., Stürzl, L.-C., Keßel, E.H., Förster, N., Dönmez, A., Klose, J., Pahl, M., Waldmann, T., Bendt, F., Kisitu, J., Suciu, I., Hübenthal, U., Mosig, A., Leist, M., Fritsche, E., 2020. Establishment of an a priori protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of developmental neurotoxicity. EFSA Supporting Publications 17, 1938E.

Mundy, W.R., Padilla, S., Breier, J.M., Crofton, K.M., Gilbert, M.E., Herr, D.W., Jensen, K.F., Radio, N.M., Raffaele, K.C., Schumacher, K., Shafer, T.J., Cowden, J., 2015. Expanding the test set: Chemicals with potential to disrupt mammalian brain development. Neurotoxicol Teratol 52, 25-35.

Products, E.Panel o.P.P., Residues, t., Hernández-Jerez, A., Adriaanse, P., Aldrich, A., Berny, P., Coja, T., Duquesne, S., Focks, A., Marinovich, M., Millet, M., Pelkonen, O., Pieper, S., Tiktak, A., Topping, C., Widenfalk, A., Wilks, M., Wolterink, G., Crofton, K., Hougaard Bennekou, S., Paparella, M., Tzoulaki, I., 2021. Development of Integrated Approaches to Testing and Assessment (IATA) case studies on developmental neurotoxicity (DNT) risk assessment. EFSA Journal 19, e06599.

Annex I – ToxTemp NPC1

Author: Stefan Masjosthusmann, Ellen Fritsche, Katharina Koch, Kristina Bartmann

Date: 31JAN2022

Version: 20200702_v1.2

1. Overview

1.1. Descriptive full-text title

Assessment of human neural progenitor cell proliferation (NPC1)

1.2. Abstract

The human developing central nervous system may be more vulnerable to adverse effects of chemical agents than the adult brain. At present, due to the knowledge gap concerning hazard identification for human neurodevelopmental toxicity (DNT), there is an urgent need for testing and subsequent regulation of chemicals for their potential to interfere with the developing nervous system. Primary human neural progenitor cells (hNPCs) cultivated as three-dimensional floating spheres are able to represent several key processes of brain development. In the neural progenitor cell proliferation assay (NPC1), hNPCs are plated in 96 well plates as 3-dimensional spheres and exposed to test compounds. Thereby the process of NPC proliferation can be studied. This DNT-specific endpoint is studied in combination with general cell viability and cytotoxicity. Cortical human NPC proliferation is a critical process during brain development that, if disturbed, may lead to alterations in brain development and cause cognitive dysfunction. Currently, cortical NPC proliferation is one of the many processes, which are assessed in the OECD TG426 by neuropathological evaluation of certain brain regions as well as neurobehavioral tests. According to the readiness criteria as published by Bal-Price et al. (2018), the neural progenitor cell proliferation assay obtained the readiness score A.

Assay summary:

toxicological target	→	developing brain
test system	→	primary human neural progenitor cells (hNPCs) from human cortex (GW16-19)
readout(s)	→	sphere size, DNA synthesis as chemiluminescence measurement, viability and cytotoxicity as fluorescence intensity
biological process(es)	→	fetal NPC proliferation viability, cytotoxicity

(human) adverse outcome(s)	→	cognitive dysfunction
hazard(s)	→	adverse effect on cell proliferation
endpoint of current regulatory studies	→	not directly
validation/evaluation	→	readiness analysis: readiness score A, according to Bal-Price et al. (2018)

2. General information

2.1. Name of test method

Neural progenitor cell proliferation assay (NPC1)

by sphere size (NPC1a)

by BrdU incorporation (NPC1b)

2.2. Version number and date of deposition

20200702_v1.2

2.3. Summary of introduced changes in comparison to previous version(s)

“original version”

2.4. Assigned data base name

NPC1a_DNT_hNPC_prol_72h_20200702v1.2

NPC1b_DNT_hNPC_prol_72h_20200702v1.2

ToxCast invitroDB name:

IUF_NPC1b_proliferation_BrdU_72hr

IUF_NPC1a_proliferation_Area_72hr

2.5. Name and acronym of the test depositor

IUF – Leibniz Research Institute for Environmental Medicine

2.6. Name and email of contact person

Ellen Fritsche: ellen.fritsche@iuf-duesseldorf.de

2.7. Name of further persons involved

Stefan Masjosthusmann: stefan.masjosthusmann@iuf-duesseldorf.de

Kristina Bartmann kristina.bartmann@iuf-duesseldorf.de

Katharina Koch katharina.koch@iuf-duesseldorf.de

2.8. Reference to additional files of relevance

Number of supporting files:

1. Standard Operation Procedure (SOP; Appendix I in Masjosthusman et al. 2020)

3. Description of general features of the test system source

3.1. Supply of source cells

Commercial supplier, Lonza, Verviers, Belgium

3.2. Overview of cell source component(s)

Primary human neural progenitor cells (hNPCs) are provided as cryopreserved 3D neurospheres from Lonza, Verviers, Belgium. The material originates from the human brain cortex of different gestational ages (GW16-19). Sex is either specified or determined before the cells are used.

3.3. Characterization and definition of source cells

$1x10^6$ hNPCs per vial are obtained from Lonza (#PT-2599) and expanded according to the SOP (Appendix I in Masjosthusmann et al. 2020). Lonza provides the cells with a viability of at least 20%. FACS analysis confirmed that proliferating neurospheres express the cell type-specific CNS neural stem cell and progenitor cell markers nestin, SRY-box 2 (SOX2), and Ki67 (Koch et al., 2022). Moreover, proliferating hNPCs react to growth factor stimuli (epidermal growth factor (EGF) and recombinant human fibroblast growth factor (FGF)) with increased proliferation, while simultaneous pharmacological inhibition of the EGF receptor (PD153035) impaired the proliferation increase. Upon transfer of hNPC neurospheres on poly D-lysin/laminin matrix and cultivation in the absence of growth factors (EGF and FGF), the hNPCs differentiate into effector cells expressing markers of neurons (β-III-tubulin), astrocytes (GFAP), radial glia cells (nestin) and oligodendrocytes (O4) (Baumann et al., 2015; Schmuck et al., 2017; Koch et al., 2022).

3.4. Acceptance criteria for source cell population

The following acceptability criteria have been tested at the supplier (Lonza) and are prerequisites for the shipment to customers:

- tested positive for TUBB3 and GFAP after differentiation
- tested free of HIV, HBV and HC
- tested negative both in sterility test and for mycoplasma contamination
- cell count of $1.2x10^6$ cells/mL
- viability of at least 20%
- Adherence of <=50%

The proliferative capacity of Lonza hNPCs was reported previously (Moors et al., 2009; Baumann et al., 2015; Klose et al., 2021a).

3.5. Variability and troubleshooting of source cells

The sphere size at day 0 of cell thawing can be different depending on the donor.

In the 3 to 4-week expansion period different donors can show differences in their proliferative capacity (spheres need longer, 3 instead of 4 weeks, to reach the acceptable min. size of 0.2 – 0.5 µm). After the first mechanical dissociation, there are no observable or measurable inter-individual differences.

Critical consumables

The proliferation medium does not contain serum or serum replacement.

The use of epidermal growth factor (EGF) and recombinant human fibroblast growth factor (FGF) is critical for sphere growth. FGF contains 1% bovine serum albumin and is thus prone to batch effects.

Critical handling

The thawing medium contains DMSO in a concentration that affects cell health which is why thawed cells should quickly be diluted in proliferation medium (30 mL of media for one vial of cells).

It is recommended to add FGF to the proliferation medium directly before thawing.

At the end of week two of the expansion period (see below), the spheres should be transferred to petri dishes coated with poly-(2-hydroxyethyl methacrylate) (poly-Hema) to prevent cell attachment.

Attached cells that are not differentiated can be gently detached using a 1000 µL pipet. To avoid repeated attachment, all cells should be transferred to a new poly-Hema coated petri dish.

Medium containing FGF should not be stored longer than 1 week at 4°C.

During the first two weeks, the medium should be removed using a 1000 µL pipet to keep the accidental removal of small spheres to a minimum. In addition, removed medium should be kept in a new petri dish under culture conditions until the next feeding day, to transfer accidentally removed spheres back to the culture.

The neurospheres should be well distributed in the petri dish to prevent aggregation. This is especially important after mechanical dissociations.

It is important to avoid frequent re-opening of the incubators, to ensure constant CO_2 and temperature levels. Furthermore, the smallest vibrations can lead to aggregations of neurospheres.

The number of passages after thawing influences the proliferation capability of neurospheres. Neurospheres should not be used for the NPC1ab assay after passage 6.

3.6. Differentiation towards the final test system

Cells are frozen in liquid nitrogen and have to be cultivated in proliferation medium at 37°C and 5% CO_2 after thawing. The medium contains Dulbecco's modified Eagle medium and Hams F12 (2:1) supplemented with 2% B27, 20 ng/mL EGF, 20 ng/mL recombinant human FGF, 1% penicillin, and streptomycin. The thawing is performed by repeated addition and removal of proliferation medium to the vial until all cells are transferred to a tissue culture flask containing proliferation medium. The cells are carefully resuspended and distributed to 10 cm petri dishes filled with fresh, prewarmed proliferation medium. The cells are fed by replacing half the medium with new medium every two to three days (Monday, Wednesday, and Friday). At each feeding day, the culture is checked for impurities (e.g. fibers or other debris). Impurities and the removed media are transferred to a new petri dish (waste dish). If spheres are mistakenly sorted out during feeding, they can be rescued and placed back in the original culture dish. After 3-4 weeks, neurospheres reach the acceptable size of 0.2 – 0.5 mm for passaging by mechanical dissociation. Therefore, neurospheres are mechanically dissociated into

pieces of 0.15 - 0.25 mm edge length (depending on the desired sphere size after passaging) using a tissue chopper, which then round-off again to uniform sized neurospheres within 1 day in proliferation medium. By using this method, neurospheres are expanded every week. Starting at week 2, poly-(2-hydroxyethyl methacrylate) (poly-Hema) coated dishes are used for the cultivation procedure.

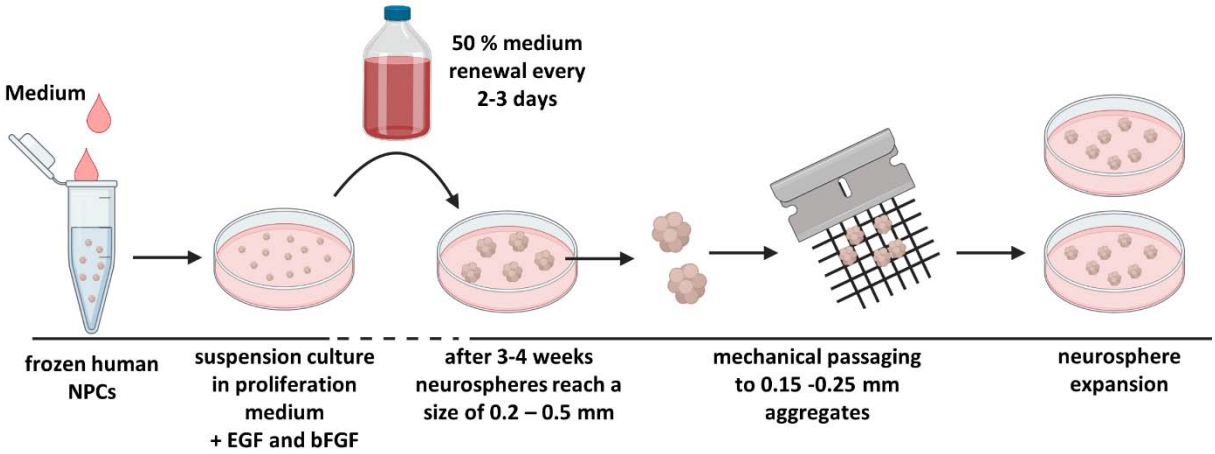


Figure 1 differentiation towards the final test system. *hNPC are thawed by repeated addition and removal of proliferation media. The resuspended cells are distributed to cell culture dishes and cultivated in proliferation media containing EGF and FGF for three to four weeks with 50% media exchange every two to three days. When the spheres reach a size of 0.2-0.5 mm they are expanded by mechanical passaging every 7 days.*

3.7. Reference/link to maintenance culture protocol

see the SOP (Appendix I in Masjosthusmann et al. 2020)

4. Definition of the test system as used in the method

4.1. Principles of the culture protocol

After the cell expansion period, the cells are cultured for up to four weeks in which they are passaged every week as described in 3.6. Between one to three days after passaging, depending on the size chosen for passaging, spheres at a size of 0.3 mm are used in the assay.

For the assessment of neural progenitor cell proliferation, the spheres are plated in poly-Hema coated 96-well U-bottom plates filled with proliferation medium containing growth factors (EGF and FGF). One 0.25 - 0.35 mm big sphere is plated in the middle of each well. Within 3 days NPCs proliferate and grow in size. Cultivation during the test method is performed at 37°C and 5% CO_2 at a pH of 7.2-7.6. As a positive control, spheres are cultivated in absence of growth factors (EGF and FGF), which dramatically reduces proliferation.

4.2. Acceptance criteria for assessing the test system at its start

To be used in the test method, neurospheres have to display a perfectly round shape with no disintegrated borders. One neurosphere with 300 µm in diameter contains around 2.6 x 10^3 cells. Additionally, the basic neurospheres culture is checked for mycoplasma contamination every three

months and controlled for fungal and bacterial contamination by visual inspection at each feeding and plating day.

4.3. Acceptance criteria for the test system at the end of compound exposure

The proliferative capacity of hNPCs is assessed by cultivating them in either medium supplemented with the human growth factors EGF and FGF basic (as described in 4.1) or deprived of them (positive control, PC). Over the 3 days in culture, hNPCs approximately increase their size on average by 33% (Koch et al. 2022; accepted). For this process the following acceptance criteria are defined for the solvent control containing the solvent of the highest test compound concentration (SC; mean of at least two replicates):

Proliferation by area (slope of sphere area)	1000-3000 pixel/day
Proliferation by BrdU (BrdU raw values):	raw values of treatment conditions must not be lower than the positive control
Proliferation by BrdU (BrdU raw values):	raw values of the SC must be significantly higher than the positive control

4.4. Variability of the test system and troubleshooting

Sources of variation:

Selection of spheres: Depending on the researcher and the availability of spheres, the size of selected spheres can differ in a range of 0.25 – 0.35 mm.

Primary hNPCs are a complex multicellular system with a self-organized sphere composition. Due to the complex multicellular and self-organizing nature, the test system is subject to some heterogenicity, which is represented as the biological variability of some of the measured endpoints.The variability for the different endpoints is shown in 8.3 “Test Performance”.

4.5. Metabolic capacity of the test system

Primary hNPC under proliferating and differentiating conditions do not express CYP1A1 and CYP1B1 (Gassmann et al., 2010).

Other metabolic pathways are not characterized.

4.6. Omics characterization of the test system

Proliferating, three day differentiated, and five day differentiated hNPCs were analyzed for changes in their transcriptomic profile. Several key neurodevelopmental processes (migration, neuronal differentiation, glial differentiation) and genes regulating these processes (Bone morphogenetic protein (BMP), Notch and EGF signaling) were identified and characterized on a functional level (Masjosthusmann et al., 2018).

Transcriptomic effects of exposure to 8 flame retardants were analyzed in hNPCs differentiated for five days (Klose et al., 2021)

4.7. Features of the test system that reflect the *in vivo* tissue

hNPCs reflect the following in vivo tissue features:

NPC1 – fetal NPC proliferation (3D, primary cells) → corresponding to in vivo growth during the fetal phase.

Proliferating hNPCs progressively increase in sphere size by on average 33% within three days of proliferation in the presence of EGF and FGF. Moreover, they express the cell type-specific CNS neural stem and progenitor cell markers nestin and SRY-box 2 (SOX2) (Koch et al., 2022). EFGR signaling is indispensable for proper brain development in vivo and increasingly expressed over time (Romano and Bucci, 2020). In line with that, exposure of proliferating hNPCs to the EGFR inhibitor PD153035 impaired the proliferative capacity (Koch et al., 2022).

4.8. Commercial and intellectual property rights aspects of cells

For the source cells, Lonza holds donor consent and legal authorization that provides permission for all research use.

4.9. Reference/link to the culture protocol

see SOP (Appendix I in Masjosthusmann et al. 2020)

5. Test method exposure scheme and endpoints

5.1. Exposure scheme for toxicity testing

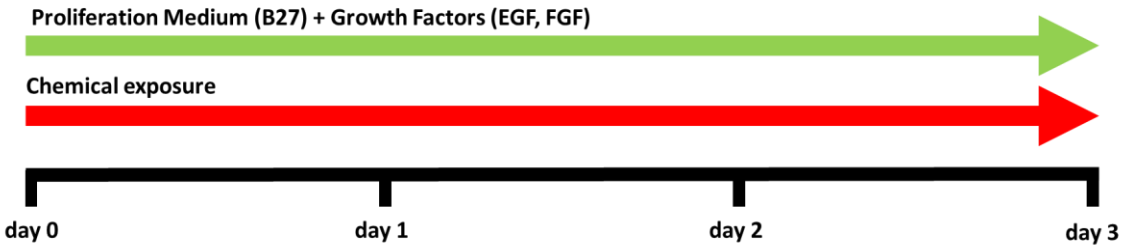


Figure 2 Exposure scheme. *Neurospheres are plated in poly-Hema coated 96-well U-bottom plates containing proliferation medium and are exposed to increasing compound concentrations over a cultivation time of 72 hours.*

0.3 mm big hNPCs are plated as described in 4.1. Cells are plated according to the plating scheme in Figure 4 in the already prepared test solutions. Exposure starts on the plating day (day 0) and is continued over three days until the experiment is terminated (Figure 2).

5.2. Endpoint(s) of the test method

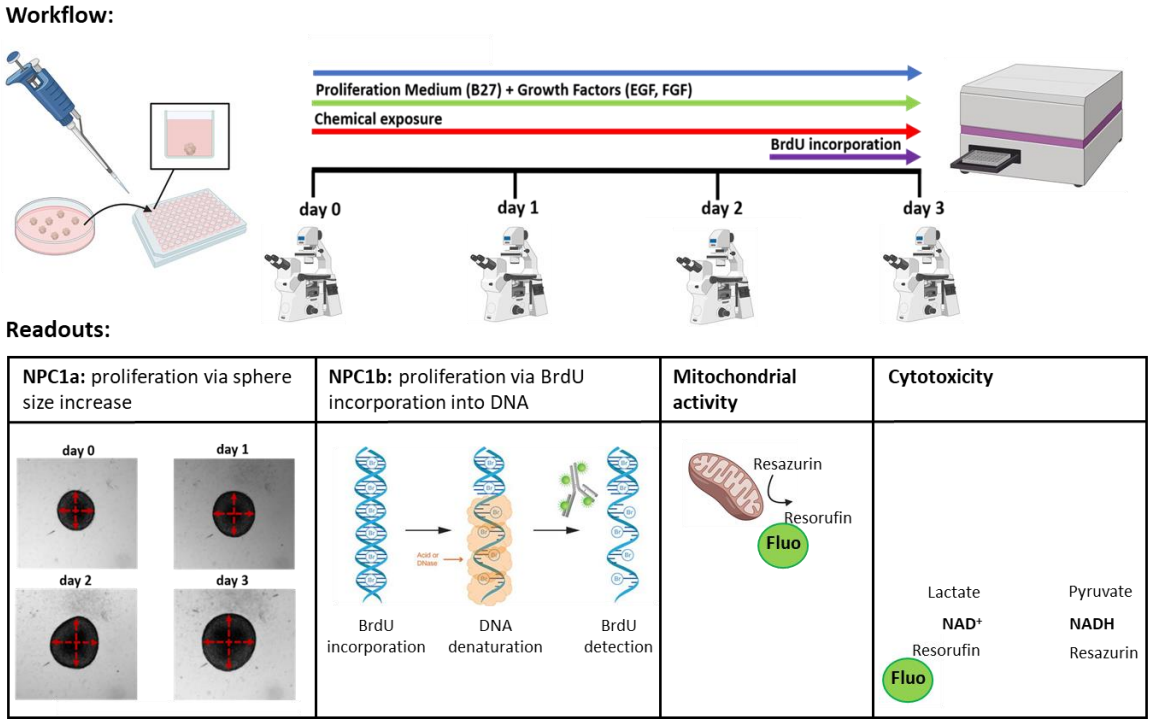


***Figure 3: Endpoint assessment**. Neurospheres are plated in 96-well U-bottom plates and exposed to increasing compound concentrations in proliferation medium over a cultivation period of 72 h. Sphere size is determined every day via brightfield images. The assay is terminated by the assessment of cell viability, cytotoxicity, and proliferation by BrdU.*

Primary DNT specific endpoints of the test method are:

1. proliferation by area (NPC1a)
2. proliferation by BrdU (NPC1b)

Secondary endpoints are:

1. cytotoxicity 72 h
2. viability 72 h

All endpoints are generated from the same experimental run and from each well/sphere in the 96-well plate.

5.3. Overview of analytical method(s) to assess test endpoint(s)

Primary endpoints:

1. Proliferation by area (72h; NPC1a) is assessed as the slope of the increase in sphere size (amount of pixel in the bright-field image, sphere area) over 72 h measured by brightfield microscopy using high content imaging at 0 h, 24 h, 48 h, and 72 h.

2. Proliferation by BrdU (72h; NPC1b) is assessed as BrdU incorporation (as an indirect measure of DNA synthesis) over the last 16 h of compound exposure. It is measured as a luminescence signal (relative luminescence unit) in a multi-plate reader after 72 h.

Secondary endpoints:

1. Cytotoxicity 72 h is assessed as membrane integrity by measuring the amount of LDH leaked from cells with damaged plasma membranes. LDH-dependent reduction of resazurin to resorufin is measured in the supernatant of each well as fluorescence of the reaction product resorufin (relative fluorescence unit) in a multiplate reader after 72 h of compound exposure.
2. Viability 72 h is assessed as mitochondrial activity by measuring the amount of resazurin reduced to fluorescent resorufin (relative fluorescence unit) in a multi-plate reader in the last two hours of the 72 h proliferation and compound exposure period.

5.4. Technical details (of e.g. endpoint measurements)

All technical details for the test method are available in the SOP (Appendix I in Masjosthusmann et al. 2020)

5.5. Endpoint-specific controls/mechanistic control compounds (MCC)

All endpoint-specific controls are run for each experiment (plate).

1. Controls for Primary endpoints:
   hNPC proliferation is diminished by the withdrawal of growth factors (EGF and FGF). Spheres are plated in medium not containing EGF and FGF. This positive control (PC) demonstrates the physiological functionality of the growth factor-dependent regulation of hNPC proliferation. Inhibition of the growth factor-dependent proliferation causes a reduction of proliferation to 0% of the solvent control (SC, see 5.7) for proliferation by area (NPC1a) and 0-40% of the SC for proliferation by BrdU (NPC1b).

2. Controls for Secondary endpoints:
   0.2 % Triton X-100 is used as a positive control for cell viability and cytotoxicity since it lyses the cell and therefore causes a maximal response for both endpoints. This positive control is run on each experimental plate.

5.6. Positive controls

The NPC1ab hNPC proliferation assay correctly identified the following compounds that are known to cause DNT in humans or in vivo (Masjosthusmann et al., 2020):

Cadmium chloride

Dexamethasone

Hexachlorophene

Chlorpromazine hydrochloride

Methylazoxymethanol acetate

all-trans-Retinoic acid

Tributyltin chloride

Sodium valproate

all-trans-Retinoic acid

5.7. Negative and unspecific controls

The solvent control (SC) is used as a negative control that is run on each experimental plate. Each solvent has to be established for its use as a solvent control by comparing the effect of the SC to the effect of medium only. Established solvent controls show the same response as the medium control. The SC is used to assess if the acceptability criteria for NPC1ab proliferation are met and for normalization of the compound exposure and the positive control response.

Established SCs are:

DMSO: 0.3% v/v; 0.2% v/v; 0.1% v/v

DPBS: 2% v/v

dsH_2O: 2% v/v

MeOH: 0.1% v/v

Other negative control compounds that were identified as negative in this assay and are known to not affect neurodevelopmental endpoints *in vivo* include (Masjosthusmann et al., 2020):

Acetaminophen

Amoxicillin

Aspirin

Buspirone

Chlorpheniramine maleate

D-Glucitol

Diethylene glycol

D-Mannitol

Doxylamine succinate

Famotidine

Ibuprofen

Metformin

Metoprolol

Penicillin VK

Saccharin

Sodium benzoate

Warfarin

5.8. Features relevant for cytotoxicity testing

Cytotoxicity and cell viability are assessed for each sphere plated in the assay.

5.9. Acceptance criteria for the test method

General acceptance criteria:

1. At least two replicate values need to be present for each condition (concentration) to be accepted for the data analysis.
2. At least five conditions and the solvent control need to be present for the experiment to be accepted in the data analysis for concentration response modeling.

5.10. Throughput estimate

The methods described here are set up in a 96-well plate format with automated image acquisition, analysis, and data evaluation. Pipetting steps such as compound dilutions, as well as the viability and cytotoxicity assays can be automated using a liquid handling system.

In the fully automated set up, 10 plates with 8 conditions (Figure 4) and 4-5 replicates per condition can be run in one week by two laboratory technicians. This results in the generation of 400 data points for each endpoint within one week (excluding all controls). The throughput is therefore estimated as medium.

6. Handling details of the test method

6.1. Preparation/addition of test compounds

The method is set up for 8 test conditions including 7 compound concentrations and one SC. The test conditions are prepared in a serial dilution from the stock solution (Figure 4).

Stock solutions are prepared by diluting the compound in the solvent (e.g. DMSO) in a concentration that allows the preparation of the highest test concentration without exceeding the highest acceptable solvent concentration (see 5.7). For DMSO the highest acceptable solvent concentration is 0.3% which means that the stock concentration needs to be at least 1000x higher than the highest test concentration.

Stock solutions in non-sterile solvents (e.g. water or PBS) have to be sterile filtrated using a sterile syringe filter (Ø 0.2 µm). Adsorption to the filter needs to be considered.

Stock solutions are aliquoted and stored at -20°C. A stock solution is not thawed more than three times.

For the preparation of the test condition, the stock solution is diluted to the highest test concentration (default 1:1000) in proliferation medium. All following dilutions are prepared by serial dilution of the highest concentration in proliferation medium with solvent (in the concentration of the highest test concentration). The default serial dilution is 1:3 which covers a concentration range from e.g. 20 µM to 27 nM (729-fold). Depending on the desired concentration range, the dilution can be adjusted to 1:2, 1:5, 1:10, or other.

The SC is prepared by adding the solvent to proliferation medium in the same concentration as the highest test concentration.

100 µL of the compound dilutions and the SC are added to each well of a 96 well plate (Figure 4). Alternatively, the serial dilution can be prepared directly in the 96 well plates.

One hNPC sphere is added to each well after the medium equilibrated for 15 to 30 min at 37°C and 5% CO_2.

To subtract the background fluorescence of the phenol red-containing medium, 5 wells with medium only (Background, Figure 5) are prepared for the mitochondrial activity and cytotoxicity assays.

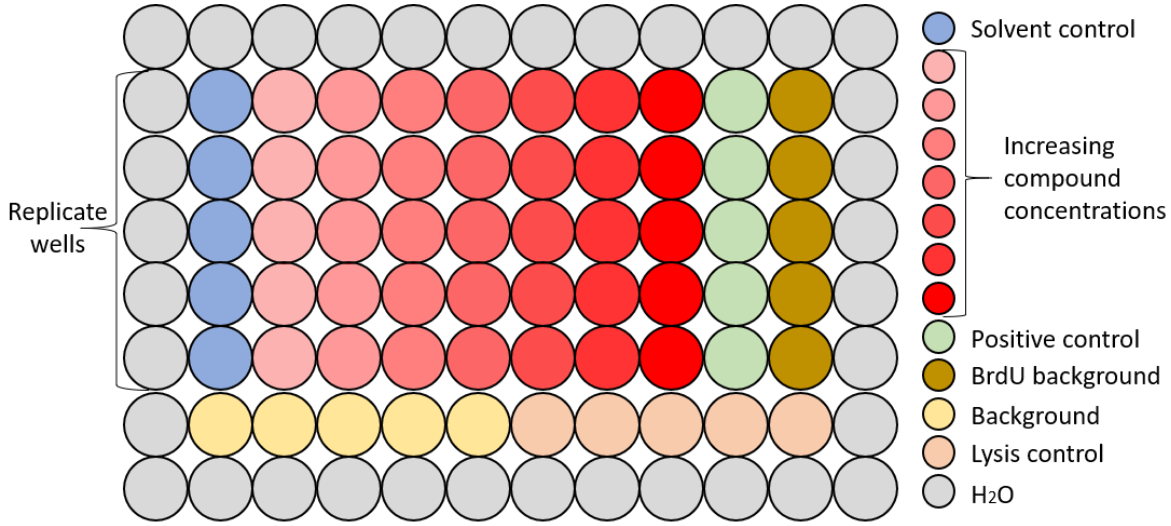


Figure 4 Plating Scheme. *The type of solvent control depends on the solvent of the compound that was tested. 7 compound concentrations are plated in a serial dilution from lowest (left) to highest (right) concentration. The positive control for cell proliferation is proliferation medium without EGF and FGF. BrdU background is used for the proliferation assay by BrdU (spheres in SC medium). Background and lysis control are used for cell viability and cytotoxicity assay.*

6.2. Day-to-day documentation of test execution

Documentation for each experiment including meta data and the experimental data that is collected in the Automated Experimental Evaluation or AXES sheet.

Meta data such as plating date, experimenter, NPC individual, NPC passage, compound, compound concentrations and a plate map are reported in these sheets. Depending on the endpoint, the experimental data is collected during or at the end of the experiment. Each raw data point (including all outliers) is collected in the AXES sheet. All deviations from the standard procedure are documented in a comment section of the AXES sheet.

6.3. Practical phase of test compound exposure

The practical phase of the test compound exposure follows the description in the SOP (Appendix I in Masjosthusmann et al. 2020). Deviation from the SOP are documented in the comment section of the AXES sheet.

Errors (e.g. pipetting in wrong well or wrong volume pipetted) are also documented in the comment section of the AXES sheets. Data points of the affected well are marked in the AXES sheet and excluded from the analysis.

6.4. Concentration settings

Starting concentrations and concentration ranges are defined based on the following factors:

- toxicological relevance of the compound (i.e. internal human exposures, effects at lowest concentrations)
- solubility of the compound
- highest useable solvent concentration

These factors are determined based on available information (databases/literature) or experimentally (e.g. solubility test).

6.5. Uncertainties and troubleshooting

Problematic compounds:

- Volatile compounds
- High lipophilicity (high K_{OW})
- Low solubility in established solvents
- Fluorescent compounds (possible interfere with viability and cytotoxicity assay)

Critical handling steps:

- If different plate types are used, the test system and method need to be re-established.
- Outer wells have to be filled with H_2O because of edge effects.
- Automating the pipetting steps using a liquid handling system for coating, preparation of the plates, viability and cytotoxicity assays reduces the variability and the user bias.

Sources of variation:

- Pipetting steps: Each pipetting step is a source of variation. Especially in the viability and cytotoxicity assay where the volume pipetted determines the final readout.

6.6. Detailed protocol (SOP)

See SOP (Appendix I in Masjosthusmann et al. 2020).

6.7. Special instrumentation

- Incubator for cell culture
- Tissue chopper
- Multiplate reader for fluorescence and luminescence measurement
- Bright-field microscopy.
- Liquid handling system (necessary to achieve the throughput described above)
- Hair dryer

6.8. Possible variations

There are no established variations of the assay.

6.9. Cross-reference to related test methods

The hNP1 assay (CCTE_Mundy_HCI_hNP1_Pro assay) uses proliferating neural stem

cells derived from a neuroepithelial cell lineage of WA09 human embryonic stem cells. The assay measures cell proliferation using BrdU labeling in combination with a immunocytochemical staining and high content imaging. In contrast to the NPC1 assay cells are plated as 2D monolayer instead of the free-floating 3D spheroids used in the NPC1 assay.

7. Data management

7.1. Raw data format

The raw data format is different depending on the endpoints.

For all endpoints assessed in a multiplate reader (viability, cytotoxicity, BrdU incorporation) the raw data format are excel files containing values (one for each endpoint, timepoint and well) measured as relative fluorescence/luminescence units. These values are transferred from the original excel file into the AXES sheet. The original excel output file is saved for traceability of the data.

The sphere size is automatically measured in the Cellomics scan software (Version 6.6.0; Thermo Scientific) and copied into the AXES sheet. Original brightfield images are archived for 10 years.

7.2. Outliers

Mathematical procedures to define outliers are not applied. Data points from wells where technical problems are known or obvious are excluded from the analysis.

Possible technical problems:

- pipetting errors
- spillover from lysis
- problems in fixation of singularized cells

All outliers are marked in the AXES sheet.

7.3. Raw data processing to summary data

If not otherwise stated, all data processing steps are performed in an R based evaluation tool that was designed for data processing, curve fitting and point of departure evaluation of in vitro concentration response toxicity data.

Data processing describes all processing steps of raw data that are necessary to obtain the final response values including the normalization, curve fitting and benchmark concentration calculation.

Processing (or pre-processing) steps depend on the endpoint and are described below:

Proliferation by BrdU: subtraction of mean BrdU background from each raw response value.

$$Background\ corrected\ response\ [RLU] = raw\ response\ [RLU] - Background\ BrdU\ [RLU]$$

Proliferation by area: slope of the sphere size over 3 days of proliferation (d0, d1, d2, d3). The calculated slope is used as raw data input for the DB and is thus not calculated in the R based evaluation tool.

Viability: subtraction of mean background from each response value.

$$Background\ corrected\ response\ [RFU] = raw\ response\ [RFU] - Background\ [RFU]$$

Cytotoxicity: no pre-processing

7.4. Curve fitting

The data is normalized to the SC and re-normalized to the starting point of the curve.

For the normalization to the SC each replicate data point is normalized to the median of the SC in the respective experiment.

For the cytotoxicity assays the following normalization is used instead of the normalization to the SC. Here again each response value is normalized using the median of the lysis control and the median of the solvent control.

$$\text{normalized response} = \frac{\text{lysis control} - \text{response}}{\text{lysis control} - \text{solvent control}}$$

The R package drc is used to calculate the optimal fit for each experiment. For calculations of curve fits and BMCs, the data from independent experiments is pooled (median of all replicate values for one concentration). Several non-linear models are run with the concentration response data of each endpoint and the Akaike´s information criteria is used to determine the best fit.

For re-normalization of the data, the response value of the curves starting point is determined and used to re-normalize all response values. Therefore, each mean response value is divided by the starting point of the curve and multiplied with 100. For the re-normalized response values the curve fitting is repeated to produce the final concentration response curve.

For deriving a reference point (RP) or point of departure (Pod) the Benchmark Concentration (BMC) approach as recommended by the EFSA Scientific Committee (Hardy et al., 2017) is applied. The BMC approach makes use of all data points that define the fitted concentration response curve. Thereby, the BMC is defined as the concentration that is associated with a specific change in response, the Benchmark Response (BMR). The BMR is a value of effect size and should be defined as an effect size that is higher than the general variability of the measured endpoint. The BMR is therefore determined based on the variability of the respective endpoint.

BMR for NPC1:

proliferation by area	BMR30
proliferation by BrdU	BMR30
cytotoxicity 72 h	BMR10
viability 72 h	BMR30

Based on the BMR and the concentration response curve, the evaluation tool calculates the BMC, as well as upper and lower confidence limits (BMCU and BMCL respectively) based on the predict function in the R package drc. The predict function calculates the prediction bands around the concentration response curve based on the deviation between independent experiments and gives an estimation of the area that is expected to enclose 95% of future data points. The BMCL is thereby defined as the intersection of the lower band and the BMR while the BMCU is defined as the intersection of the upper band and the BMR. The confidence intervals are used to access the uncertainty of the BMC. If the BMCU is 1.5 times above the test range, the original BMCU is replaced by 1.5x the highest tested concentration.

7.5. Internal data storage

All raw data is stored on a server with a daily server back up for at least 10 years.

7.6. Metadata

All metadata is collected in the AXES sheet (see 6.2) together with all raw data.

The metadata gives information on:

The experiment:

- start and end date of the experiment
- experimenter

The cell source:

- human individual
- cells thawing date
- passage of cells
- date of cell passaging

The compound:

- compound identity
- stock concentration
- all dilution steps
- solvent and solvent concentration

The controls:

- control identity
- preparation of controls

7.7. Metadata file format

All metadata is collected in an Excel format.

8. Prediction model and toxicological application

8.1. Scientific principle, test purpose and relevance

Primary hNPCs are isolated from the fetal brain cortices and can be used to measure proliferation, a process of brain growth during the fetal phase of prenatal development.

The test system therefore measures adverse events in the young (fetal) developing brain.

Different types of NPC exist in the developing brain. Besides ventricular zone NPC, radial glia cells serve as cortical progenitor cells responsible for cortical expansion and folding. As whole cortices were used for cell preparation, this is not a specific NPC type but rather a mix of NPCs found in fetal human cortex during development.

The toxicological events that are modeled concern events that influence proliferation of NPCs found in human cortex during the fetal phase.

8.2. Prediction model

Two prediction models (PM) are applied for the NPC1 assay. One PM for a downregulation (PM downregulation) and one PM for an upregulation (PM upregulation) in cell proliferation.

PM downregulation

The PM is based on a hit definition followed by comparison of the confidence intervals (CI) for the BMC of the DNT-specific endpoint (BMCs) and the unspecific endpoint (cytotoxicity/viability; BMCus).

Thereby the following four hit classifications apply:

“no hit”	The compound is not defined as a hit.
“specific hit”:	The compound is defined as hit and the CI’s do not overlap, meaning that the upper confidence limit of the specific endpoint (BMCUs) is lower than the lower confidence limit of the unspecific endpoint (BMCLus).
“borderline hit”:	The compound is defined as hit and the CI of the specific endpoint overlaps by less than, or equal to 10% with the CI of the unspecific endpoint.
“unspecific hit”	The compound is defined as hit and the CI of the specific endpoint overlaps by more than 10% with the CI of the unspecific endpoint.

The compound is classified as a hit, if the concentration response curve generates a BMC and if the CI is within the test range. In case the CI spans above the test range, the compound is only classified as a hit, if the highest test concentration is significantly different from the lowest test concentration. The adjusted significance ($p<0.05$) is thereby determined using a Tukey HSD test.

The decision process for the prediction model is described in the flow chart in Figure 5.

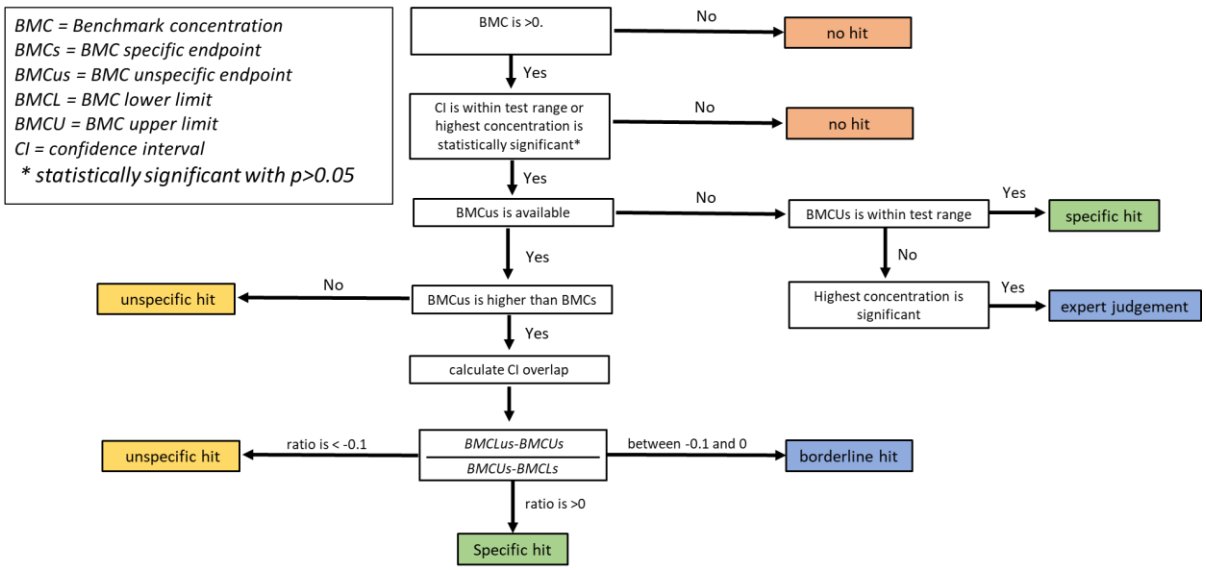


Figure 5: Decision tree for the PM for down regulation applied for the test method NPC1. Overview of the decisions leading to the classification of a compound in one of four categories: "no hit", "specific hit", "borderline hit" and "unspecific hit".

Specific consideration for the prediction model:

In case no confidence limits are available for the unspecific endpoint because the BMR is not reached, the BMCLus is assumed to be the highest tested concentration. If the CI of the specific endpoint additional spans above the test range for a compound identified as "hit" based on statistical significance, expert judgement is applied to define if the "hit" is specific or unspecific. Therefore, all datapoints of the unspecific and specific endpoint of the highest test concentration are compared. If these data points do not overlap, the compound is classified as "specific hit" otherwise as "unspecific hit".

In general, BMCs based on the same BMR are compared (e.g. BMC_{30us} vs BMC_{30s}). In case the BMC_{30us} is not available for an endpoint that allows the generation of a BMC_{10us}, the BMC_{10us} is used instead. If the classification of this comparison is "unspecific hit", the compound will be flagged as "check manually" as the BMC_{10us} is lower than the BMC_{30us} leading to a higher probability of a false classification. To avoid such false classifications, expert judgment is needed.

Compounds can also be flagged as "check manually", if the classification, based on the viability is different from the classification based on the cytotoxicity and if the confidence interval is very wide (BMCU/BMCL > 25), which means a high uncertainty for the BMC estimation. In both cases expert judgment is needed to decide on the classification.

The expert judgement is an individual decision process that accounts for effect size, curve progression, statistical significance and overall standard deviation. If the concentration response curves do not give enough information for decision by expert judgement, additional testing or testing in a different concentration range should be performed.

PM upregulation

In contrast to the PM for downregulation, the PM for upregulation is based on a hit definition without the comparison of confidence intervals (CI) between the specific and unspecific endpoint. The reason is, that specific and unspecific endpoints do not have the same relationship during an induction,

compared to a reduction in the endpoint. A loss in general cell health will likely result in an effect in cell proliferation, while an induction in cell health (measured as mitochondrial activity) does not necessarily increase cell proliferation.

The following three hit classifications apply:

"no hit"	The compound is not defined as hit.
"specific hit":	The compound is defined as hit and the effect is no artifact due to loss in cell health.
"unspecific hit"	The compound is defined as hit in only the unspecific endpoints.

The compound is classified as hit, if the concentrations response curve generates a BMC and if the CI is within the test range. In case the CI spans above the test range, the compound is only classified as hit, if the highest test concentration is significantly different from the lowest test concentration. The adjusted significance ($p<0.05$) is thereby determined using a Tukey HSD test.

The decision process for the prediction model is described in the flow chart in Figure 6.

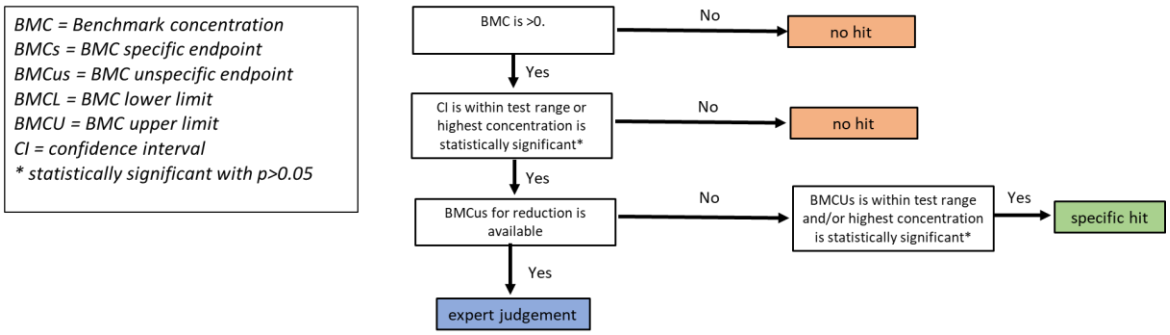


Figure 6: Decision tree for the PM for upregulation applied for the test method NPC1. Overview of the decisions leading to the classification of a compound in one of three categories: "no hit", "specific hit", and "unspecific hit".

In case a compound is classified as hit together with a reduction in an unspecific endpoint, it needs to be clarified if the induction is a specific effect or an artifact, due to a loss in cell viability or an increase in cytotoxicity. Therefore, expert judgement is applied, which accounts for effect size, curve progression, statistical significance, overall standard deviation but also morphological changes of the spheres. Expert judgement is additionally applied, if the confidence interval is very wide (BMCU/BMCL > 25) which means a high uncertainty for the BMC estimation.

8.3. Prediction model setup

The prediction model is set up as a statistical model which uses the 95 % confidence intervals (assessed based on the prediction bands around the concentration response curve) to determine the uncertainty of a hit definition. In the case of high uncertainty in the confidence interval, (e.g. because the CI spans above the tested concentration range) the model additionally considers the statistical significance between the highest and the lowest test concentrations. Next to the hit definition, the uncertainty given by the 95 % CI is also considered in the specificity analyses (see 8.2 for a more detailed description).

The model has been tested on a set of 17 DNT negative- and 9 DNT positive compounds (see 5.6 and 5.7). All negative compounds were correctly classified as “no hit” and 3 positive compounds were correctly classified as “specific hit” for DNT. Here it is important to mention, that it is not expected that the NPC1 assay identifies all DNT positive compounds as not all of those compounds act via the mode-of-action (MoA) ‘cell proliferation’ on the developing brain. For most compounds the exact MoA for their neurodevelopmental adversity is not precisely known, yet other mechanisms like neuron/glia differentiation, neurite outgrowth, synaptogenesis and neuronal network formation are amongst the known DNT MoAs.

8.4. Test Performance

The following parameters were assessed to quantify the assay variability:

Intra-experimental variation (SC) is the mean CV±SD of the CV of all replicates of the solvent control from one experiment across all (n>400) experiments.

Inter-experimental variation (raw) is the variability across all independent experiments (n>400) before normalization.

Inter-experimental variation (low conc.) is the variability across all independent experiments (n>400) after normalization based on the response of the lowest test concentration. It is assumed that the lowest test concentration does not affect any of the endpoints measured.

Inter-experimental variation (positive controls) is the variability of the positive control across all independent experiments (n>40) after normalization.

Positive control:

proliferation media without growth factors (EGF, FGF)

Table 1 summarizes the assay performance in terms of variability of each endpoint in the assay.

Table 1: Assay variability quantified as coefficient of variance (CV).

Endpoint	Intra-experimental variation (SC)	inter-experimental variation (raw)	Inter-experimental variation (low con.)	Inter-experimental variation (positive controls)
proliferation by area (NPC1a)	19.4 ±9.4 %	34.3%	21.4%	1.1%
proliferation by BrdU (NPC1b)	13.2 ±5.9 %	65.4%	28.1%	10.3%
cytotoxicity [72h]	1.3 ±1.6 %	37.7%	7.9%	
viability [72h]	5.8 ±3.5 %	11.6%	9.1%	

Sensitivity and specificity of the NPC1 assay are determined based on a set of 9 predicted human DNT positive compounds and 17 predicted human DNT negative compounds (Masjosthusmann et al., 2020).

Based on this compound set the following performance parameters are obtained for the NPC1 assay.

Specificity: 100 %

Sensitivity: 33%

Here it is important to mention, that it is not expected that the NPC1 assay identifies all DNT positive compounds as not all of those compounds act via the mode-of-action (MoA) 'cell proliferation' on the developing brain. For most compounds the exact MoA for their neurodevelopmental adversity is not precisely known, yet other mechanisms like neuron/glia differentiation, neurite outgrowth, synaptogenesis and neuronal network formation are amongst the known DNT MoAs. It is therefore recommended that this assay is run as one part of an in vitro DNT battery.

8.5. In vitro – in vivo extrapolation (IVIVE)

Parameters for in vitro – in vivo extrapolation are not yet determined.

8.6. Applicability of test method

Toxicological applicability domain

The following compound classes have been tested successfully:

- Industrial chemicals
- pesticides and biocides
- cosmetics ingredients
- pharmaceuticals

Compounds need to be soluble in a solvent at a solubility where the solvent does not produce effects by itself in the test systems (see 5.7 for established solvents).

Compounds that are volatile or have a high lipophilicity have not been tested and might need more sophisticated exposure methods such as 'passive dosing'.

Biological applicability domain

Neural progenitor cell proliferation is based on primary hNPC obtained from the fetal human cortex. As mentioned in 8.1 "Scientific principle" The method represents NPC proliferation during the fetal period.

Next to the endpoints represented by this test method, there are several other necessary neurodevelopmental endpoints which need to be studied using other test methods.

Neurodevelopmental processes not represented by this test method:

Neural Crest Cell (NCC) Migration

NPC apoptosis

Neuronal migration

Oligodendrocyte migration

Radial glia migration

Neuronal differentiation

Oligodendrocyte differentiation

Neuronal morphology

Synaptogenesis

Neuronal network formation

Neural Rosette Formation

hiPSC-derived NPC proliferation

hiPSC-NPC neuronal differentiation

Neuronal subtype differentiation

Astrocyte Differentiation and Maturation

Astrocyte Reactivity

Microglia reactivity

Myelination

For a complete assessment of developmental neurotoxicity, the test method needs to be part of test battery.

The information on signaling pathways modulating the neurodevelopmental endpoints of the test method is summarized in Table 2. This describes the so far tested biological application domain of the assay.

Table 2: Signaling pathways studied in the test method.

DNT Assay	Signaling Pathway	Model Compound	
NPC1	EGFR	lack of EGF, EGFR antagonist	Masjosthusmann et al., 2018
	COX-2	Celecoxib	unpublished
	PDGFR	CP-673451	
	mTOR	Everolimus, MHY1485	
	CREB	KG-501	
	ETC complex I	MPP+	Masjosthusmann et al., 2020
	RAR	all-*trans* retinoic acid	
	GR	Dexamethasone	
	RHO	Narciclasine	
	PKC	Bis-I	

8.7. Incorporation in test battery

To assess the hazard for developmental neurotoxicity it is recommended that this assay is used as one assay in a battery of assays (see 8.5 “Applicability of test methods”) For the assessment of chemical action on the endpoints represented by this test method, the test method can be used as stand-alone test method.The test method is currently used in the set-up of a DNT test battery.

9. Publication/validation status

9.1. Availability of key publications

Key Publications concerning the test method are:

Koch et al. 2022

Klose et al., 2021

Masjosthusmann et al., 2020

Nimtz et al., 2019

Masjosthusmann et al., 2018

Baumann et al., 2016

Baumann et al., 2016

Baumann et al., 2014

Fritsche et al., 2011

Moors et al., 2009

9.2. (Potential) linkage to AOPs

No AOP linkage.

9.3. Steps towards mechanistic validation

See:

3.3 Characterization and definition of source cells

4.6 Omics characterization of the test system

4.7 Features of the test system that reflect the in vivo tissue

8.5 Applicability of test method

9.4. Pre-validation or validation

To date, 123 unique compounds (as defined by unique DTXSIDs) have been tested successfully in this assay.

No formal OECD 34 validation study has been done (eg., ring trials with a standard set of known positive and negative controls).

9.5. Linkage to (e.g. OECD) guidelines/regulatory use

Test is not linked to regulatory guidelines.

10. Test method transferability

10.1. Operator training

For operators with a basic training in cell culture practices a four-week training period for handling of the test system and training in the assay is recommended. The operators should have basic understating in image analysis and data evaluation with respect to concentration response fitting.

10.2. Transfer

The test method has been used by multiple operators over a period of 18 month. However, inter operator variability has not been determined.

11. Safety, ethics and specific requirements

11.1. Specific hazards; issues of waste disposal

No specific requirements.

11.2. Safety data sheet (SDS)

Reference to MSDS is given in the SOP (Appendix I in Masjosthusmann et al. 2020)

11.3. Specific facilities/licenses

No specific facilities are required. No specific ethical approval is required.

11.4. Commercial aspects/intellectual property of material/procedures

There are no commercial aspects or intellectual properties to be considered.

12. References

Bal-Price, A., Hogberg, H.T., Crofton, K.M., Daneshian, M., FitzGerald, R.E., Fritsche, E., Heinonen, T., Hougaard Bennekou, S., Klima, S., Piersma, A.H., Sachana, M., Shafer, T.J., Terron, A., Monnet-Tschudi, F., Viviani, B., Waldmann, T., Westerink, R.H.S.S., Wilks, M.F., Witters, H., Zurich, M.-G.G., Leist, M., 2018. Recommendation on test readiness criteria for new approach methods in toxicology: Exemplified for developmental neurotoxicity. ALTEX 35, 306–352. https://doi.org/10.14573/altex.1712081

Baumann, J., Barenys, M., Gassmann, K., Fritsche, E., 2014. Comparative human and rat "neurosphere assay" for developmental neurotoxicity testing. Curr. Protoc. Toxicol. 1, 1–24. https://doi.org/10.1002/0471140856.tx1221s59

Baumann, Jenny, Dach, K., Barenys, M., Giersiefer, S., Goniwiecha, J., Lein, P.J., Fritsche, E., 2016. Application of the Neurosphere Assay for DNT Hazard Assessment: Challenges and Limitations. Methods Pharmacol. Toxicol. 1–29. https://doi.org/10.1007/7653

Baumann, J., Gassmann, K., Masjosthusmann, S., DeBoer, D., Bendt, F., Giersiefer, S., Fritsche, E., 2016. Comparative human and rat neurospheres reveal species differences in chemical effects on neurodevelopmental key events. Arch. Toxicol. 90. https://doi.org/10.1007/s00204-015-1568-8

Fritsche, E., Gassmann, K., Schreiber, T., 2011. Neurospheres as a Model for Developmental

Neurotoxicity Testing, in: Costa, L.G., Giordano, G., Guizzetti, M. (Eds.), In Vitro Neurotoxicology: Methods and Protocols. Humana Press, Totowa, NJ, pp. 99–114. https://doi.org/10.1007/978-1-61779-170-3_7

Gassmann, K., Abel, J., Bothe, H., Haarmann-Stemmann, T., Merk, H.F., Quasthoff, K.N., Rockel, T.D., Schreiber, T., Fritsche, E., 2010. Species-specific differential ahr expression protects human neural progenitor cells against developmental neurotoxicity of PAHs. Environ. Health Perspect. 118, 1571–1577. https://doi.org/10.1289/ehp.0901545

Hardy, A., Benford, D., Halldorsson, T., Jeger, M.J., Knutsen, K.H., More, S., Mortensen, A., Naegeli, H., Noteborn, H., Ockleford, C., Ricci, A., Rychen, G., Silano, V., Solecki, R., Turck, D., Aerts, M., Bodin, L., Davis, A., Edler, L., Gundert-Remy, U., Sand, S., Slob, W., Bottex, B., Abrahantes, J.C., Marques, D.C., Kass, G., Schlatter, J.R., 2017. Update: use of the benchmark dose approach in risk assessment. EFSA J. 15, 1–41. https://doi.org/10.2903/j.efsa.2017.4658

Klose, J., Pahl, M., Bartmann, K., Bendt, F., Blum, J., Dolde, X., Förster, N., Holzer, A.K., Hübenthal, U., Keßel, H.E., Koch, K., Masjosthusmann, S., Schneider, S., Stürzl, L.C., Woeste, S., Rossi, A., Covaci, A., Behl, M., Leist, M., Tigges, J., Fritsche, E., 2021. Neurodevelopmental toxicity assessment of flame retardants using a human DNT in vitro testing battery. Cell Biol. Toxicol. https://doi.org/10.1007/S10565-021-09603-2

Masjosthusmann, S., Becker, D., Petzuch, B., Klose, J., Siebert, C., Deenen, R., Barenys, M., Baumann, J., Dach, K., Tigges, J., Hübenthal, U., Köhrer, K., Fritsche, E., 2018. A transcriptome comparison of time-matched developing human, mouse and rat neural progenitor cells reveals human uniqueness. Toxicol. Appl. Pharmacol. https://doi.org/10.1016/j.taap.2018.05.009

Masjosthusmann, S., Blum, J., Bartmann, K., Dolde, X., Holzer, A., Stürzl, L., Keßel, E.H., Förster, N., Dönmez, A., Klose, J., Pahl, M., Waldmann, T., Bendt, F., Kisitu, J., Suciu, I., Hübenthal, U., Mosig, A., Leist, M., Fritsche, E., 2020. Establishment of an a priori protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of developmental neurotoxicity. EFSA Support. Publ. 17. https://doi.org/10.2903/sp.efsa.2020.en-1938

Moors, M., Rockel, T.D., Abel, J., Cline, J.E., Gassmann, K., Schreiber, T., Shuwald, J., Weinmann, N., Fritsche, E., 2009. Human neurospheres as three-dimensional cellular systems for developmental neurotoxicity testing. Environ. Health Perspect. 117, 1131–1138. https://doi.org/10.1289/ehp.0800207

Nimtz, L., Klose, J., Masjosthusmann, S., Barenys, M., Fritsche, E., 2019. The neurosphere assay as an in vitro method for developmental neurotoxicity (DNT) evaluation, in: Neuromethods. Humana Press Inc., pp. 141–168. https://doi.org/10.1007/978-1-4939-9228-7_8

Schmuck, M.R., Temme, T., Dach, K., de Boer, D., Barenys, M., Bendt, F., Mosig, A., Fritsche, E., 2017. Omnisphero: a high-content image analysis (HCA) approach for phenotypic developmental neurotoxicity (DNT) screenings of organoid neurosphere cultures in vitro. Arch. Toxicol. 91, 2017–2028. https://doi.org/10.1007/s00204-016-1852-2

Annex II – ToxTemp NPC2-5

Author: Stefan Masjosthusmann, Ellen Fritsche; Katharina Koch, Kristina Bartmann

Date: 31JAN2021

Version: 20200702_v1.2

1. Overview

1.1. Descriptive full-text title

Assessment of human neural progenitor cell migration and differentiation (NPC2-5)

1.2. Abstract

The human developing central nervous system may be more vulnerable to the adverse effects of chemical agents than the adult brain. At present, due to the knowledge gap concerning hazard identification for human neurodevelopmental toxicity (DNT), there is an urgent need for testing and subsequent regulation of chemicals for their potential to interfere with the developing nervous system. Primary human neural progenitor cells (hNPCs) cultivated as three-dimensional floating spheres are able to represent several key processes during brain development. In the neural progenitor cell migration and differentiation assay (NPC2-5), hNPCs are plated on an extracellular matrix, and migrate and differentiate out of the sphere core. Thereby the processes radial glia migration, migration of neurons and oligodendrocytes as well as differentiation into neurons and oligodendrocytes and neurite outgrowth can be studied. Those DNT-specific endpoints are studied in combination with general cell viability and cytotoxicity. Cell migration and differentiation are critical processes during brain development that, if disturbed, lead to alterations in brain development and may cause cognitive dysfunction. Currently, cortical NPC migration- and differentiation-related processes are some of the many processes, which are assessed in the OECD TG426 by neuropathological evaluation of certain brain regions as well as neurobehavioral tests. According to the readiness criteria as published by Bal-Price et al. (2018), the neural progenitor cell migration and differentiation assay obtained readiness scores between A and B depending on the endpoint.

Assay summary:

toxicological target	→	developing brain
test system	→	primary human neural progenitor cells (hNPCs) from human cortex (GW16-19)
readout(s)	→	migration distance, cell number (all cells) number of neurons/oligodendrocytes, neurite length, neurite area, fluorescence intensity
biological process(es)	→	radial glia/neuronal/oligodendrocyte migration, neuronal/oligodendrocyte differentiation,

		neuronal morphology, viability, cytotoxicity
(human) adverse outcome(s)	→	cognitive dysfunction
hazard(s)	→	adverse effect on cell migration and differentiation
endpoint of current regulatory studies	→	not directly
validation/evaluation	→	readiness analysis: readiness score A and B (depending on the endpoint), according to Bal-Price et al. (2018)

2. General information

2.1. Name of test method

Neural progenitor cell migration and differentiation assay (NPC2-5)

radial glia migration (NPC2a)

neuronal migration (NPC2b)

oligodendrocyte migration (NPC2c)

neuronal differentiation (NPC3)

neuronal morphology (NPC4)

oligodendrocyte differentiation (NPC5)

2.2. Version number and date of deposition

20200317_v1.2

2.3. Summary of introduced changes in comparison to previous version(s)

“original version”

2.4. Assigned data base name

NPC2a_DNT_hNPC_mig_72h_20200317v1.1

NPC2a_DNT_hNPC_mig_120h_20200317v1.1

NPC2b_DNT_hNPC_mig_120h_20200317v1.1

NPC2c_DNT_hNPC_mig_120h_20200317v1.1

NPC3_DNT_hNPC_diff_120h_20200317v1.1

NPC4_DNT_hNPC_diff_120h_20200317v1.1

NPC5_DNT_hNPC_diff_120h_20200317v1.1

ToxCast invitroDB name:

IUF_NPC2a_radial_glia_migration_72hr

IUF_NPC2a_ radial_glia_migration_120hr

IUF_NPC2b_neuronal_migration_120hr

IUF_NPC2c_oligodendrocyte_migration_120hr

IUF_NPC3_neuronal_differentaition_120hr

IUF_NPC4_neurite_length_120hr

IUF_NPC4_neurite_area_120hr

IUF_NPC5_oligodendrocyte_differentiation_120hr

2.5. Name and acronym of the test depositor

IUF – Leibniz Research Institute for Environmental Medicine

2.6. Name and email of contact person

Ellen Fritsche ellen.fritsche@iuf-duesseldorf.de

2.7. Name of further persons involved

Stefan Masjosthusmann stefan.masjosthusmann@iuf-duesseldorf.de

Kristina Bartmann kristina.bartmann@iuf-duesseldorf.de

Katharina Koch katharina.koch@iuf-duesseldorf.de

2.8. Reference to additional files of relevance

Number of supporting files:

2. Standard Operation Procedure (Appendix I)

3. Description of general features of the test system source

3.1. Supply of source cells

Commercial supplier, Lonza, Verviers, Belgium

3.2. Overview of cell source component(s)

Primary human neural progenitor cells (hNPCs) are provided as cryopreserved 3D neurospheres from Lonza, Verviers, Belgium. The material originates from the human brain cortex of different gestational ages (GW16-19). Sex is either specified or determined before the cells are used.

3.3. Characterization and definition of source cells

$1x10^6$ hNPCs per vial are obtained from Lonza (#PT-2599) and expanded according to the SOP (Appendix J in Masjosthusmann et al. 2020). Lonza provides the cells with a viability of at least 20%. FACS analysis confirmed that proliferating neurospheres express the cell type-specific CNS neural stem cell and progenitor cell markers nestin, SRY-box 2 (SOX2), and Ki67 (Koch et al., 2022). Moreover, proliferating hNPCs react to growth factor stimuli (epidermal growth factor (EGF) and recombinant human fibroblast growth factor (FGF)) with increased proliferation, while simultaneous pharmacological inhibition of the EGF receptor (PD153035) impaired the proliferation increase. Upon transfer of hNPC neurospheres on poly D-lysin/laminin matrix and cultivation in the absence of growth factors (EGF and FGF), the hNPCs differentiate into effector cells expressing markers of neurons (β-III-tubulin), astrocytes (GFAP), radial glia cells (nestin) and oligodendrocytes (O4) (Baumann et al., 2015; Schmuck et al., 2017, Koch et al. 2022).

3.4. Acceptance criteria for source cell population

The following acceptability criteria have been tested at the supplier (Lonza) and are prerequisites for the shipment to customers:

- tested positive for TUBB3 and GFAP after differentiation
- tested free of HIV, HBV and HC
- tested negative both in sterility test and for mycoplasma contamination
- cell count of $1.2x10^6$ cells/mL
- viability of at least 20%

The proliferative capacity of Lonza hNPCs was reported previously (Moors et al., 2009; Baumann et al., 2015; Klose et al., 2021a).

3.5. Variability and troubleshooting of source cells

The sphere size at day 0 of cell thawing can be different depending on the donor.

In the 3 to 4-week expansion period different donors can show differences in their proliferative capacity (spheres need longer, 3 instead of 4 weeks, to reach the acceptable min. size of 0.2 – 0.5 µm). After the first mechanical dissociation, there are no observable or measurable inter-individual differences.

Critical consumables

The proliferation medium does not contain serum or serum replacement.

The use of epidermal growth factor (EGF) and recombinant human fibroblast growth factor (FGF) is critical for sphere growth. FGF contains 1% bovine serum albumin and is thus prone to batch effects.

Critical handling

The thawing medium contains DMSO in a concentration that affects cell health which is why thawed cells should quickly be diluted in proliferation medium (30 mL of media for one vial of cells).

It is recommended to add FGF into the proliferation medium directly before thawing.

At the end of week two of the expansion period (see below), the spheres should be transferred to petri dishes coated with poly-(2-hydroxyethyl methacrylate) (poly-Hema) to prevent cell attachment.

Attached cells that are not differentiated can be gently detached using a 1000 µL pipet. To avoid repeated attachment, all cells should be transferred to a new poly-Hema coated petri dish.

Medium containing FGF should not be stored longer than 1 week at 4°C.

During the first two weeks, the medium should be removed using a 1000 µL pipet to keep the accidental removal of small spheres to a minimum. In addition, removed medium should be kept in a new petri dish under culture conditions until the next feeding day, to transfer accidentally removed spheres back to the culture.

The neurospheres should be well distributed in the petri dish to prevent aggregation. This is especially important after mechanical dissociations.

It is important to avoid frequent re-opening of the incubators, to ensure constant CO_2 and temperature levels. Furthermore, the smallest vibrations can lead to aggregations of neurospheres.

3.6. Differentiation towards the final test system

Cells are frozen in liquid nitrogen and have to be cultivated in proliferation medium at 37°C and 5% CO_2 after thawing. The medium contains Dulbecco's modified Eagle medium and Hams F12 (2:1) supplemented with 2% B27, 20 ng/mL EGF, 20 ng/mL recombinant human FGF, 1% penicillin, and streptomycin. The thawing is performed by repeated addition and removal of proliferation medium to the vial until all cells are transferred to a tissue culture flask containing proliferation medium. The cells are carefully resuspended and distributed to 10 cm petri dishes filled with fresh, prewarmed proliferation medium. The cells are fed by replacing half the medium with new medium every two to three days (Monday, Wednesday, and Friday). At each feeding day, the culture is checked for impurities (e.g. fibers or other debris). Impurities and the removed media are transferred to a new petri dish (waste dish). If spheres are mistakenly sorted out during feeding, they can be rescued and placed back in the original culture dish. After 3-4 weeks neurospheres reach the acceptable size of 0.2 – 0.5 mm for passaging by mechanical dissociation. Therefore, neurospheres are mechanically dissociated into pieces of 0.15 - 0.25 mm edge length (depending on the desired sphere size after passaging) using a tissue chopper, which then round-off again to uniform sized neurospheres within 1 day in proliferation medium. By using this method, neurospheres are expanded every week. Starting at week poly-(2-hydroxyethyl methacrylate) (poly-Hema) coated dishes are used for the cultivation procedure.

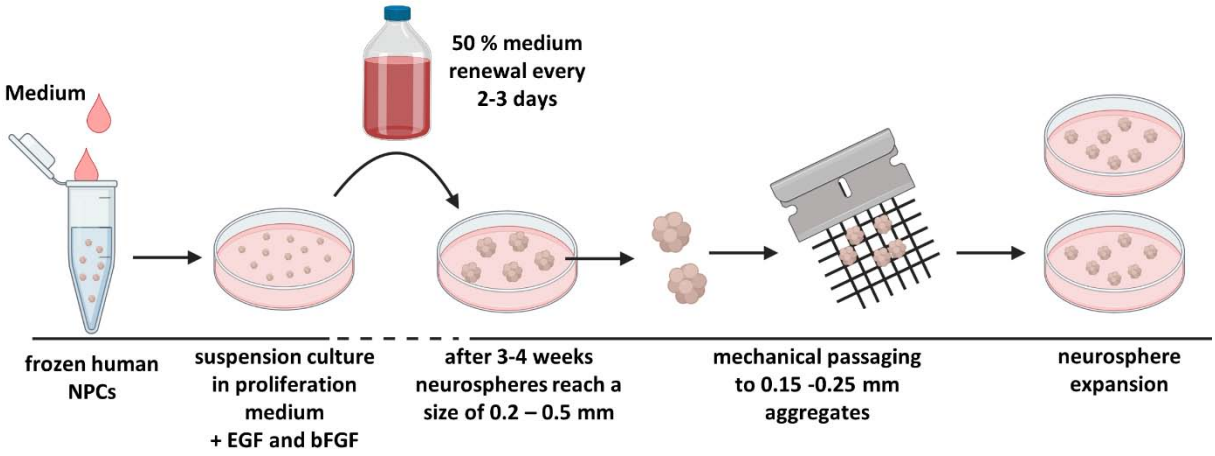


Figure 7 differentiation towards the final test system. *hNPC are thawed by repeated addition and removal of proliferation media. The resuspended cells are distributed to cell culture dishes and cultivated in proliferation media containing EGF and FGF for three to four weeks with 50% media exchange every two to three days. When the spheres reach a size of 0.2-0.5 mm they are expanded by mechanical passaging every 7 days.*

3.7. Reference/link to maintenance culture protocol

see the SOP (Appendix J in Masjosthusmann et al. 2020)

4. Definition of the test system as used in the method

4.1. Principles of the culture protocol

After the cell expansion period, the cells are cultured for up to four weeks in which they are passaged every week as described in 3.6. Between one to three days after passaging spheres at a size of 0.3 mm are used in the assay.

For the assessment of neural progenitor cell migration and differentiation, the spheres are plated on poly-D-lysine/laminin coated 96-well flat bottom plates in differentiation medium (N2) to initiate migration and differentiation. Therefore, one sphere of 0.3 mm diameter is plated in the middle of a well. The differentiation medium consists of DMEM and Ham's F12 at a ratio of 2 to 1 supplemented with 1% N2 and 1% penicillin and streptomycin. Within 5 days NPCs radially migrate out of the sphere core and differentiate into radial glia cells (nestin positive), neurons (β-III-tubulin positive), oligodendrocytes (O4 positive) and astrocytes (GFAP positive). Cultivation during the test method is performed at 37°C and 5% CO2 at a pH of 7.2-7.6.

4.2. Acceptance criteria for assessing the test system at its start

The spheres need to be rounded and have a size of 0.25 – 0.35 mm to be used in the test method.

Additionally, the basic neurospheres culture is checked for mycoplasma contamination every three month and controlled for fungal and bacterial contamination by visual inspection at each feeding and plating day.

4.3. Acceptance criteria for the test system at the end of compound exposure

As described in 4.1 the cells radially migrate out of the sphere core and differentiate into the effector cells. For this process the following acceptance criteria are defined for the solvent control (SC; median of at least two replicates):

Radial glia migration 72h:	700 – 1500 µm
Radial glia migration 120h:	700 – 1600 µm
Cell number:	2500 – 6000
Percentage of neurons	≥ 1.5 %
Percentage of Oligodendrocytes	≥ 1.5 %

4.4. Variability of the test system and troubleshooting

Sources of Variation:

Selection of spheres: Depending on the researcher and the availability of spheres, the size of selected spheres can differ.

Primary hNPCs are a complex multicellular system with a self-organized sphere composition as well as migration and differentiation. Due to the complex multicellular and self-organizing nature, the test system is subject to some heterogenicity which is represented as biological variability of some of the measured endpoints.

The variability for the different endpoints is shown in 8.4 “Test Performance”.

4.5. Metabolic capacity of the test system

Primary hNPCs under proliferating and differentiating conditions do not express CYP1A1 and CYP1B1 (Gassmann et al., 2010).

Primary hNPCs during differentiation, have the capacity to up-regulate glutathione-dependent protective strategies upon reactive oxygen species (ROS) exposure (Masjosthusmann *et al*, 2019).

Gene expression levels of genes involved in the antioxidative defense (glutathione peroxidase 1 (GPX1), superoxide dismutase 1 (SOD1), catalase (CAT)) were comparable between the *in vitro* system and developing human brains *in vivo* and show similar expression levels (Masjosthusmann et al., 2019).

Other metabolic pathways are not characterized.

4.6. Omics characterization of the test system

Proliferating, three day differentiated, and five day differentiated hNPCs were analyzed for changes in their transcriptomic profile. Several key neurodevelopmental processes (migration, neuronal differentiation, glial differentiation) and genes regulating these processes (Bone morphogenetic protein (BMP), Notch and EGF signaling) were identified and characterized on a functional level (Masjosthusmann et al., 2018).

4.7. Features of the test system that reflect the *in vivo* tissue

hNPCs reflect the following in vivo tissue features:

In differentiation culture:

NPC2a – radial glia cell migration → corresponding to radial glia cell migration during corticogenesis in vivo.

The migrated hNPCs exhibit the characteristic elongated RG-like morphology and express the RG-markers nestin and GFAP as well as the proliferation marker Ki-67 (Koch et al., 2022). In accordance with in vivo studies, exposure to EGF (0.5-1 ng/mL) after neurosphere plating enhances hNPC migration compared to the solvent control (Koch et al., 2022). As a second human-relevant key regulator of migration, hNPCs respond to SRC-family kinase inhibition (PP2) with reduced migration (Moors et al., 2007; Koch et al., 2022). The migration speed of hNPCs in vitro is in the same range as migrating mouse granule cells in vivo (Baumann et al., 2016, Fahrion et al., 2012).

NPC2b – migration of young cortical neurons on a radial glia scaffolds → corresponding to cortical neuronal radial migration.

NPC2c – oligodendrocyte migration → corresponding to oligodendrocyte migration during corticogenesis.

NPC3 – fetal neuronal differentiation into young neurons → corresponding to cortical neurogenesis in vivo.

Over the time course of differentiation, neurons expressing β(III)tubulin progressively appear in the migration zone representing approximately 20% of the mixed culture after 5 days. In line with observations in vivo, inhibition of the Notch signaling pathway during hNPC differentiation increases neuronal numbers compared to the solvent control (Koch et al., 2022). Moreover, narciclasine, an activator of RhoA, reduces neuronal differentiation of hNPCs cultured for 5 days in a concentration-dependent manner (Masjosthusmann et al., 2020). Likewise, RhoA inactivation stimulated axon regeneration and recovery of hindlimb function after spinal cord injury in mice (Dergham et al., 2002).

NPC4 – neurite length, neurite area of young primary fetal neurons → corresponding to axon/dendrite formation in vivo.

During the 5 days of hNPC differentiation, neurite maturation is characterized by an elongation of neurites and an increase in neurite area (Koch et al., 2022). Increase in RhoA activity caused morphological alterations in rat cortical neurons in vivo (Chen et al., 2018). In line with that, RhoA activator narciclasine reduced both neurite area and neurite length (Masjosthusmann et al., 2020).

NPC5 – oligodendrocyte formation from fetal hNPCs → corresponding to oligodendrogenesis during the fetal phase of brain development.

Differentiation of hNPCs over 5 days progressively generates cells expressing the oligodendrocyte-marker O4, which exhibit the typical oligodendrocyte morphology with multiple branched processes necessary to ensheath neuronal axons (Moors et al., 2009; Koch et al., 2022). Differentiation of hNPCs in presence of the Notch inhibitor DAPT concentration-dependently decreases the percentage of O4-positive cells compared to the solvent control (Koch et al., 2022). Moreover, oligodendrocyte differentiation is negatively influenced by bone morphogenic protein (BMP) 7 (Baumann et al., 2015) and BMP2 (Masjosthusmann et al., 2018), proteins of the transforming growth factor β family. These data demonstrate that two major developmental pathways, i.e. Notch and BMP, are functional in hNPCs.

Primary hNPCs are self-organized and produce auto- and paracrine cues like HB-EGF and neuregulins that guide migration and differentiation. Moreover, they express extracellular matrix (ECM) proteins like laminin subunits, fibronectin and collagens (microarray data from Klose et al. 2021). Hence, effects concerning the ECM like adhesion defects can be assessed with this assay (Barenys et al. 2017). Various signaling cues that guide migration and differentiation processes of the neurosphere assay are summarized in Koch et al., 2022.

4.8. Commercial and intellectual property rights aspects of cells

For the source cells, Lonza holds donor consent and legal authorization that provides permission for all research use.

4.9. Reference/link to the culture protocol

See 3.7.

5. Test method exposure scheme and endpoints

5.1. Exposure scheme for toxicity testing

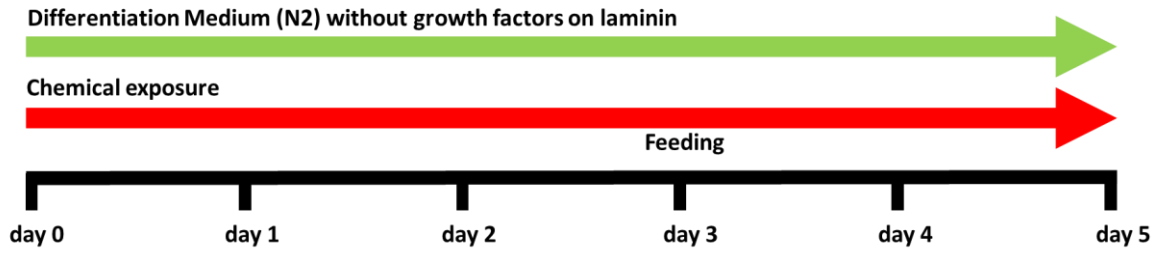


Figure 8 Exposure scheme. Spheres are plated in PDL/Laminin-coated 96 well F-bottom plates in differentiation medium and exposed to increasing compound concentrations over a cultivation time of 120 h. Half of the medium is replaced after 72h of cultivation.

hNPCs of 0.3 µm diameter are plated as described in 4.1. Cells are plated according to the plating scheme in Figure 4 in the already prepared test conditions. Exposure starts at day 0 of differentiation and is continued over five days of differentiation until the experiment is terminated. Cells are fed with fresh medium on day 3 of differentiation (Figure 2). Therefore, half of the test condition solution (e.g. solvent control or compound dilution) is replaced by freshly prepared test condition solution.

5.2. Endpoint(s) of the test method

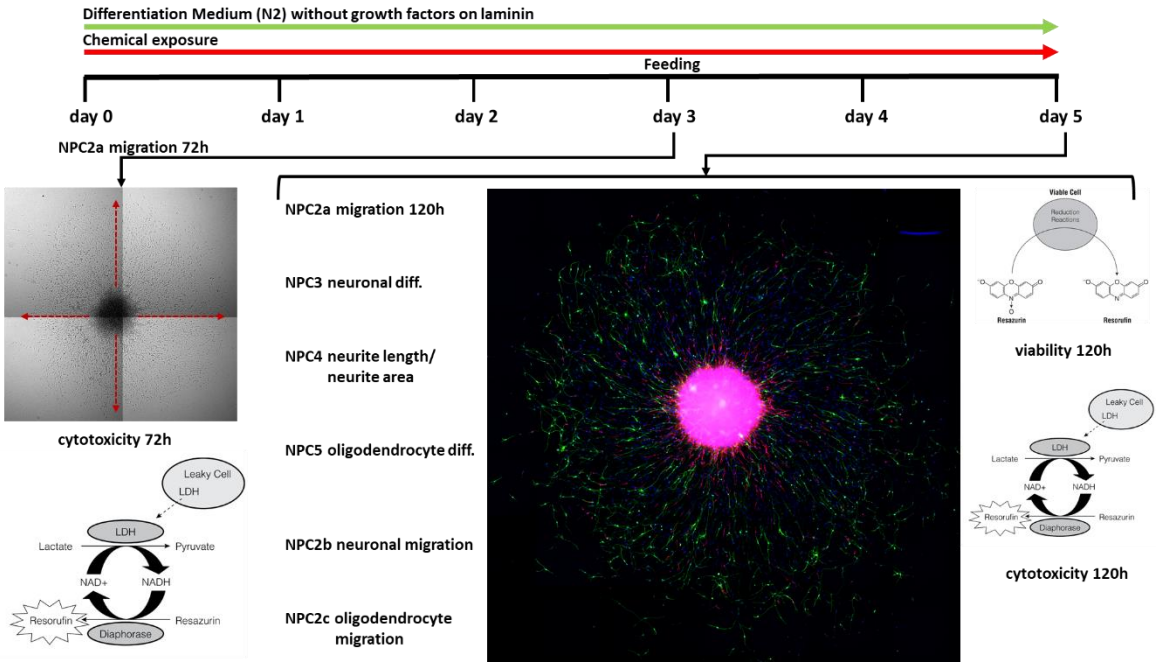


Figure 9 Endpoint assessment. *Migration distance and cytotoxicity is determined after 72 h. The assay is terminated by the assessment of cell viability and cytotoxicity as well as cell fixation after 120 h. Immunocytochemistry is performed to assess Hoechst positive nuclei, β(III)-tubulin-positive neurons and O4-positive oligodendrocytes. On the ICC images, migration after 120 h, neuronal and oligodendrocyte number, neuronal morphology and neuron/oligodendrocyte specific migration is assessed*

Primary DNT specific endpoints of the test method are:

3 . radial glia migration 72 h (NPC2a)
4 . radial glia migration 120h (NPC2a)
5 . neuronal migration 120 h (NPC2b)
6 . oligodendrocytes migration 120h (NPC2c)
7 . neuronal differentiation 120 h (NPC3)
8 . neurite length 120 h (NPC4)
9 . neurite area 120 h (NPC4)
1 0 . oligodendrocyte differentiation 120h (NPC5)

Secondary endpoints are:

3 . cell number 120 h
used for normalization of neuronal and oligodendrocyte differentiation
4 . cytotoxicity 72 h
5 . cytotoxicity 120 h
6 . viability 120 h

All endpoints are generated from the same experimental run and from each well/sphere in the 96 well plate.

5.3. Overview of analytical method(s) to assess test endpoint(s)

Primary endpoints:

1. Radial glia migration 72 h is assessed as migration distance in µm from the edge of the sphere core to the edge of the migration area based on brightfield images of each well (Figure 3). Therefore, each plate is scanned after 72 h in culture in an automated high content imaging device. Images are exported and the sphere size in four directions is measured manually using ImageJ.

All other primary endpoints are assessed based on an immunocytochemical staining (ICC) image for each sphere. Therefore, cells are fixated after 120 h in culture and an ICC staining with Hoechst for nuclei, β(III)-tubulin for neurons and O4 for oligodendrocytes is performed. The plates are scanned using an automated high content imaging device and all nuclei and their positions are determined automatically based on their intensity and size. Images are imported to the Omnisphero software (https://omnisphero.com) to run the image analysis that measures the following endpoints.

2. Radial glia migration 120h is assessed as the migration distance in µm between the sphere core and the edge of the migration area based on ICC images of Hoechst-positive nuclei. By identification of each nuclei's position in relation to the sphere core, the migration distance can be calculated. Therefore, a density distribution mask is calculated. By scanning the images of the nuclei channel, the algorithm can determine relatively more or less dense image areas. By identifying the densest area in the image, the sphere core can be detected. For identification of the migration, it is assumed, that the nuclei density decreases with increasing distance to the sphere core. Once the density hits a pre-defined threshold, the outer boundaries are determined and the sphere itself can be mapped out in a polynomial bounding box. Derived from this box, the actual size and migration distance can be calculated for each well.

3. Neuronal migration 120 h is the mean distance of all neurons from the edge of the sphere core in relation to the radial glia migration and is determined based on the position of each neuron (see neuronal differentiation).

4. Oligodendrocyte migration 120 h is the mean distance of all oligodendrocytes from the edge of the sphere core in relation to the radial glia migration and is determined based on the position of each oligodendrocytes (see oligodendrocyte differentiation).

5. Neuronal differentiation is defined as the number of all β(III)-tubulin-positive cells in percent of the cell number (Hoechst-positive cells) in the migration area after 120 h of differentiation. The identification of neurons is done automatically using a convolutional neural network (CNN). Training of the CNN was done based on manually annotated experiments (Förster et al., 2021).

6. Neurite length 120 h is the mean length in µm of all neurons (see neuronal differentiation) that are identified by the skeletonization algorithm in Omnisphero (https://omnisphero.com).

7 . Neurite area 120 h is the mean area in pixel (without nuclei) of all neurons (see neuronal differentiation) that are identified by the skeletonization algorithm in Omnisphero (https://omnisphero.com).

8 . Oligodendrocyte differentiation 120 h is defined as the number of all O4-positive cells in percent of the cell number (Hoechst-positive cells) within the migration area after 120 h of differentiation. The identification of oligodendrocytes is done automatically using a convolutional neural network (CNN). Training of the CNN was done based on manually annotated experiments (Förster et al., 2021).

Secondary endpoints:

3 . Cell number is the number of all Hoechst-positive nuclei detected within the area between the sphere core and the outer boundaries of the migration area (see radial glia migration 120 h).

4 . Cytotoxicity 72/120 h is assessed as membrane integrity by measuring the amount of LDH leaked from cells with damaged plasma membranes. LDH-dependent reduction of resazurin to resorufin is measured in the supernatant of each well as fluorescence of the reaction product resorufin (relative fluorescence unit) in a multiplate reader after 72/120 h of differentiation and compound treatment.

5 . Viability 120 h is assessed as mitochondrial activity by measuring the amount of resazurin reduced to fluorescent resorufin (relative fluorescence unit) in a multiplate reader in the last two hours of the 120 h differentiation and compound treatment period.

5.4. Technical details (of e.g. endpoint measurements)

All technical details for the test method are available in the SOP (Appendix J in Masjosthusmann et al. 2020).

5.5. Endpoint-specific controls/mechanistic control compounds (MCC)

Primary Endpoints:

All endpoint-specific controls are run with each experimental run (experimental day) and for each human individual and passage number (Figure 10).

3. Radial glia migration 72 h: The endpoint-specific control for radial glia migration after 72 h is the SRC kinase inhibitor PP2. SRC family kinases represent one pathway that regulates radial gilia migration of differentiating hNPCs (Moors et al., 2007). Inhibition of this pathway with PP2 (10 µM) causes a reduction of radial glia cell migration to values between 0 and 60 % of the solvent control.

4. Neuronal differentiation: The endpoint specific control for neuronal differentiation is EGF. EGF is a growth factor that stimulates radial glia proliferation and migration and inhibits neuronal differentiation (Ayuso-Sacido et al., 2010; Jenny Baumann et al., 2016). 20 ng/mL EGF reduces neuronal differentiation to values between 0 and 50% of the solvent control.

5. Oligodendrocyte differentiation: The endpoint specific control for oligodendrocyte differentiation is BMP7. BMP7 promotes the BMP signaling cascade which upregulates astroglia differentiation and maturation and inhibits oligodendrocyte formation (Jenny Baumann et al., 2016; Gross et al., 1996; Mabie et al., 1997). 100 ng/mL BMP7 reduces oligodendrocyte differentiation to values between 0 and 60 % of the SC.

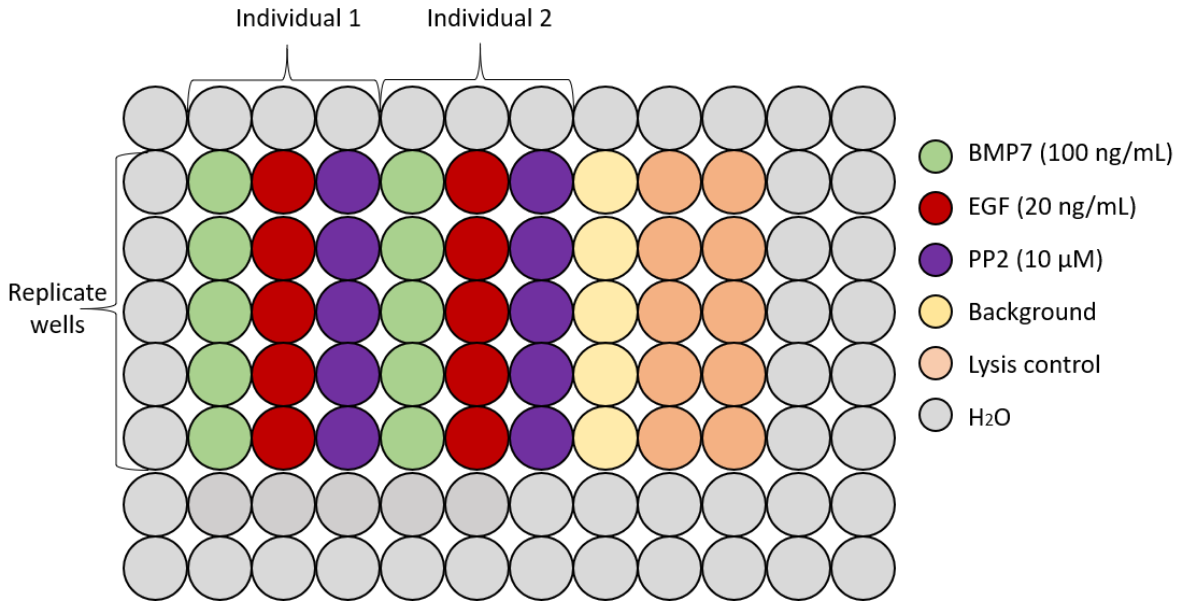


Figure 10 Plating scheme of control plate. *The control plate is prepared for each experimental run with each control for one individual or passage if different individuals or passages where used in the experimental run.*

Secondary Endpoints:

0.2 % Triton X-100 is used as a positive control for cell viability and cytotoxicity since it lyses the cell and therefore causes a maximal response for both endpoints. This positive control is run on each experimental plate (Figure 4).

5.6. Positive controls

The NPC2-5 hNPC migration and differentiation assay correctly identified the following compounds that are known to cause DNT in humans or in vivo (Masjosthusmann et al., 2020):

Methylmercury(II) chloride

Cadmium chloride

Hexachlorophene

2,2',4,4'-Tetrabromodiphenyl ether

Dexamethasone

Manganese(II) chloride

Chlorpromazine hydrochloride

Haloperidol

Paraquat dichloride hydrate

Trichlorfon

Deltamethrin

Sodium valproate

Tebuconazole

Tributyltin chloride

Potassium perfluorooctanesulfonate

5.7. Negative and unspecific controls

The solvent control (SC) is used as negative control that is run on each experimental plate. Each SC has to be established by comparing the effect of the SC to the effect of the media control. Established solvent controls show the same response as the media control.

The SC is used to assess if the acceptability criteria for radial glia migration, neuronal and oligodendrocyte differentiation are met and to normalize the compound treatment and the positive control response.

Established solvent controls are:

DMSO: 0.1 % v/v

DPBS: 2 % v/v

dsH2O: 2 % v/v

MeOH: 0.1% v/v

Other negative control compounds that were identified as negative in this assay and are known to not affect neurodevelopmental endpoints in vivo include (Masjosthusmann et al., 2020:

Acetaminophen

Amoxicillin

Aspirin

Buspirone

Chlorpheniramine maleate

D-Glucitol

Diethylene glycol

D-Mannitol

Doxylamine succinate

Famotidine

Ibuprofen

Metformin

Metoprolol

Penicillin VK

Saccharin

Sodium benzoate

Warfarin

5.8. Features relevant for cytotoxicity testing

Differentiating hNPCs are a multicellular system consisting of radial glia, neurons (1.5-16 %), oligodendrocytes (1.5-11%) and astrocytes within the migration area. The measurement of cytotoxicity and viability therefore always represents all cells within the migration area.

Because of the higher percentage of GFAP-positive radial glia and astrocytes, these two cell types are overrepresented in the assessment of cytotoxicity and viability.

The measure of cell viability assessed by the Alamar blue assay (mitochondrial reductase activity) strongly depends on the number of cells in the migration area. Therefore, a reduction in cell number either due to a reduced migration distance or a lower cell number in the migration area will lead to a reduction in the Alamar blue signal despite the cell viability is not necessarily affected (Figure 3 in Nimtz et al. 2019). Therefore, as soon as the migration distance or cell number in the migration area are reduced, the Alamar blue (viability) assay is not used as a measure for viability, but the cytotoxicity assay (based on LDH release) is taken as a reference for the specific DNT endpoints to determine if the effect is specific or not.

5.9. Acceptance criteria for the test method

General acceptance criteria:

1. At least two replicate values need to be present for each condition to be accepted for the data analysis.
2. At least five conditions need to be present for the experiment to be accepted in the data analysis for concentration response modeling.

Endpoint dependent acceptance criteria:

The acceptance criteria described below is the MEAN response of at least two replicates of the SC:

- radial glia migration 72 h: 700 – 1500 µm
- radial glia migration 120h: 700 – 1600 µm
- neuronal differentiation: ≥1.5 % neurons
- oligodendrocyte differentiation: ≥1.5 % oligodendrocytes
- cell number: 2500 – 6000

The acceptance criteria described below is the response for each replicate and is applied to all conditions:

1 . neuronal migration: ≥ 5 neurons
2 . oligodendrocyte migration ≥ 5 oligodendrocytes

General acceptance criteria:

1. At least two replicate values need to be present for each condition to be accepted for the data analysis.
2. At least five conditions and the solvent control need to be present for the experiment to be accepted in the data analysis for concentration response modeling.

5.10. Throughput estimate

The methods described here are set up in a 96-well plate format with automated image acquisition, analysis, and data evaluation. Pipetting steps such as coating of 96-well plates, compound dilutions, feeding, cell viability and cytotoxicity assay can be automated using a liquid handling system.

In the fully automated setup, 10 plates with 8 conditions (Figure 5) and 4-5 replicates per condition can be run in one week by two technical assistance. This results in the generation of 400 data points for each endpoint within one week (excluding all controls). The throughput is therefore estimated as medium.

6. Handling details of the test method

6.1. Preparation/addition of test compounds

The method is set up for 8 test conditions including 7 compound concentrations and one SC. The test conditions are prepared in a serial dilution from the stock solution (Figure 4).

Stock solutions are prepared by diluting the compound in the solvent (e.g. DMSO) in a concentration that allows the preparation of the highest test concentration without exceeding the highest acceptable solvent concentration (see 5.7). For DMSO the highest acceptable solvent concentration is 0.1% which means that the stock concentration needs to be at least 1000x higher than the highest test concentration.

Stock solutions in non-sterile solvents (e.g. water or PBS) have to be sterile filtrated using a sterile syringe filter (Ø 0.2 µm). Adsorption to the filter needs to be considered.

Stock solutions are aliquoted and stored at -20°C. A stock solution is not thawed more than three times.

For the preparation of the test condition the stock solution is diluted to the highest test concentration (default 1:1000) in differentiation medium. All following dilutions are prepared by serial dilutions of the highest concentration in differentiation medium containing solvent (in the concentration of the highest test concentration). The default serial dilution is 1:3 which covers a concentration range from e.g. 20 µM to 27 nM (729-fold). Depending on the desired concentration range, the dilution can be changed to 1:2, 1:5, 1:10 or other.

The SC is prepared by adding the solvent to differentiation media in the same concentration as the highest test concentration.

100 µL of the compound dilutions and the SC are added to a 96-well plate (Figure 4).

The serial dilution can also be prepared directly in the 96 well plates.

hNPCs are added to each well after a 15 to 30 min equilibration period at 37°C and 5 % CO_2.

To subtract the background fluorescence of the phenol red-containing medium, 5 wells with medium only (Background, Figure 5) are prepared for the mitochondrial activity and cytotoxicity assay.

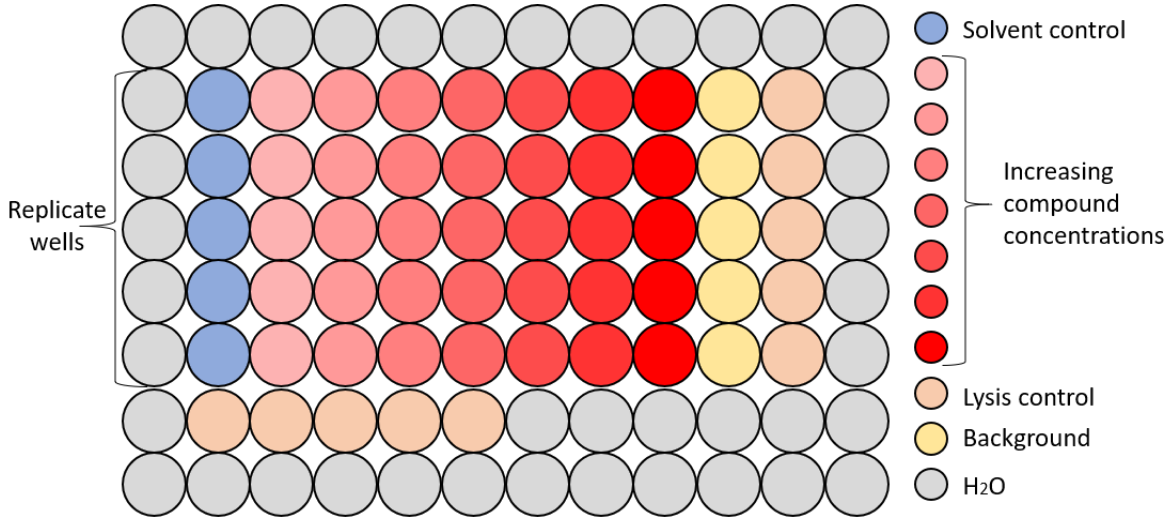


Figure 11 Plating Scheme. *The solvent control depends on the solvent of the compound that is tested. 7 compound concentrations are plated in a serial dilution from lowest (left) to highest (right) concentration. Lysis control and background control (5 replicates for each timepoint; 72 and 120 h) are used for the viability and cytotoxicity assays.*

6.2. Day-to-day documentation of test execution

Documentation for each experiment including meta data and the experimental data that is collected in the Automated Experimental Evaluation or AXES sheet.

Meta data such as plating date, experimenter, NPC individual, NPC passage, compound, compound concentrations and a plate map are reported in these sheets. Depending on the endpoint, the experimental data is collected during or at the end of the experiment. Each raw data point (including all outliers) is collected in the AXES sheet. All deviations from the standard procedure are documented in a comment section of the AXES sheet.

6.3. Practical phase of test compound exposure

The practical phase of the test compound exposure follows the description in the SOP (Appendix J in Masjosthusmann et al. 2020). Deviation from the SOP are documented in the comment section of the AXES sheet.

Errors (e.g. pipetting in wrong well or wrong volume pipetted) are also documented in the comment section of the AXES sheets. Data points of the affected well are marked in the AXES sheet and excluded from the analysis.

6.4. Concentration settings

Starting concentrations and concentration ranges are defined based on the following factors:

- toxicological relevance of the compound (i.e. internal human exposures, effects at lowest concentrations)

- solubility of the compound
- highest useable solvent concentration

These factors are determined based on available information (databases/literature) or experimentally (e.g. solubility test).

6.5. Uncertainties and troubleshooting

Problematic compounds:

- Volatile compounds
- High lipophilicity (high K_{OW})
- Low solubility in established solvents
- Fluorescent compounds (possible interfere with viability and cytotoxicity assay)

Critical handling steps:

- The poly-D-lysine/laminin coating as well as the plate format and plate type are critical for cell migration and differentiation. Coated plates should be stored at 4°C for no longer than 7 days. If different plate types are used, the test system and test method need to be re-established.
- Outer wells have to be filled with H_2O because of edge effects.
- Automating the pipetting steps using a liquid handling system for coating, preparation of the plates, feeding of the plate, viability and cytotoxicity assay reduces the variability and the user bias.

Sources of variation:

- Pipetting steps: Each pipetting step is a source of variation. Especially in the viability and cytotoxicity assay where the volume pipetted determines the final readout.
- Immunocytochemical staining: The staining consists of multiple washing steps. As pipetting errors add up in each washing step, there is variation in the dilution of blocking solution and antibodies.
- NPC differentiation: The biological variation between the differentiation of different NPC spheres cannot be controlled and can lead to variations in the endpoints neuronal and oligodendrocyte differentiation.

Known Pitfalls:

- Spill over from the lysis control wells to other wells can happen and needs to be controlled by the operator. It is important to completely empty the 3-day lysis control wells after they have been used and fill them with water (see SOP; Appendix J in Masjosthusmann et al. 2020), to avoid spill over during the last 2 days of the experiment.
- Multiple washing steps in the ICC staining lead to added pipetting errors and may cause differences in the volume in each well. Here the operator needs to control that the spheres are always covered by PBS. Differences between the wells can be corrected by the operator.

6.6. Detailed protocol (SOP)

See the SOP (Appendix J in Masjosthusmann et al. 2020).

6.7. Special instrumentation

- Incubator for cell culture
- Tissue chopper
- Multiplate reader for fluorescence measurement
- High content imaging device for automated fluorescence microscopy
- Liquid handling system (necessary to achieve the throughput described above)

6.8. Possible variations

Other endpoints:

- Measurement of BMP2 dependent astrocyte maturation (Masjosthusmann et al., 2018)
- Distribution of the neuronal density within the migration area (Schmuck et al., 2017)
- Oligodendrocyte maturation (Katharina Dach et al., 2017)
- Migration pattern (Barenys et al., 2016)

Other exposure schemes:

- 24 h migration (Jenny Baumann et al., 2016)
- 72 h neuronal differentiation (Jenny Baumann et al., 2016)

6.9. Cross-reference to related test methods

No related test methods

7. Data management

7.1. Raw data format

The raw data format is different depending on the endpoints.

For all endpoints assessed in a multiplate reader (viability and cytotoxicity), the raw data format are excel files containing values (one for each endpoint, timepoint and well) measured as relative fluorescence units. These values are transferred from the original excel file into the AXES sheet. The original excel output files are saved for traceability of the data.

The radial glia migration after 72 h, which is measured manual in ImageJ, is directly copied into the AXES sheet as values in µm. Original brightfield images are archived for 10 years.

All other raw data is computed from the ICC images in the Omnisphero software and is exported and saved as one csv file. From there the values are again transferred to the AXES sheets. The following data is exported from Omnisphero:

- number of all cells in the migration area (cell number)
- number of all neurons in the migration area
- number of all oligodendrocytes in the migration area
- radial glia migration (µm)
- mean neuronal migration (µm)
- mean oligodendrocyte migration (µm)
- neurite length (µm or pixel)
- neurite area (pixel)

All original ICC images are archived for 10 years.

7.2. Outliers

Mathematical procedures to define outliers are not applied. Data points from wells where technical problems are known or obvious are excluded from the analysis.

Possible technical problems:

- pipetting errors
- spillover from lysis
- washed-off sphere without indication of cytotoxicity or reduction in cell viability
- wrong or no identification of migration area and sphere core
- problems in ICC staining
    - cells dried out
    - wrong illumination
    - blurry pictures

All outliers are marked in the AXES sheet.

7.3. Raw data processing to summary data

If not otherwise stated, all data processing steps are performed in an R based evaluation tool that was designed for data processing, curve fitting and point of departure evaluation of in vitro concentration response toxicity data.

Data processing describes all processing steps of raw data that are necessary to obtain the final response values including the normalization, curve fitting and benchmark concentration calculation.

Processing (or pre-processing) steps depend on the endpoint and are described below:

radial glia migration 72 h: The mean of four replicate measures of each sphere. The mean of four measures per well is used as raw data input and is not calculated in the R based evaluation tool.

radial glia migration 120 h: no pre-processing.

cell number 120 h: no pre-processing.

neuronal differentiation: the number of all neurons is divided by the number of all cells (in the migration area).

$$neuronal\ differentiation\ [\%] = \frac{\#\ neurons}{\#\ cells} * 100\ \%$$

neuronal migration 120 h: mean neuronal migration (in the migration area) divided by radial glia migration 120 h.

$$migration\ distance\ neurons\ [\%] = \frac{mean\ migration\ distance\ all\ neurons\ [\mu m]}{migration\ distance\ radial\ glia\ 120\ h\ [\mu m]} * 100\ \%$$

oligodendrocyte differentiation: number of all oligodendrocyte is divided by the number of all cells (in the migration area).

$$oligodendrocyte\ differentiation\ [\%] = \frac{\#\ oligodendrocytes}{\#\ cells} * 100\ \%$$

oligodendrocyte migration 120 h: mean oligodendrocyte migration (in the migration area) divided by the radial glia migration 120 h.

$$migration\ distance\ oligo.\ [\%] = \frac{mean\ migration\ distance\ all\ oligo.\ [\mu m]}{migration\ distance\ radial\ glia\ 120\ h\ [\mu m]} * 100\ \%$$

Viability: subtraction of mean background from each response value.

$$Background\ corrected\ response\ [RFU] = raw\ response\ [RFU] - Background\ [RFU]$$

where RFU is a relative fluorescent unit of the plate reader output

Cytotoxicity: no pre-processing

Neurite Area: no pre-processing

Neurite length: no pre-processing

7.4. Curve fitting

The data is normalized to the SC and re-normalized to the starting point of the curve. For the normalization to the SC each replicate data point is normalized to the median of the SC in the respective experiment. For the cytotoxicity assays the following normalization is used instead of the normalization to the SC. Here again each response value is normalized using the median of the lysis control and the median of the solvent control.

$$\text{normalized response} = \frac{\text{lysis control} - \text{response}}{\text{lysis control} - \text{solvent control}}$$

The R package drc is used to calculate the optimal fit for each experiment. For calculations of curve fits and BMCs, the data from independent experiments is pooled (median of all replicate values for one concentration). Several non-linear models are run with the concentration response data of each endpoint and the Akaike´s information criteria is used to determine the best fit.

For re-normalization of the data, the response value of the curves starting point is determined and used to re-normalize all response values. Therefore, each mean response value is divided by the starting point of the curve and multiplied with 100. For the re-normalized response values the curve fitting is repeated to produce the final concentration response curve.

For deriving a reference point (RP) or point of departure (Pod) the Benchmark Concentration (BMC) approach, as recommended by the EFSA Scientific Committee (Hardy et al., 2017), is applied. The BMC approach makes use of all data points that define the fitted concentration response curve. Thereby, the BMC is defined as the concentration that is associated with a specific change in response, the Benchmark Response (BMR). The BMR is a value of effect size and should be defined as an effect size that is higher than the general variability of the measured endpoint. The BMR is therefore determined based on the variability of the respective endpoint.

BMR for NPC2-5:

Primary DNT specific endpoints of the test method are:

radial glia migration 72 h (NPC2a)	BMR10
radial glia migration 120h (NPC2a)	BMR10
neuronal migration 120 h (NPC2b)	BMR30
oligodendrocyte migration 120h (NPC2c)	BMR30
neuronal differentiation 120 h (NPC3)	BMR30
neurite length 120 h (NPC4)	BMR30
neurite area 120 h (NPC4)	BMR30
oligodendrocyte differentiation 120h (NPC5)	BMR30

Secondary endpoints are:

cell number 120 h	BMR30
cytotoxicity 72 h	BMR10
cytotoxicity 120 h	BMR10
viability 120 h	BMR30

Based on the BMR and the concentration response curve, the evaluation tool calculates the BMC, as well as upper and lower confidence limits (BMCU and BMCL respectively) based on the predict function in the R package drc. The predict function calculates the prediction bands around the concentration response curve based on the deviation between independent experiments and gives an estimation of the area that is expected to enclose 95% of future data points. The BMCL is thereby defined as the intersection of the lower band and the BMR while the BMCU is defined as the intersection of the upper band and the BMR. The confidence intervals are used to access the uncertainty of the BMC. If the BMCU is 1.5 times above the test range, the original BMCU is replaced by 1.5x the highest tested concentration.

7.5. Internal data storage

All raw data is stored on a server with a daily server back up for at least 10 years.

7.6. Metadata

All metadata is collected in the AXES sheet (see 6.2) together with all raw data.

The metadata gives information on:

The experiment:

- start and end date of the experiment
- experimenter

The cell source:

- human individual
- cells thawing date
- passage of cells
- date of cell passaging

The compound:

- compound identity

- stock concentration
- all dilution steps
- solvent and solvent concentration

The controls:

- control identity
- preparation of controls

7.7. Metadata file format

All metadata is collected in an Excel format.

8. Prediction model and toxicological application

8.1. Scientific principle, test purpose and relevance

Primary hNPCs are isolated from the fetal brain and can be used to model neurodevelopmental processes like migration of radial glia, neurons and oligodendrocytes, neurite outgrowth and differentiation into neurons and oligodendrocytes within this test method.

The test system therefore measures adverse events in the young (fetal) developing brain.

Thereby, the biological processes that are modeled are:

1. radial glia cell migration (NPC2a)
2. migration of young cortical neurons (NPC2b)
3. oligodendrocyte migration (NPC2c)
4. fetal neuronal differentiation into young neurons (NPC3)
5. neurite outgrowth of young primary fetal neurons (NPC4)
6. oligodendrocyte formation of fetal NPC (NPC5)

The toxicological events that are modeled include events that impact the above-mentioned biological processes in any direction (increase or decrease). Thereby, neurodevelopmental processes represented by the NPC2-NPC5 assay are guided by a variety of pathways known to contribute to the respective processes in vivo as assessed by hNPC exposure to in vivo relevant signaling pathway modulators (summarized in Table13, Masjosthusmann et al. 2020; Koch et al., 2022). All the endpoints can be assessed in a high content format within one experiment due to the multi-cellularity and high developmental dynamic of the neurospheres. In addition, the strength of the system is that it allows KE-based assessment of human neurodevelopment without the need to perform species extrapolation. Moreover, the self-organized mixed culture system contains cell-cell interactions e.g. via gap junction channels as cells die when cell-cell communication is blocked by gap junction blockers (unpublished observations). Hence, the system is superior to just plated single-cell 2D systems, since due to its nature it mimics a multitude of neurodevelopmental processes. One limitation of the hNPC model is the restricted timing. These endpoints represent early processes of neurodevelopment during the fetal period. Neurons stay immature hence later processes like synaptogenesis cannot be studied. Moreover, correct positioning, like one can study in vivo with cortical layering, cannot be studied here, since the assay mimics cortical radial glia and neuronal migration but does not cortical layer formation.

The test method predicts the hazard to induce developmental neurotoxicity in the form of neurophysiological, functional and behavioral changes in the developing nervous system.

8.2. Prediction model

Two prediction models (PM) are applied for the NPC2-5 assay. One PM for a downregulation (PM downregulation) and one PM for an upregulation (PM upregulation) in cell differentiation and migration.

PM down regulation

The PM is based on a hit definition followed by comparison of the confidence intervals (CI) for the BMC of the DNT-specific endpoint (BMCs) and the unspecific endpoint (cytotoxicity/viability; BMCus).

Thereby the following four hit classifications apply:

"no hit"	The compound is not defined as hit.
"specific hit":	The compound is defined as hit and the CI's do not overlap, meaning that the upper confidence limit of the specific endpoint (BMCUs) is lower than the lower confidence limit of the unspecific endpoint (BMCLus).
"borderline hit":	The compound is defined as hit and the CI of the specific endpoint overlaps by less than or equal to 10% with the CI of the unspecific endpoint.
"unspecific hit"	The compound is defined as hit and the CI of the specific endpoint overlaps by more than 10% with the CI of the unspecific endpoints.

The compound is classified as a hit, if the concentration response curve generates a BMC and if the CI is within the test range. In case the CI spans above the test range, the compound is only classified as hit, if the highest test concentration is significantly different from the lowest test concentration. The adjusted significance ($p<0.05$) is thereby determined using a Tukey HSD test.The decision process for the prediction model is described in the flow chart in Figure 12.

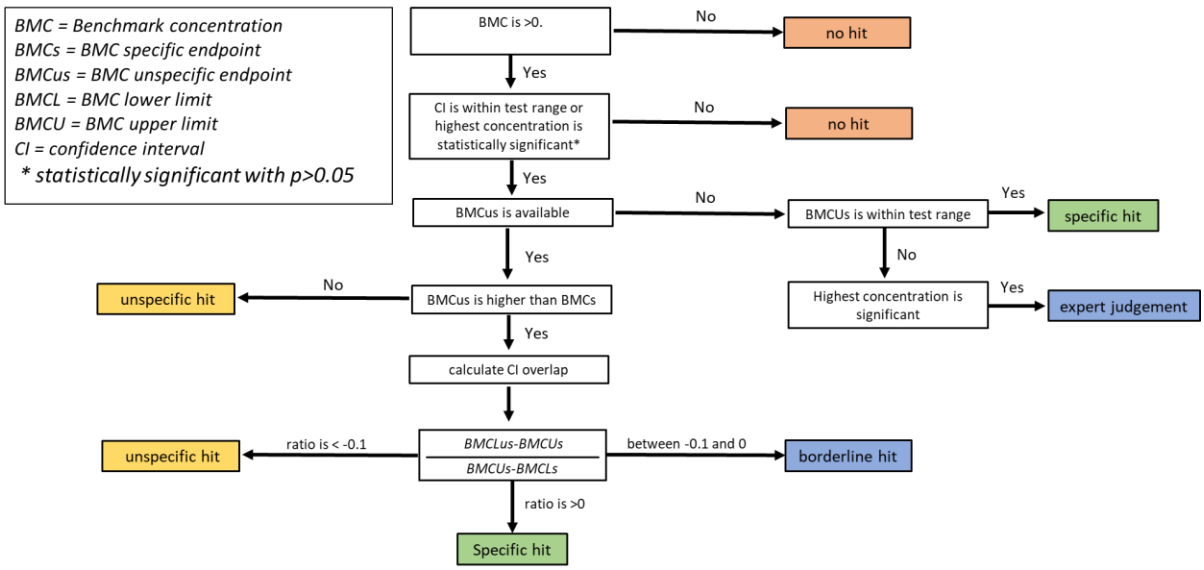


Figure 12: Decision tree for the PM for downregulation applied for test method NPC2-5. Overview of the decisions leading to the classification of a compound in one of four categories: "no hit", "specific hit", "borderline hit" and "unspecific hit".

Specific consideration for the prediction model:

In case no confidence limits are available for the unspecific endpoint because the BMR is not reached, the BMCLus is assumed to be the highest tested concentration. If the CI of the specific endpoint additional spans above the test range for a compound identified as "hit" based on statistical significance, expert judgement is applied to define, if the "hit" is specific or unspecific. Therefore, all datapoints of the unspecific and specific endpoint of the highest test concentration are compared. If these data points do not overlap, the compound is classified as "specific hit", otherwise as "unspecific hit".

In general, BMCs based on the same BMR are compared (e.g. BMC_{30us} vs BMC_{30s}). In case the BMC_{30us} is not available for an endpoint that allows the generation of a BMC_{10us}, the BMC_{10us} is used instead. If the classification of this comparison is "unspecific hit", the compound will be flagged as "check manually" as the BMC_{10us} is lower than the BMC_{30us} leading to a higher probability of a false classification. To avoid such false classification expert judgement is needed.

Compounds are also be flagged as "check manually", if the classification based the viability is different from the classification based on the cytotoxicity and if the confidence interval is very wide (BMCU/BMCL > 25), which means a high uncertainty for the BMC estimation. In both cases expert judgement is needed to decide on the classification.

The expert judgement is an individual decision process, that accounts for effect size, curve progression, statistical significance and overall standard deviation. If the concentration response curves do not give enough information for decision by expert judgement, additional testing or testing in a different concentration range should be performed.

For compounds that reduce the cell number or radial glia migration the viability is not used as unspecific endpoint. The reason is that an effect in both endpoints indirectly affects cell viability. In this case only the cytotoxicity is used for the classification.

PM upregulation

In contrast to the PM for downregulation, the PM for upregulation is based on a hit definition without the comparison of confidence intervals (CI) between the specific and unspecific endpoint. The reason is that specific and unspecific endpoint do not have the same relationship during an induction compared to a reduction in the endpoint. A loss in general cell health will likely result in an effect in e.g cell migration or differentiation, while an induction in cell health (measured as mitochondrial activity) does not necessarily increase these endpoints.

The following three hit classifications apply:

"no hit"	The compound is not defined as hit.
"specific hit":	The compound is defined as hit and the effect is no artifact due to loss in cell health.
"unspecific hit"	The compound is defined as hit in only the unspecific endpoints.

The compound is classified as hit, if the concentrations response curve generates a BMC and if the CI is within the test range. In case the CI spans above the test range, the compound is only classified as hit, if the highest test concentration is significantly different form the lowest test concentration. The adjusted significance ($p<0.05$) is thereby determined using a Tukey HSD test.

The decision process for the prediction model is described in the flow chart in figure 7.

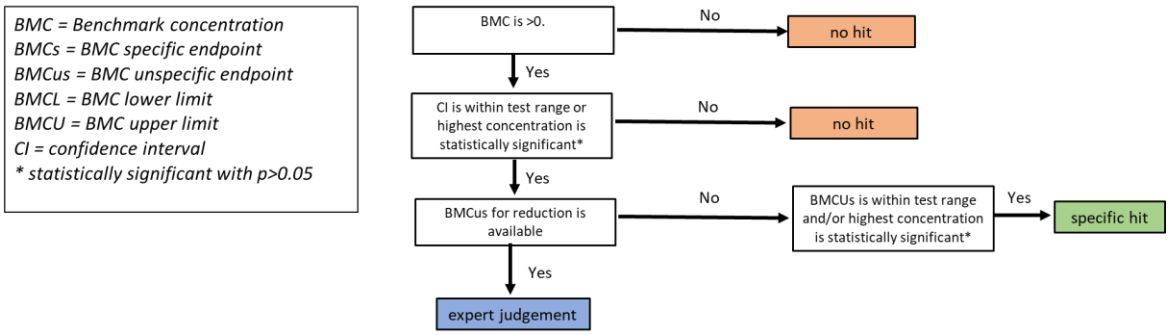


Figure 7: Decision tree for the PM for upregulation applied for test method NPC2-5. Overview of the decisions leading to the classification of a compound in one of three categories: "no hit", "specific hit", and "unspecific hit".

In case a compound is classified as hit together with a reduction in an unspecific endpoint, it needs to be clarified if the induction is a specific effect or an artifact due to a loss in cell viability or an increase in cytotoxicity. Therefore, expert judgement is applied which accounts for effect size, curve progression, statistical significance, overall standard deviation but also morphological changes of the spheres. Expert judgement is additionally applied if the confidence interval is very wide (BMCU/BMCL > 25) which means a high uncertainty for the BMC estimation.

8.3. Prediction model setup

The prediction model is set up as a statistical model which uses the 95 % confidence intervals (assessed based on the prediction bands around the concentration response curve) to determine the uncertainty of a hit definition. In the case of high uncertainty in the confidence interval, (e.g. because the CI spans

above the tested concentration range) the model additionally considers the statistical significance between the highest and the lowest test concentration. Next to the hit definition the uncertainty given by the 95 % CI is also considered in the specificity analyses (see 8.2 for a more detailed description).

The model has been tested on a set of 17 DNT negative- and 9 human DNT positive compounds (see 5.6 and 5.7). All negative compounds were correctly classified as "no hit" while 6 positive compounds were correctly classified as "specific hit". Here it is important to mention, that it is not expected that the NPC2-5 assay identifies all DNT positive compounds as not all of those compounds act via the mode-of-action (MoA) represented by the NPC2-5 assay. For most compounds the exact MoA for their neurodevelopmental adversity is not precisely known, yet other mechanisms like synaptogenesis and neuronal network formation are amongst the known DNT MoAs.

8.4. Test Performance

Table 1 summarizes the assay performance in terms of reproducibility of the assay.

Intra-experimental variation (SC) is the mean CV±SD of the CV of all replicates of the solvent control from each experiment across n>400 experiments.

Inter-experimental variation (raw) is the variability across all independent experiments (n>400) before normalization.

Inter-experimental variation (low conc.) is the variability across all independent experiments (n>400) after normalization based on the response of the lowest test concentration. It is assumed that the lowest test concentration does not affect any of the endpoints measured.

Inter-experimental variation (positive controls) is the variability of the respective positive controls across all independent experiments (n>40) after normalization.

Positive controls:

radial glia migration 72h (NPC2a) → PP2

neuronal differentiation (NPC3) → EGF

oligodendrocyte differentiation (NPC5) → BMP7

Table 3: Assay variability quantified as coefficient of variance (CV)

Endpoint	Intra-experimental variation (SC)	inter-experimental variation (raw)	Inter-experimental variation (low con.)	Inter-experimental variation (positive controls)
radial glia mig. 72h (NPC2a)	5.1 ±1.9%	11.3 %	5.2 %	15.5 %
radial glia mig. 120h (NPC2a)	5.6 ±2.3%	9.6 %	6.1 %	
neuronal mig. (NPC2b)	10.7 ±4.3%	19.8 %	11.5 %	
oligodendrocyte mig. (NPC2c)	9.4 ±4.9%	13.1 %	10.9 %	
cell number (NPC2c)	12.4 ±4.8%	23.3 %	14.4 %	
neuronal diff. (NPC3)	23.0 ±8.9%	48.4 %	30.5 %	15.8 %

neurite length (NPC4)		25.9 %	23.31	
neurite area (NPC4)		28.6 %	23.1 %	
oligodendrocyte diff. (NPC5)	35.1 ±15.9%	54.3 %	34.2 %	22.2 %
cytotoxicity [72h]	1.4 ±1.8%	24.4 %	8.6 %	
cytotoxicity [120h]	1.5 ±1.8%	37.6 %	6.6 %	
viability [120h]	8.4 ±3.3%	15.7 %	9.2 %	

Sensitivity and specificity of the NPC1 assay are determined based on a set of 9 predicted human DNT positive compounds and 17 predicted human DNT negative compounds (Masjosthusmann et al., 2020).

Based on this compound set the following performance parameters are obtained for the NPC2-5 assay.

Specificity: 100 %

Sensitivity: 66.7%

Here it is important to mention, that it is not expected that the NPC2-5 assay identifies all DNT positive compounds as not all of those compounds act via the mode-of-action (MoA) represented by the NPC2-5 assay. For most compounds the exact MoA for their neurodevelopmental adversity is not precisely known, yet other mechanisms like synaptogenesis and neuronal network formation are amongst the known DNT MoAs. It is therefore recommended that this assay is run as one part of an in vitro DNT battery.

8.5. *In vitro – in vivo* extrapolation (IVIVE)

Parameters for in vitro – in vivo extrapolation are not yet determined.

8.6. Applicability of test method

Toxicological applicability domain

The following compound classes have been tested successfully:

- Industrial chemicals
- cosmetics ingredients
- pharmaceuticals

Compounds need to be soluble in a solvent at a solubility where the solvent does not produce effects by itself in the test systems (5.7 for established solvents).

Compounds that are volatile or have a high lipophilicity have not been tested and might need more sophisticated exposure methods such as ‘passive dosing’.

Biological applicability domain

The neural progenitor cell migration and differentiation assay (NPC2-5) is based on primary hNPCs obtained from the human fetal cortex. As mentioned in 8.1 “Scientific principle”, the method depicts migration of radial glia, neurons and oligodendrocytes, neurite outgrowth and differentiation into neurons and oligodendrocytes from fetal hNPCs.

Next to the endpoints represented by this test method there are several other necessary neurodevelopmental endpoints which need to be studied using other test methods.

Neurodevelopmental processes not represented by this test method:

- Neural Crest Cell (NCC) Migration
- NPC apoptosis
- Neuronal morphology
- Synaptogenesis
- Neuronal network formation
- Neural Rosette Formation
- hiPSC-derived NPC proliferation
- hiPSC-NPC neuronal differentiation
- Neuronal subtype differentiation
- Astrocyte Differentiation and Maturation
- Astrocyte Reactivity
- Microglia reactivity
- Myelination

For a complete assessment of developmental neurotoxicity, the test method needs to be part of test battery.

The information on signaling pathways modulating the neurodevelopmental endpoints of the test method is summarized in Table 2. This describes the so far tested biological applicability domain of the assay.

Table 4: Signaling pathways studied in the test method.

DNT Assay	Signaling Pathway	Model Compound	Literature
NPC2a	SRC	PP2	Baumann et al., 2015; Masjosthusmann et al., 2018; Moors et al., 2007
	EGFR	AG1478, PD153035, EGF	
	PKC	PMA/Bis-I	
	NO-cGMP	7-NI, ODQ, Rp-8-Br-cGMP	Tegenge et al., 2011
	ROCK	Y27632	
	GSK3	CHIR	unpublished
	WNT	CHIR	
	RHO	Narciclasine	Masjosthusmann et al., 2020
	ETC complex I	Rotenone	
	PKC	Bis-I	
NPC3	Notch	DAPT	Baumann et al., 2015; Masjosthusmann et al., 2018
	EGFR	EGF, PD153035	
	RHO	Narciclasine	Masjosthusmann et al., 2020
	ETC complex I	Rotenone	
	THR	Triiodothyronine	unpublished
	RXR	Bexarotene	
	mTOR	Everolimus	
	GSK3	CHIR	
	WNT	CHIR	
NPC4	PKC	Bis-I	Masjosthusmann et al., 2020
	RHO	Narciclasine	
	WNT	CHIR	unpublished
	GSK3	CHIR	
	AKT	LY294002	
NPC5	BMP	BMP2, BMP7	Dach et al., 2017; Masjosthusmann et al., 2018
	-	Ascorbic Acid	
	Notch	DAPT	
	EGFR	PD153035	
	WNT	CHIR, IWP2	unpublished
	COX-2	Celecoxib	
	mTOR	MHY1485, Everolimus	
	GSK3	CHIR	
	CREB	cAMP	
	AKT	LY294002	
	PLC	m-3M3FBS	
	PDGFR	CP-673451	
	PPARβ/δ	GW0742	
	LXR	GW3965	
	PKC	Bis-I	Masjosthusmann et al., 2020
	RHO	Narciclasine	

8.7. Incorporation in test battery

To assess the hazard for developmental neurotoxicity it is recommended that this assay is used as one assay in a battery of assay (see 8.5 “Applicability of test methods”)

For the assessment of chemical action on the endpoints represented by this test method, the test method can be used as stand-alone test method.

The test method is currently used in the setup of a DNT test battery.

9. Publication/validation status

9.1. Availability of key publications

Key Publications concerning the test method are:

Koch et al. 2022

Klose et al., 2021

Masjosthusmann et al., 2020

Nimtz et al., 2019

Masjosthusmann et al., 2019

Masjosthusmann et al., 2018

Dach et al., 2017

Schmuck et al., 2017

Baumann et al., 2016

Jenny Baumann et al., 2016

Baumann et al., 2014

Fritsche et al., 2011

Moors et al., 2009

9.2. (Potential) linkage to AOPs

NPC2 is linked to the AOP "Disrupted laminin-beta1-integrin interaction leading to developmental neurotoxicity". The AOP is part of the OECD Project with the ID1.83 and the Coaching number C3-#8.

NPC5 is linked to an AOP on the binding to voltage gated sodium channels, which leads to impaired behavioral function (Hernández-Jerez et al., 2021).

9.3. Steps towards mechanistic validation

See:

3.3 Characterization and definition of source cells

4.6 Omics characterization of the test system

4.7 Features of the test system that reflect the in vivo tissue

8.5 Applicability of test method

9.4. Pre-validation or validation

To date, 123 unique compounds (as defined by unique DTXSIDs) have been tested successfully in this assay.

No formal OECD 34 validation study has been done (eg., ring trials with a standard set of known positive and negative controls).

9.5. Linkage to (e.g. OECD) guidelines/regulatory use

Test is not linked to regulatory guidelines.

10. Test method transferability

10.1. Operator training

For operators with a basic training in cell culture practices a four-week training period for handling of the test system and training in the assay is recommended. The operators should have basic understating in image analysis and data evaluation with respect to concentration response fitting.

10.2. Transfer

The test method has been used by multiple operators over a period of 18 month. However, inter operator variability was not been determined.

The test method is currently part of project that involves the lab to lab transfer and testing of a set of 35 chemicals.

11. Safety, ethics and specific requirements

11.1. Specific hazards; issues of waste disposal

No specific requirements.

11.2. Safety data sheet (SDS)

Reference to MSDS is given in the SOP (Appendix J in Masjosthusmann et al. 2020).

11.3. Specific facilities/licenses

No specific facilities are required.

No specific ethical approval is required.

11.4. Commercial aspects/intellectual property of material/procedures

There are no commercial aspects or intellectual properties to be considered.

12. References

Ayuso-Sacido, A., Moliterno, J.A., Kratovac, S., Kapoor, G.S., O'Rourke, D.M., Holland, E.C., García-Verdugo, J.M., Roy, N.S., Boockvar, J.A., 2010. Activated EGFR signaling increases proliferation, survival, and migration and blocks neuronal differentiation in post-natal neural stem cells. J. Neurooncol. 97, 323–37. https://doi.org/10.1007/s11060-009-0035-x

Bal-Price, A., Hogberg, H.T., Crofton, K.M., Daneshian, M., FitzGerald, R.E., Fritsche, E., Heinonen, T., Hougaard Bennekou, S., Klima, S., Piersma, A.H., Sachana, M., Shafer, T.J., Terron, A., Monnet-Tschudi, F., Viviani, B., Waldmann, T., Westerink, R.H.S.S., Wilks, M.F., Witters, H., Zurich, M.-G.G., Leist, M., 2018. Recommendation on test readiness criteria for new approach methods in toxicology: Exemplified for developmental neurotoxicity. ALTEX 35, 306–352. https://doi.org/10.14573/altex.1712081

Barenys, M., Gassmann, K., Baksmeier, C., Heinz, S., Reverte, I., Schmuck, M., Temme, T., Bendt, F., Zschauer, T.C., Rockel, T.D., Unfried, K., Watjen, W., Sundaram, S.M., Heuer, H., Colomina, M.T., Fritsche, E., 2016. Epigallocatechin gallate (EGCG) inhibits adhesion and migration of neural progenitor cells in vitro. Arch Toxicol. https://doi.org/10.1007/s00204-016-1709-8

Baumann, J., Barenys, M., Gassmann, K., Fritsche, E., 2014. Comparative human and rat "neurosphere assay" for developmental neurotoxicity testing. Curr. Protoc. Toxicol. 1, 1–24. https://doi.org/10.1002/0471140856.tx1221s59

Baumann, Jenny, Dach, K., Barenys, M., Giersiefer, S., Goniwiecha, J., Lein, P.J., Fritsche, E., 2016. Application of the Neurosphere Assay for DNT Hazard Assessment: Challenges and Limitations. Methods Pharmacol. Toxicol. 1–29. https://doi.org/10.1007/7653

Baumann, J., Gassmann, K., Masjosthusmann, S., DeBoer, D., Bendt, F., Giersiefer, S., Fritsche, E., 2016. Comparative human and rat neurospheres reveal species differences in chemical effects on neurodevelopmental key events. Arch. Toxicol. 90. https://doi.org/10.1007/s00204-015-1568-8

Chen, Y.-A., Lu, I.-L., Tsai, J.-W., 2018. Contactin-1/F3 Regulates Neuronal Migration and Morphogenesis Through Modulating RhoA Activity. Front. Mol. Neurosci. 11. https://doi.org/10.3389/fnmol.2018.00422

Dach, Katharina, Bendt, F., Huebenthal, U., Giersiefer, S., Lein, P.J., Heuer, H., Fritsche, E., 2017. BDE-99 impairs differentiation of human and mouse NPCs into the oligodendroglial lineage by species-specific modes of action. Sci. Rep. 7, 44861. https://doi.org/10.1038/srep44861

Dach, K, Bendt, F., Huebenthal, U., Giersiefer, S., Lein, P.J., Heuer, H., Fritsche, E., 2017. BDE-99 impairs differentiation of human and mouse NPCs into the oligodendroglial lineage by species-specific modes of action. Sci Rep 7, 44861. https://doi.org/10.1038/srep44861

Dergham, P., Ellezam, B., Essagian, C., Avedissian, H., Lubell, W.D., McKerracher, L., 2002. Rho Signaling Pathway Targeted to Promote Spinal Cord Repair. J. Neurosci. 22, 6570–6577. https://doi.org/10.1523/JNEUROSCI.22-15-06570.2002

Fahrion, J.K., Komuro, Y., Li, Y., Ohno, N., Littner, Y., Raoult, E., Galas, L., Vaudry, D., Komuro, H., 2012. Rescue of neuronal migration deficits in a mouse model of fetal Minamata disease by increasing neuronal Ca 2+ spike frequency. Proc. Natl. Acad. Sci. U. S. A. 109, 5057–5062. https://doi.org/10.1073/PNAS.1120747109/-/DCSUPPLEMENTAL

Förster, N., Butke, J., Keßel, H.E., Bendt, F., Pahl, M., Li, L., Fan, X., Leung, P. chung, Klose, J., Masjosthusmann, S., Fritsche, E., Mosig, A., 2021. Reliable identification and quantification of neural cells in microscopic images of neurospheres. Cytom. Part A. https://doi.org/10.1002/CYTO.A.24514

Fritsche, E., Gassmann, K., Schreiber, T., 2011. Neurospheres as a Model for Developmental Neurotoxicity Testing, in: Costa, L.G., Giordano, G., Guizzetti, M. (Eds.), In Vitro Neurotoxicology: Methods and Protocols. Humana Press, Totowa, NJ, pp. 99–114. https://doi.org/10.1007/978-1-61779-170-3_7

Gassmann, K., Abel, J., Bothe, H., Haarmann-Stemmann, T., Merk, H.F., Quasthoff, K.N., Rockel, T.D., Schreiber, T., Fritsche, E., 2010. Species-specific differential ahr expression protects human neural progenitor cells against developmental neurotoxicity of PAHs. Environ. Health Perspect. 118, 1571–1577. https://doi.org/10.1289/ehp.0901545

Gross, R.E., Mehler, M.F., Mabie, P.C., Zang, Z., Santschi, L., Kessler, J.A., 1996. Bone morphogenetic proteins promote astroglial lineage commitment by mammalian subventricular zone progenitor cells. Neuron 17, 595–606.

Hardy, A., Benford, D., Halldorsson, T., Jeger, M.J., Knutsen, K.H., More, S., Mortensen, A., Naegeli, H., Noteborn, H., Ockleford, C., Ricci, A., Rychen, G., Silano, V., Solecki, R., Turck, D., Aerts, M., Bodin, L., Davis, A., Edler, L., Gundert-Remy, U., Sand, S., Slob, W., Bottex, B., Abrahantes, J.C., Marques, D.C., Kass, G., Schlatter, J.R., 2017. Update: use of the benchmark dose approach in risk assessment. EFSA J. 15, 1–41. https://doi.org/10.2903/j.efsa.2017.4658

Hernández-Jerez, A., Adriaanse, P., Aldrich, A., Berny, P., Coja, T., Duquesne, S., Focks, A., Marinovich, M., Millet, M., Pelkonen, O., Pieper, S., Tiktak, A., Topping, C., Widenfalk, A., Wilks, M., Wolterink, G., Crofton, K., Hougaard Bennekou, S., Paparella, M., Tzoulaki, I., 2021. Development of Integrated Approaches to Testing and Assessment (IATA) case studies on developmental neurotoxicity (DNT) risk assessment. EFSA J. 19, e06599. https://doi.org/10.2903/j.efsa.2021.6599

Klose, J., Pahl, M., Bartmann, K., Bendt, F., Blum, J., Dolde, X., Förster, N., Holzer, A.K., Hübenthal, U., Keßel, H.E., Koch, K., Masjosthusmann, S., Schneider, S., Stürzl, L.C., Woeste, S., Rossi, A., Covaci, A., Behl, M., Leist, M., Tigges, J., Fritsche, E., 2021. Neurodevelopmental toxicity assessment of flame retardants using a human DNT in vitro testing battery. Cell Biol. Toxicol. https://doi.org/10.1007/S10565-021-09603-2

Mabie, P.C., Mehler, M.F., Marmur, R., Papavasiliou, A., Song, Q., Kessler, J.A., 1997. Bone morphogenetic proteins induce astroglial differentiation of oligodendroglial-astroglial progenitor cells. J. Neurosci. 17, 4112–20.

Masjosthusmann, S., Becker, D., Petzuch, B., Klose, J., Siebert, C., Deenen, R., Barenys, M., Baumann, J., Dach, K., Tigges, J., Hübenthal, U., Köhrer, K., Fritsche, E., 2018. A transcriptome comparison of time-matched developing human, mouse and rat neural progenitor cells reveals human uniqueness. Toxicol. Appl. Pharmacol. https://doi.org/10.1016/j.taap.2018.05.009

Masjosthusmann, S., Blum, J., Bartmann, K., Dolde, X., Holzer, A., Stürzl, L., Keßel, E.H., Förster, N., Dönmez, A., Klose, J., Pahl, M., Waldmann, T., Bendt, F., Kisitu, J., Suciu, I., Hübenthal, U., Mosig, A., Leist, M., Fritsche, E., 2020. Establishment of an a priori protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of

developmental neurotoxicity. EFSA Support. Publ. 17. https://doi.org/10.2903/sp.efsa.2020.en-1938

Masjosthusmann, S., Siebert, C., Hübenthal, U., Bendt, F., Baumann, J., Fritsche, E., 2019. Arsenite interrupts neurodevelopmental processes of human and rat neural progenitor cells: The role of reactive oxygen species and species-specific antioxidative defense. Chemosphere 235, 447–456. https://doi.org/10.1016/j.chemosphere.2019.06.123

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

Moors, M., Rockel, T.D., Abel, J., Cline, J.E., Gassmann, K., Schreiber, T., Shuwald, J., Weinmann, N., Fritsche, E., 2009. Human neurospheres as three-dimensional cellular systems for developmental neurotoxicity testing. Environ. Health Perspect. 117, 1131–1138. https://doi.org/10.1289/ehp.0800207

Nimtz, L., Klose, J., Masjosthusmann, S., Barenys, M., Fritsche, E., 2019. The neurosphere assay as an in vitro method for developmental neurotoxicity (DNT) evaluation, in: Neuromethods. Humana Press Inc., pp. 141–168. https://doi.org/10.1007/978-1-4939-9228-7_8

Schmuck, M.R., Temme, T., Dach, K., de Boer, D., Barenys, M., Bendt, F., Mosig, A., Fritsche, E., 2017. Omnisphero: a high-content image analysis (HCA) approach for phenotypic developmental neurotoxicity (DNT) screenings of organoid neurosphere cultures in vitro. Arch. Toxicol. 91, 2017–2028. https://doi.org/10.1007/s00204-016-1852-2

Annex III – ToxTemp UKN2 (cMINC) assay

1. Overview

1.1 Descriptive full-text title

Assay to test impairment of migration of human neural crest cells (cMINC; UKN2) – V2.0

1.2 Abstract

This in vitro test method uses human neural crest cells (NCCs) generated from induced pluripotent stem cells (iPSC). It assesses disturbances of NCC migration during fetal development. The number of migrated NCC, as well as cell viability are measured simultaneously using high content imaging. Thereby the processes of cell migration and cell death are measured in cells exposed to potential toxicants for 24 h. The data of this method are meant to predict developmental disorders and malformations e.g. neural tube defects or craniofacial malformations caused by compound exposure during fetal development. The method has a well-established prediction model, but it has not undergone formal validation and it has not been part of a ring trial. It has been used in the screening of several medium-sized compound libraries. According to the readiness criteria as published by Bal-Price et al. (2018) the neural crest cell migration assay obtained the readiness score A-.

2. General information

2.1 Name of test method

Circular migration inhibition of NCC (cMINC) assay, UKN2

2.2 Version number and date of deposition

This is Version 2.0 of the protocol "Assay to test impairment of migration of human neural crest cells (cMINC; UKN2) – V2.0). It was assembled and deposited in February 2022. A previous version was assembled in 2019 in the context of the EU-ToxRisk project (see publication Krebs et al., 2020)

2.3 Summary of introduced changes in comparison to previous version(s)

Changes compared to V1 refer mainly to the generation of the test system and the cell line used. Test procedures and parameters remain unchanged.

2.4 Assigned data base name

Normal text names often do not uniquely define the method. Therefore, each method should be assigned a clearly and uniquely defined data base name.

Here this is exemplified by the names generated in the EU-ToxRisk project:

UKN1a_DART_NPC_Diff_6D_02

UKN1b_DART_NPC_Diff_4D_01

UKN2a_DART_NC_Migr_24h_04

The name is assembled (in more generic terms) from the following elements:

Axa_B_C_D_E

Axa: mandatory part of the identifier allowing unambiguous identification

A: Abbreviation / acronym of the partner depositing the assay

x: Consecutive number (referring to the partner's assay number)

a: Sub-specifier (for variants, i.e. very similar assays, but e.g. different readout or medium); not mandatory, but 'Axa' must be specific (i.e. clearly identifying) for each assay variant

B: Indication of the main intended use (max 5 letters), e.g. DART, Neuro, Liver, Lung, Renal, Redox, Stress...

C: Specifier for test system, e.g. cell type, e.g. NPC (neural precursor cells), NC (neural crest), Hep (liver cells), REN (kidney cells), PUL (lung cells) (max. 4 letters)

D: Identification of test endpoint, e.g. Diff_6D = Differentiation for 6 days; exp_24 h = exposure for 24 hours; RNA_6h = transcriptome after 6 hours; and so on (use max. 15 signs altogether; if desired in 2-3 blocks), name (and acronym) of the project partner home organisation.

E: version number.

UKN2a_DART_NC_cMIGR_24h_02

2.5 Name and acronym of the test depositor

University of Konstanz (UKN), Germany

2.6 Name and email of contact person

Prof. Dr. Marcel Leist
marcel.leist@uni-konstanz.de
Tel: +49-7531885037

2.7 Name of further persons involved

Xenia Dolde (PhD student, experimenter)
xenia.dolde@uni-konstanz.de
Jonathan Blum (PhD student)
jonathan.blum@uni-konstanz.de

Heidrun Leisner (experimenter)
heidrun.leisner@uni-konstanz.de

2.8 Reference to additional files of relevance

- An important reference is the DB-ALM Protocol n° 195 → The original iPSC are in the meantime cultured feeder-free (see 3.2)
- Raw data file
- Data processing file

3. Description of general features of the test system source*

Note: *this section might be redundant with section 4. It is meant to describe the procedure of generating the cells, which are eventually used in the test method. This applies e.g. for stem cells, which have to be differentiated towards the cell type with which the method is conducted (e.g. neurons or hepatocytes derived from iPSC). See scheme for illustration. If this is not applicable to your test system, go directly to section 4.*

Overview of cell culture stages relevant to their use in an *in vitro* test method

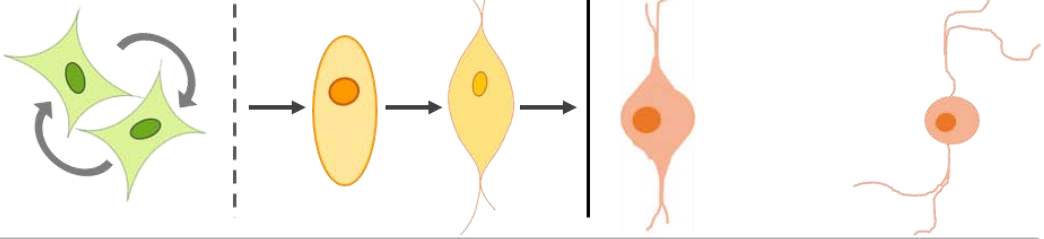
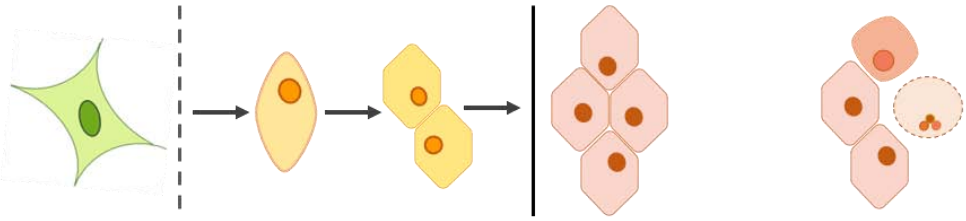
Stage:	Original cell source	Test system generation / maturation	Test run: start of test	Test run: end of test
Chapter:	3.1 – 3.5	3.6	4.1 + 4.2	4.4 – 4.8 / 4.3
				
Example 1:	iPSC	generation of neurons	neuronal maturation	
Example 2:	HEK 293	re-plating	proliferation	
				
Example 3:	primary hepatocytes	spreading, partial de-differentiation, bile duct formation	spontaneous cell death, de-differentiation	
Example 4:	primary astrocytes	re-plating, secondary culture	spontaneous activation, metabolic shifts	
Processes:	maintenance, isolation, thawing	differentiation, proliferation, aggregation, self-organization, re-plating, maturation, de-differentiation	maturation, proliferation, cell death, de-differentiation / differentiation	

Figure 13: Overview of test system stages and where to find / deposit corresponding information in this document.

Note: Refer to overview figure to connect question numbers and cell culture stages.

3.1 Supply of source cells

The human induced pluripotent stem cell (hiPSC) line IMR90_clone_#4 has been bought from WiCell, Wisconsin in 2012 and a masterstock has been frozen. From the masterstock several working stocks have been prepared. The working stocks are regularly thawed and can be continously maintained due to self-renewal and pluripotency capabilities of the cells. The cells are maintained up to 10 passages before a new vial of the working stock is thawed.

3.2 Overview of cell source component(s)

Undifferentiated hiPSC cells (IMR90, WiCell) are maintained as monoculture on Laminin-521 coating in essential 8 (E8) medium. The cells grow in colonies, and are split weakly. The cells show self-renewal and pluripotency characteristics (regular testing). The cells can be differentiated into several different cell types.

3.3 Characterization and definition of source cells

- **ATCC number**: CCL-186
- **Origin**: Homo sapiens, human
- **Tissue**: lung
- **Cell type**: fibroblast
- **Gender**: female
- **Morphology**: fibroblast
- **Culture properties:** adherent
- **Disease:** no disease was diagnosed
- **Age:** 16 weeks gestation
- **Ethnicity:** Caucasian
- **Expression:** Cells express the pluripotency markers Oct4, nanog and Tra-1-60
- **STR signature:** Cell line identity by STR conforms 100% loci homology:

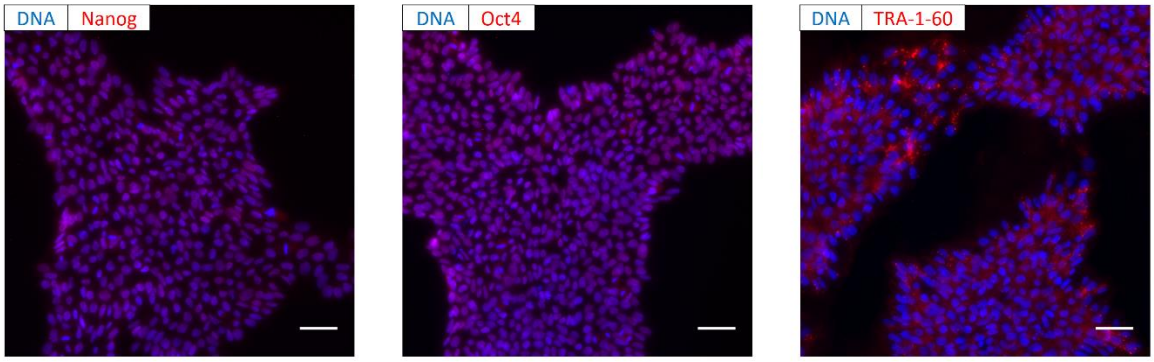
	ATCC (CCL-186)	University of Konstanz; Dept. of In-Vitro-Toxikology and Biomedicin
Cell line	**IMR-90**	**IMR90**
Date		09.11.2018
D5	12	12
D5'	13	13
D13	11	11
D13'	13	13
D7	9	9
D7'	12	12
D16	10	10
D16'	13	13
vWA	16	16
vWA'	19	19
TH01	8	8
TH01'	9,3	9,3
TPOX	8	8
TPOX'	9	9

CSF1	11	11
CSF1'	13	13
Amel	X	X
Amel'	X	X

3.4 Acceptance criteria for source cell population

The cells have to be pathogen-free (regular testing for mycoplasma).

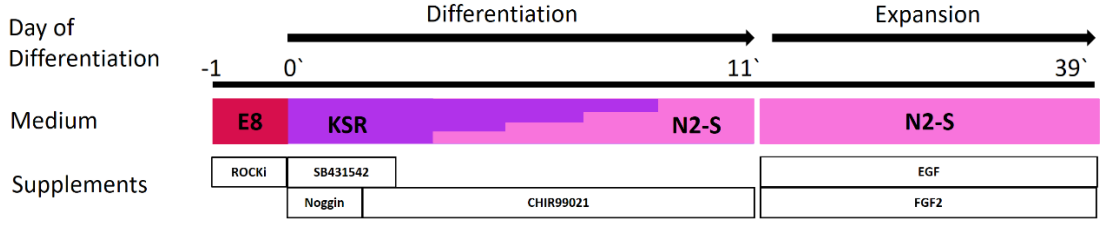
The iPSC maintenance is regularly checked for expression of pluripotency markers (Oct4, Nanog, Tra-1-60) by immunocytochemistry.



3.5 Variability and troubleshooting of source cells

- hiPSC can be maintained up to 10 passages, high passage number might influence performance of cells
- Too little or too high cell density leads to detachment of cells or spontaneous differentiation
- If cells start to differentiate, cells should be discarded immediately
- Cells have to be maintained as colonies and not as single cells. Therefore splitting should be performed as fast as possible
- Batch effects of critical additives (e.g. holo-transferrin, TGF-β) can lead to differentiation of cells at low passage number

3.6 Differentiation towards the final test system



NCCs were differentiated from hiPSCs following the modified protocol of Mica et al. (2013). Therefore, IMR90 iPSCs were plated on Matrigel coated 6-well plates at a density of 100'000 cells/cm^2 in E8 medium containing 10 µM ROCK-inhibitor (Y-27632). After one day cells reached a confluency of 80-100% and differentiation was initiated (day 0') by a medium change to KSR medium (Knock out DMEM, 15% knock out serum replacement, 1% GlutaMax, 1% MEM NEAA solution, 50 µM 2-mercaptoethanol) supplemented with 20 ng/ml Noggin and 10 µM SB431542. From day 2 on cells were treated with 3 µM CHIR 99021. Noggin and SB431542 were withdrawn at day 3 and 4, respectively. Beginning at day 4, the KSR medium was gradually replaced with N2-S medium (DMEM/F12, 1.55 mg/ml glucose, 1% GlutaMax, 0.1 mg/ml apotransferin, 25 µg/ml insulin, 20 nM progesterone, 100 µM putrescine, 30 nM selenium) in 25% increments. Cells were collected at day 11, resuspended in N2-S medium supplemented with 20 ng/ml EGF and 20 ng/ml FGF2) and seeded as droplets (10 µl) on poly-L-ornithine (PLO)/Laminin/Fibronectin coated 10 cm dishes. Cells were expanded by weekly splitting. From now on seeding as droplets was not necessary and medium was changed every second day. After 35-39 days, cells were cryopreserved at a concentration of 4*10^6 cells/ml in 90% N2-S medium and 10% dimethyl sulfoxide (DMSO) (Merck Millipore) until further use.

3.7 Reference / link to maintenance culture protocol

DB-ALM Protocol n° 195

4. Definition of the test system as used in the method

Note: *this section refers to the stage of the test system, which is then used for the test method. See scheme for illustration. If you have cells that do not need prior differentiation, give their basic characteristics here. If your test system undergoes significant changes between the maintenance culture and the use for testing, please also fill in section 3.*

4.1 Principles of the culture protocol

A highly homogeneous pre-differentiated population of neural crest cells is added to coated wells. In the middle of the well is a silicone stopper that prevents cells from settling in a circular area in the middle of the well. The cells are kept viable and alive by the presence of EGF and FGF in the medium, and they become adherent overnight. When the stopper is removed, the cells form a dense monolayer in the culture dish, with a sharply demarcated circular area in the middle that is free of cells. Due to their natural spontaneous migration behaviour, the cells move into the cell-free area. The cells are still in a proliferative state. Proliferation at the edge of the cell free area contributes to a small extent to "apparent migration behaviour". This has been characterized and quantified in detail (PMID: 27463612). It is accounted for by counting cells in a narrower circle than the original circular area.

4.2 Acceptance criteria for assessing the test system at its start

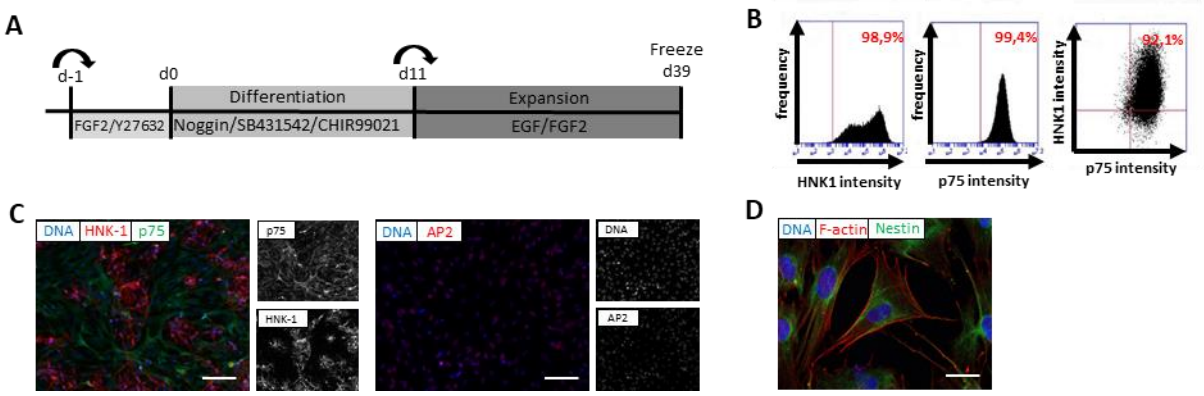


Figure 14: Characterization of the cellular system.
(A) Differentiation schema of NCCs from human induced pluripotent stem cells. (B) Expression of the NCC markers HNK-1 and p75 was monitored by FACS analysis. (C) After thawing the cells were immunofluorescently labelled for the typical NCC markers p75, HNK1 and AP2. Scale bar: 50 µm.
(D) Typical mesenchymal like cell morphology of NCCs. Cells were double-stained for nestin and f-actin. Scale bar: 10 µm.

The derived NCCs are ≥ 90% positive for the NCC marker human natural killer-1 (HNK1) and the nerve low affinity nerve growth factor receptor p75. Additionally, the cells express the stem cell microfilament protein nestin and AP2. On the other hand, none of the cells show expression of the astrocyte marker GFAP, the central nervous system precursor cell marker Pax6 and the neuronal marker βIII tubulin. Furthermore, the functional capability of the cells is checked by performing the cMINC assay with endpoint specific controls. The cells have to be pathogen-free to be used in further experiments (regular testing for mycoplasma after thawing). After thawing the viability should be > 90%.

4.3 Acceptance criteria for the test system at the end of compound exposure

After compound treatment, the negative controls should fulfil the following:
- cells should be migrated into the cell free area
- cell viability should be > 90%

4.4 Variability of the test system and troubleshooting

Causes of variability:

- High passage number of iPSC maintenance might influence NCC differentiation
- Too little cell density at start of differentiation can cause problems (confluency should be 80-100%)
- The freezing process should be performed as fast as possible. The number of viable cells after thawing can be decreased due to slow freezing.

4.5 Metabolic capacity of the test system

No specific information available.

4.6 Omics characterization of the test system

Transcriptomics data (unpublished) will become available from the originator lab (Leist) upon request

4.7 Features of the test system that reflect the in vivo tissue

- The cells are able to migrate mostly as single cells; they show some evidence of collective migration
- Cells show invasive behaviour in a 3D environment (e.g. transwell assay)

4.8 Commercial and intellectual property rights aspects of cells

The cells are not protected by patents or any other licences.

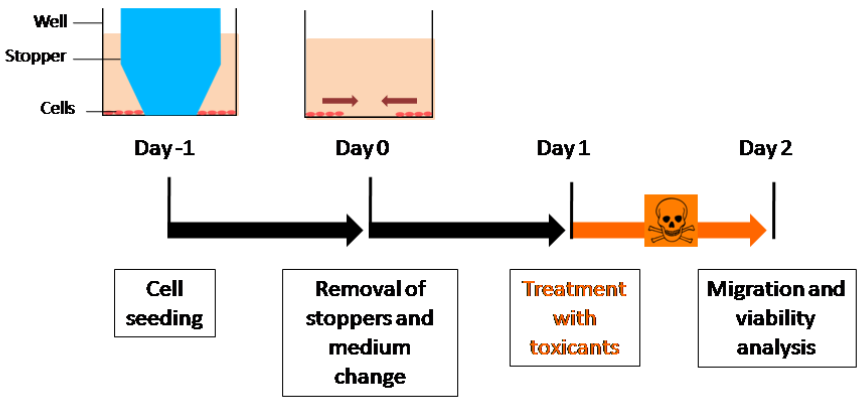
4.9 Reference / link to the culture protocol

Chapter 3 has been answered.
The maintenance is described in the DB-ALM SOP n° 195 available at:
https://ec.europa.eu/jrc/en/scientific-tool/database-alternative-methods-animal-experimentation

A lab-internal handling protocol is also available upon request to the Leist-lab.

5. Test method exposure scheme and endpoints

5.1 Exposure scheme for toxicity testing



Day -1: Silicone stoppers are placed into the wells of a 96-well plate. The wells are coated with PLO/laminin/fibronectin. Cells are seeded around the stoppers and allowed to attach.

Day 0: The silicone stoppers are removed and the medium is replaced with pre-warmed, fresh N2-S medium containing the cytokines EGF and FGF. Cells are allowed to migrate into the cell free area.

Day 1: 25 µl of the 5x concentrated toxicants are added to the wells. The cells plus toxicants are incubated for 24 h.

Day 2: Cells are stained with calcein-AM and H-33342 for 30 min before imaging with a high content imaging microscope. Quantification of migration and viability is done by high content imaging analysis.

5.2 Endpoint(s) of the test method

Specific Endpoint: Migration inhibition

Reference endpoint: Cell viability

5.3 Overview of analytical method(s) to assess test endpoint(s)

Migration inhibition: NCCs are plated around silicone stoppers in a culture dish and are allowed to migrate into the cell free area upon removal of the stopper. The number of migrated cells into the cell free zone is quantified 24 h after toxicant treatment. Migration inhibition of NCCs after treatment with toxicants is measured relative to control conditions (untreated control cells). For the quantification the cells are stained with calcein-AM and H-33342 for 30 min at 37 °C. The center of the well (migration zone) is imaged in four tiles with a 5x lens. Afterwards the four images are stitched together to obtain one image. For migration quantification, a software tool has been developed (http://invitrotox.uni-konstanz.de/). With the help of this software the previously cell-free area can be estimated and the number of H-33342 and calcein double-positive cells in the region of interest (ROI) can be counted. The diameter of the ROI was chosen so that 150 to 300 cells were in the ROI in untreated conditions. An Excel table containing the number of viable cells in the ROI of all wells of the plate is generated by the software.

Cell viability: Cell viability is measured after 48 h outside the migration zone in the same well. The cells are stained with calcein-AM and H-33342 and four fields outside the migration zone are imaged with a 10x lens. Viability is defined as the number of H-33342 and calcein double-positive cells, viable cells are determined by an automated algorithm. An excel file is generated with the number of viable cells in each well.

Migration and Viability are normalized to untreated controls.

5.4 Technical details (of e.g. endpoint measurements)

Quantification of migration

An automated microplate reading microscope (Array-ScanII HCS Reader, Cellomics, PA) equipped with a Hamamatsu ORCA-ER camera (resolution 1024 x 1024; run at 2 x 2 binning) was used for image acquisition. Four fields per well were imaged. Images were recorded in 2 channels using a 5x objective and excitation/emission wavelengths of 365 ± 50/535 ± 45 to detect H-33342 in channel 1 and 474 ± 40/535 ± 45 to detect calcein in channel 2. Pictures were exported from the microscope and for migration quantification, a software tool has been developed (http://invitrotox.uni-konstanz.de/). With the help of this software the previously cell-free area can be estimated and the number of H-33342 and calcein double-positive cells in the region of interest (ROI) can be counted. The diameter of the ROI was chosen so that 150 to 300 cells were in the ROI in untreated conditions. An Excel table containing the number of viable cells in the ROI of all wells of the plate is generated by the software.

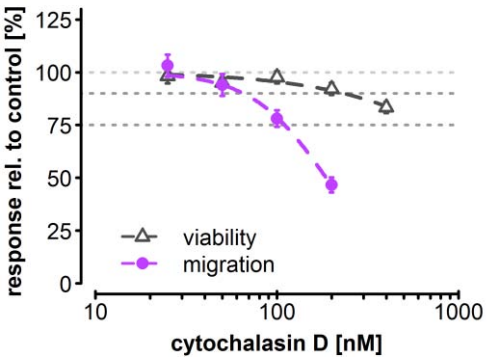
Quantification of individual viable cells by imaging

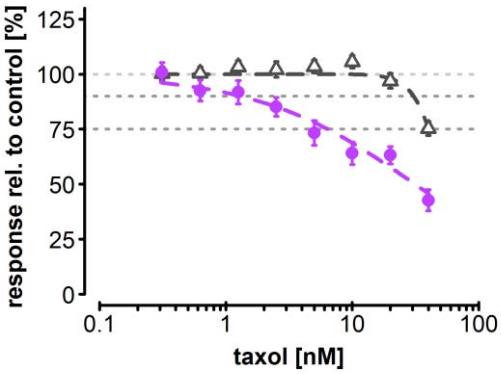
For a quantitative assessment of viable cells, the same images used to assess migration were analyzed using another image analysis algorithm: nuclei were identified in channel 1 as objects according to their size, area, shape, and intensity. Nuclei of apoptotic cells with increased fluorescence were excluded. A VCSA was defined around each nucleus by expanding it by 0.3 µm into each direction. Calcein-AM staining, labeling live cells, was detected in channel 2. The algorithm quantified the calcein intensity in the VCSA areas. Cells having an average calcein signal intensity in the VCSAs below a predefined threshold were classified by the program as “not viable”. Valid nuclei with a positive calcein signal in their cognate VCSA were counted as viable cells. A positive calcein signal was based on measurements of the average intensity (normal cells: 1300 ± 115, threshold: < 50) and the total integrated intensity (normal cells: 186,000 ± 23,600, threshold < 1000) of cells.

5.5 Endpoint-specific controls / mechanistic control compounds (MCC)

Example 1: Cell migration requires dynamic variability of the cytoskeleton, e.g. actin reorganisation. Cytochalasin D is a known inhibitor of actin polymerisation and therefore inhibits cell migration at non-cytotoxic concentrations.

Example 2: Migration depends on the dynamic instability of microtubules in the leading edge of a migrating cell. Taxol inhibits breakdown of microtubules and leads to migration inhibition at non-cytotoxic concentrations.





5.6 Positive controls

Positive control: Cytochalasin D (200 nM), LiCl (10 mM), taxol (10 nM)

5.7 Negative and unspecific controls

Negative control: solvent (0.1% DMSO final concentration), paracetamol, ASS

5.8 Features relevant for cytotoxicity testing*

Cells still proliferate after thawing and proliferation can interfere with the migration result. About 30% of the cells undergo mitosis during the assay.

5.9 Acceptance criteria for the test method

A test is discarded if the positive control did not inhibit migration, e.g. if the cell number in the migration zone is ≥ 75% of control.

If the negative control decreases migration ≥ 10%, the test is discarded.

5.10 Throughput estimate*

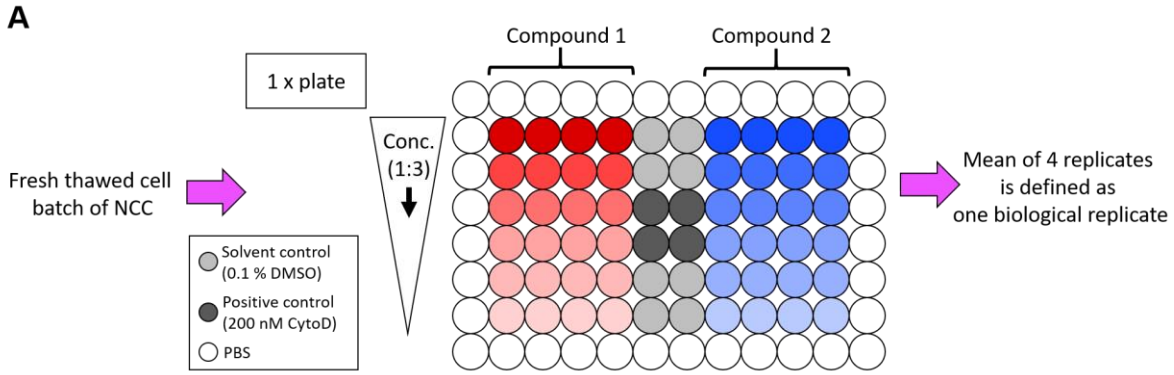
Data point = one biological replicate (→ usually 4 technical replicates); each concentration/condition of a compound counts as data point

768 data points per month

2 compounds per plate, 6 different concentrations of each compound per plate (see figure) → 12 data points (1 plate)

16 plates can be done per week → correlates to 32 compounds → 192 data points per week
4 weeks per month → 768 data points per month.

Typical plate layout:



6. Handling details of the test method

6.1 Preparation / addition of test compounds

- Compounds are stored according to the manufacturer's instructions (e.g. 4°C, room temperature, -20°C).

- Preferable solvent is DMSO. The used DMSO is stored in a lightproof, air-tight bottle at room temperature.

- Final DMSO concentration on the cells is 0.1%

- After dissolving the compounds which are delivered in a solid/powder form, all compound solutions are aliquoted into volumes sufficient for one experiment (i.e. one biological replicate). In this way repeated freezing and thawing and therefore damaging the compound's stability and efficiency can be avoided.

- For conducting an experiment, a compound aliquot is thawn and diluted with 'DMEM/F12 Advanced' without supplement in a separate deepwell-plate.

- All compound dilutions in the master plate contain 0.5 % DMSO, so that a final concentration of 0.1 % DMSO is reached on the cells. The highest compound concentration is diluted with medium 1:200 without DMSO as 0.5 % is already reached with the DMSO the compound is solved in, the serial dilution is done with N2-S medium supplemented with 20 ng/ml EGF and FGF and 0.5 % DMSO.

- 25 µl of the 5x concentrated compound dilutions are added to the cells using a multichannel pipette.

6.2 Day-to-day documentation of test execution

Plate maps are defined prior to the experiment and documented in the lab book and files (Excel files) are stored on the work group server.
Concentrations and compound dilutions are calculated prior to the experiment.
Experimental procedures are noted manually in a paper lab book.

6.3 Practical phase of test compound exposure

The experimenter plans the experiment according to Cellomics microscope availability (has to be booked in advance).

Pipetting errors are marked directly on the plate maps and are documented in the lab book.

The paper lab book is taken to cell culture rooms and errors are documented in there right away.

The technical replicates were pipetted from left to right. The highest concentration is located at the top row.

6.4 Concentration settings

2 compounds per plate with 6 concentrations

As default a serial dilution 1:3 is used, i.e. a concentration range from e.g. 100 µM → 0.4 µM.

Serial dilutions of compounds are prepared in a separate deepwell-plate, from which 25 µl are transferred to the according plates with attached cells using a multichannel-multistepper pipette. N2-S medium supplemented with 20 ng/ml EGF and FGF is used for dilution.

Dilution steps can be adapted to be more narrow (e.g. 1:1.5)

6.5 Uncertainties and troubleshooting*

- Compound solubility in stock and during dilution is too low (stock solved in 100% DMSO, final concentration of the solvent on the cells is 0.1% DMSO)
- Some compounds show autofluorescence and interfere with the detection of calcein-AM or H-33342.
- To prevent negative edge effects, only the inner 60 wells of a 96-well plate are used and the edge wells were filled with PBS.
- Focusing failure of Array Scan VTI HCS Reader (Cellomics, PA) can be a problem that produces outliers; as well as imaging only one channel.
- Highly trained/automated handling with multichannel and multistepper pipette is necessary to achieve little variance.

- Different cell batches vary in the cell number due to freezing conditions and the differing proliferation rates of the lots. The variation between the plates of one experiment is lower than 5% and the variation between different experiments is lower than 20%. The results were always normalized to untreated controls.
- Operators can get trained within 2-4 weeks. Cell seeding and medium change should be performed as fast as possible to keep cells as short as possible at room temperature. The more practice an operator has, the faster the critical steps can be performed.
- Substances are added when pipette tips are touching the wall of the wells right above the medium surface. When the substance solution is pipetted too high above the medium surface, the droplet may just stick to the wall of the well without flowing down into the medium.

6.6 Detailed protocol (SOP)

Protocol n° 195 in DB-ALM data base
See the SOP (Appendix K in Masjosthusmann et al. 2020).

6.7 Special instrumentation

- The method requires a Cellomics Array Scan VTI HCS high content reader that may not be present in the standard lab.
- Silicone stoppers (Platypus Technologies, Madison, Wi, US).

6.8 Possible variations*

a) further additional endpoints:
- EdU staining can be performed within the assay to measure proliferation
b) other analytical endpoints:
cell viability by:
- fluorimetric measurement of resazurin conversion
- measurement of extracellular LDH
- measurement of luminescence indicating ATP content
c) other exposure:
- The exposure time for toxicants can be increased up to 48 h. But longer exposure time increases the effect of toxicants on viability and cell proliferation.

d) experimental variation:
- AraC addition to prevent proliferation effects on migration result. About 30% of the cells divide during the 24 h of the assay period. Therefore a cell proliferation inhibiting compound would reduce the cell number by 25% in the migration zone and therefore results in 80% viability and 75% migration measurement. But a reduction of migration by > 25% is unlikely to be explained by effects on proliferation (Nyffeler J. et al., 2016. PMID: 27463612).
- other cells may be used (derived from different iPSC)

6.9 Cross-reference to related test methods*

The scratch assay is another method to analyse cell migration. In comparison to the cMINC assay the throughput of the scratch assay is low (Zimmer B. et al., 2012. PMID: 22571897).

7. Data management

7.1 Raw data format

Raw data is extracted by copy-paste in Excel files (example file available).

Data from all technical replicates are collected in one file.

7.2 Outliers

1. Mathematical procedures to define outliers have not been defined. Data points that 'look' very far off are discarded. Biological outliers do practically not exist, most far data points are the result of technical problems (focus not found, only one channel imaged, etc.)
2. All raw data (incl. outliers) are stored.
3. Technical outliers make up 1-0.1%.

7.3 Raw data processing to summary data

- Array Scan VTI HCS Reader (Cellomics, PA) takes images (optionally bitmap or tiff-format; 512 x 512 pixels, 8bit or 16bit)

- For migration quantification images from the Array Scan VTI HCS Reader are exported in 8-bit tiff format and loaded to the "Ringassay" software (http://invitrotox.uni-konstanz.de/) for calculating the number of migrated cells.

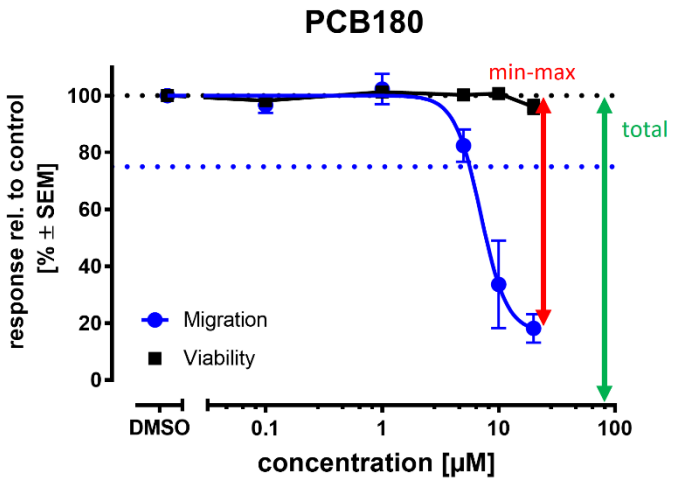
- Images are locally analyzed using the Array Scan software, algorithm quantifies cell count (viable cells)

- data are copy-pasted into an Excel sheet, further analysis is done with Excel + GraphPad Prism

7.4 Curve fitting

The data are analyzed with Excel and represented with GraphPad Prism.
For the concentration curve, a nonlinear regression fit is calculated. The fitting method is least squares. If a non-linear curve fit is not possible, a linear curve fit is performed. The curve deriving from the fit is a 4-parameter log function. To calculate the EC25 value, this log-function is solved for y=25% of the total scale, not for 25% of the min-max scale (see example below). Treated concentrations are analyzed for deviation from control. Sometimes it is analyzed whether the deviation of migration is different from the deviation of viability. This is done by two-way ANOVA + Tukey-Kramer post hoc testing. Statistics applied are one-way ANOVA (and nonparametric) with Dunnett's post test.

BMC values with their upper and lower confidence intervals (BMCU and BMCL) are calculated via the publically available online software:
http://invitrotox.uni-konstanz.de/BMCeasy/



7.5 Internal data storage*

The data are firstly stored on the microscope computer and then exported to other servers (lab group server and university server), which are back-upped regularly.

7.6 Metadata

The metadata are documented, stored and exported as text document (log)-files to the according scheme: (local PC)_descriptor(date and time)_XXX.log:

The following metadata are stored:

- cellinsight-pc_160429130003_AutomationControllerIni
- cellinsight-pc_160429130003_kineticprotocol
- cellinsight-pc_160429130003_protocol
- cellinsight-pc_160429130003_scan
- cellinsight-pc_160429130003_ScanIni
- cellinsight-pc_160429130003_spooling
- cellinsight-pc_160429130003.spooled

7.7 Metadata file format

Metadata files are available.

8. Prediction model and toxicological application

8.1 Scientific principle, test purpose and relevance

Migration of NCCs is an essential process during fetal development. Impaired NCC migration triggered genetically or by toxicants can lead to malformations and disorders e.g. Hirschsprung's disease or Treacher-Collins syndrome.

The cMINC assay models the effects of NCC migration under toxicant exposure.

The test captures endpoints like spina bifida or cleft palate also measured during developmental toxicity regulatory studies.

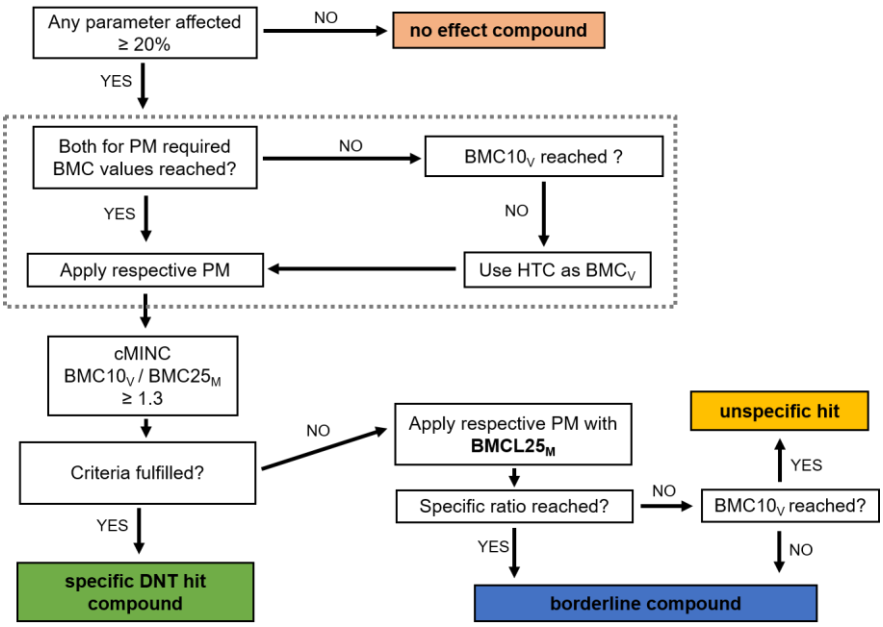
8.2 Prediction model

Three different models are used:

1. prediction model for screening:
hit = inhibition of NCC migration while viability is not changed:
Migration ≤ 80% of DMSO control
Viability ≥ 90% of DMSO control

2. prediction model for compound hazard evalaution:
hit confirmation testing: EC10 Viability (V) / EC 25 Migration (M) ≥ 1.3 → specific migration inhibitor of NCCs

3. prediction model for borderline compounds:
A ratio of BMC10 Viability (V) / BMC**L**25 Migration (M) ≥ 1.3 is considered a borderline hit. In some scenarios the viability does not reach the BMC10 Viability (V) necessary for the ratio calculations. In this case the highest tested concentration (HTC) was used. Schematic representation of the complete prediction model is shown in scheme below.



8.3 Prediction model setup

a) The prediction model was established using the following compounds (Nyffeler et al. 2016):
- acrylamide
- As_2O_3
- $CdCl_2$
- LiCl
- PCB180

- retinoic acid
  ➔ Positive controls
- cytochalsin D
- taxol
- colchicine
  ➔ Endpoint-specific controls
- $AgNO_3$
- AraC
- aphidicolin
- L-homocysteine
- MG-132
- staurosporine
- triton X-100
  ➔ Unspecific toxicants

b) The prediction model has been applied to screen 80 compound library of NTP. The prediction model including the borderline classification has been applied to screen a 120 compound library (Masjosthusmann et al., 2020).

c) The process is documented in Nyffeler et al. 2016

d) Sensitivity and specificity require still definition of a good standard and has therefore not been done.

8.4 Test performance

Some background on the test performance is given in chapters 8.2/8.3 (prediction model).

Several performance parameters for the test were obtained in several separate evaluation rounds.

A first evaluation was done during the first publication of the model and its applications (Nyffeler et al. 2016). Here, a panel of well-selected positive and negative controls have been tested. Accordingly, the specificity was 100% and the sensitivity was > 90 %. In dedicated experiments, S/N ratios of > 20 and a z' of > 0.5 have been determined. The compound used as positive control cytochalasin D was run across 35 different assay plates. The migration percentage relative to the solvent control varied between 30 and 70% across all plates (Nyffeler et al. 2016, Supplementary S2).

Later, the test has been used in screening campaigns, and real-live performance data under broader screen conditions have been obtained. The different performance data need to be considered, when a compound is a hit in a screen, or whether it has been specifically evaluated in a hit follow-up or a mechanistic project.

A first screen application has been the NTP80 screen (80 compounds provided by the US NTP). Data are published in Nyffeler et al. 2017.

A second screen application has been the cross systems case study of the EU-ToxRisk project. The baseline variation is indicated in Krebs et al., 2020. Moreover, an overview is given for 19 compounds on the BMC/BMCL ratio as measure of readout certainty.

A third screen was performed in the context of the EFSA DNT test battery evaluation with 120 compounds (Masjosthusmann et al. 2020). From this screen the following performance indicators were obtained:

A: **Sensitivity: 100%**

→ With cMINC as standalone assay in 17 `tool negatives´ tested (Masjosthusmann et al. 2020).

B: **Specificity: 82.7%**

→ With a selected set of 27 positive compounds with evidence for DNT (Masjosthusmann et al. 2020).

C: **Baseline variation (intra-experimental)**

Migration: 14.8 ± 4.3%

Cell viability: 6 ± 3%

D: **Baseline variation (inter-experimental)**

Migration: 15.6%

Cell viability: 9.7%

E: **Variation of a positive control run on each (inter-experimental)**

Migration: 32.2%

Definition of values C-E

C: **Baseline variation (intra-experimental)** is the mean CV±SD of the CV of all replicates of the solvent control from each experiment across n>200 experiments.

D: **Baseline variation (inter-experimental)** is the variability across all independent experiments (n>200) after normalization based on the response of the lowest test concentration. It was assumed that the lowest test concentration does not affect any of the endpoints measured.

E: **Variation of a positive control run on each (inter-experimental)** is the variability of the positive control across all independent experiments (n>40) after normalization. Example for a positive control that on average reduced the specific endpoint down to 40% (relative to solvent control) and a calculated variability of 50%: 0.5 x 40% = +/- 20% → The positive control with mean of 40% varies from 20% to 60%.

8.5 In vitro – in vivo extrapolation (IVIVE)

1. Estimated lipid content and albumin concentration in in vitro test media and human plasma:

Medium	Lipid content (ml/ml)	Albumin concentration (µM)
UKN2	2.8E-6	5.6
Human plasma	6000	600

2. A three-step (physiology-based) pharmacokinetic (PBPK) modelling strategy has been used to evaluate the clinical relevance of the in vitro concentrations, which impair NCC migration in the scratch assay. The in vivo plasma concentrations of the tested compounds were within the same range as the concentrations used in the scratch assay (Zimmer B. et al., 2014. PMID: 24691702).

3. No special considerations known.

8.6 Applicability of test method*

Pesticides, flame retardants, polychlorinated biphenyls (PCBs), drugs have been detected.

The exact applicability domain is not yet clear.

Volatile compounds and substances that are not water-soluble cannot be measured.

8.7 Incorporation in test battery*

a) Strengths:

- Medium throughput
- Automated microscopy

b) The comparable scratch assay is a low throughput assay.

c) The assay has been considered as part of the ESNATS screen battery.

d) The test can to some extent be stand-alone (positive hits are meaningful). The 'negative hits' provide little information and require other tests in a developmental toxicity test battery

9. Publication / validation status

9.1 Availability of key publications

Neurodevelopmental toxicity assessment of flame retardants using a human DNT in vitro testing battery.
Klose, J. et al. Cell Biol Toxicol. 2021; PMID: 33969458

Establishment of an a priori protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of developmental neurotoxicity.
Masjosthusmann, S. et al. EFSA Supporting Publications. 2020; 17(10): 1938E.

The EU-ToxRisk method documentation, data processing and chemical testing pipeline for the regulatory use of new approach methods.
Krebs, A. et al. Arch. Toxicol., 2020. PMID: **32632539**

A structure-activity relationship linking non-planar PCBs to functional deficits of neural crest cells: new roles for connexins.
Nyffeler, J. et al. Arch. Toxicol., 2018. PMID: 29164306

Combination of multiple neural crest migration assays to identify environmental toxicants from a proof-of-concept chemical library.
Nyffeler, J. et al. Arch. Toxicol., 2017. PMID: 28477266

Design of a high-throughput human neural crest cell migration assay to indicate potential developmental toxicants.
Nyffeler, J. et al. ALTEX, 2016. PMID: 27463612

Identification of transcriptome signatures and biomarkers specific for potential developmental toxicants inhibiting human neural crest cell migration.
Pallocca, G. et al. Arch.Toxicol., 2016. PMID: 26705709

Profiling of drugs and environmental chemicals for functional impairment of neural crest migration in a novel stem cell-based test battery.
Zimmer, B. et al. Arch. Toxicol., 2014. PMID: 24691702

Evaluation of developmental toxicants and signaling pathways in a functional test based on the migration of human neural crest cells.
Zimmer, B. et al. Environ. Health Perspect., 2012. PMID: 22571897

9.2 (Potential) linkage to AOPs*

Test method is not linked to AOP.

9.3 Steps towards mechanistic validation*

a) Cells express typical NCC markers, are of human origin and are able to migrate

b) Cell migration requires dynamic variability of the cytoskeleton, e.g. actin reorganization and the dynamic instability of microtubules in the leading edge. If this is inhibited by cytochalasin D or taxol migration of NCCs is inhibited.

c) A formal mechanistic validation has not been performed.

9.4 Pre-validation or validation*

To date, 141 unique compounds (as defined by unique DTXSIDs) have been tested successfully in this assay.

No formal OECD 34 validation study has been done (eg., ring trials with a standard set of known positive and negative controls).

In total, >200 different compounds were tested in the cMINC assay. The test method was developed using a compound training set (Nyffeler et al. 2016). It was used for an 80 compound screening library from the US National Toxicology Program (NTP) (Nyffeler et al. 2017). The test method was part of a DNT hazard assessment for 120 compounds in a DNT testing battery. The later compound set includes potential DNT positive and DNT negative compounds (Masjosthusmann 2020).

9.5 Linkage to (e.g. OECD) guidelines / regulatory use*

Test is not linked to regulatory guidelines.

10. Test method transferability*

10.1 Operator training*

Experiences are required in:
- cell culture
- multichannel/multistep pipetting
- handling of Array Scan VTI HCS Reader (Cellomics, PA) and its software
- Microsoft Excel
- Ringassay software
- GraphPad Prism

Operator is trained and guided by a highly experienced instructor. Approximately 4 weeks will be needed for a smooth assay performance.
Learning iPSC culture and cell differentiation takes several months.

10.2 Transfer*

The assay hasn't been transferred or applied in other labs.

11. Safety, ethics and specific requirements

11.1 Specific hazards; issues of waste disposal

No specific requirements.

11.2 Safety data sheet (SDS)

SDS are available in the university of Konstanz DaMaRIS database (**Da**ngerous **Ma**terials **R**egistry **I**nformation **S**ystem).

11.3 Specific facilities / licenses

Work requires S1 cell culture laboratories (genetically modified cells).
No specific facilities are required.
No specific ethical approval is required.

11.4 Commercial aspects / intellectual property of material / procedures*

To our best knowledge, no elements needed to conduct the experimental part of the test method are protected. Programs used to conduct the analysis of the data (Microsoft Excel and GraphPad Prism) need to be purchased or obtained by license agreement, however data analysis and plotting can be done with other, freely available tools.

Annex IV – ToxTemp UKN4 (NeuriTox) assay

1. Overview

1.1 Descriptive full-text title

Assay to test compound-derived impairment in neurite outgrowth in human dopaminergic neurons (NeuriTox; UKN4) – V2.0

1.2 Abstract

This *in vitro* test method is based on human neurons (LUHMES cells) at a stage of neurite growth. Is assesses (a) disturbances in the development of the nervous system/brain structures, and (b) direct damage to the adult nervous system, by exposure to toxicants. The neurite area (which serves as indirect measurement of neuronal interconnectivity) of stained differentiating neurons as well as cellular viability are measured simultaneously using high content imaging. The processes of neurite outgrowth and cell viability are assessed. The data of this method are meant to predict (a) developmental disorders in children caused by compound exposure during fetal development, and (b) damage to the developing nervous system, in particular to dopaminergic parts of the nervous system. The method has not undergone formal validation and has not been part of a ring trial. It predicts some aspects of neurotoxicity, but not all aspects covered by an in vivo neurotoxicity study (TG424). It has been used in the screening of medium-sized compound libraries, has undergone some mechanistic evaluation, and has been linked to AOP id3 in aopwiki.org (parkinsonian motor deficits). According to the readiness criteria as published by Bal-Price et al. (2018) the NeuriTox assay obtained the readiness score A.

2. General information

2.1 Name of test method

NeuriTox test, UKN4

2.2 Version number and date of deposition

This is Version 2.0 of the protocol “Assay to test compound-derived impairment in neurite outgrowth in human dopaminergic neurons (NeuriTox; UKN4) – V2.0)”. It was assembled and deposited in February 2022. A previous version was assembled in October 2019 in the context of the EU-ToxRisk project (see publication Krebs et al., 2020).

2.3 Summary of introduced changes in comparison to previous version(s)

Changes compared to V1 refer mainly to changed parameters in the prediction model.

Test procedures remain unchanged.

2.4 Assigned data base name

Normal text names often do not uniquely define the method. Therefore, each method should be assigned a clearly and uniquely defined data base name.

Here this is exemplified by the names generated in the EU-ToxRisk project:

UKN1a_DART_NPC_Diff_6D_02

UKN1b_DART_NPC_Diff_4D_01

UKN2a_DART_NC_Migr_24h_04

The name is assembled (in more generic terms) from the following elements:

Axa_B_C_D_E

Axa: mandatory part of the identifier allowing unambiguous identification

A: Abbreviation / acronym of the partner depositing the assay

x: Consecutive number (referring to the partner's assay number)

a: Sub-specifier (for variants, i.e. very similar assays, but e.g. different readout or medium); not mandatory, but 'Axa' must be specific (i.e. clearly identifying) for each assay variant

B: Indication of the main intended use (max 5 letters), e.g. DART, Neuro, Liver, Lung, Renal, Redox, Stress...

C: Specifier for test system, e.g. cell type, e.g. NPC (neural precursor cells), NC (neural crest), Hep (liver cells), REN (kidney cells), PUL (lung cells) (max. 4 letters)

D: Identification of test endpoint, e.g. Diff_6D = Differentiation for 6 days; exp_24 h = exposure for 24 hours; RNA_6h = transcriptome after 6 hours; and so on (use max. 15 signs altogether; if desired in 2-3 blocks), name (and acronym) of the project partner home organisation.

E: version number.

UKN4a_DART_LUH_neurite_24h

2.5 Name and acronym of the test depositor

University of Konstanz (UKN), Germany

2.6 Name and email of contact person

Prof. Dr. Marcel Leist
marcel.leist@uni-konstanz.de
Tel: +49-7531885037

2.7 Name of further persons involved

Jonathan Blum (PhD student, experimenter)
jonathan.blum@uni-konstanz.de

Anna-Katharina Ückert (PhD student, experimenter)
anna-katharina.ueckert@uni-konstanz.de

2.8 Reference to additional files of relevance

- An important reference is the DB-ALM Protocol n° 200 (the new prediction model remains to be updated)
- Raw data file
- Data processing file
- Internal SOP

3. Description of general features of the test system source*

Note: *this section might be redundant with section 4. It is meant to describe the procedure of generating the cells, which are eventually used in the test method. This applies e.g. for stem cells, which have to be differentiated towards the cell type with which the method is conducted (e.g. neurons or hepatocytes derived from iPSC). See scheme for illustration. If this is not applicable to your test system, go directly to section 4.*

Overview of cell culture stages relevant to their use in an *in vitro* test method

Stage:	Original cell source	Test system generation / maturation	Test run: start of test	Test run: end of test
Chapter:	3.1 – 3.5	3.6	4.1 + 4.2	4.4 – 4.8 / 4.3
Example 1:	iPSC	generation of neurons	neuronal maturation	
Example 2:	HEK 293	re-plating	proliferation	
Example 3:	primary hepatocytes	spreading, partial de-differentiation, bile duct formation	spontaneous cell death, de-differentiation	
Example 4:	primary astrocytes	re-plating, secondary culture	spontaneous activation, metabolic shifts	
Processes:	maintenance, isolation, thawing	differentiation, proliferation, aggregation, self-organization, re-plating, maturation, de-differentiation	maturation, proliferation, cell death, de-differentiation / differentiation	

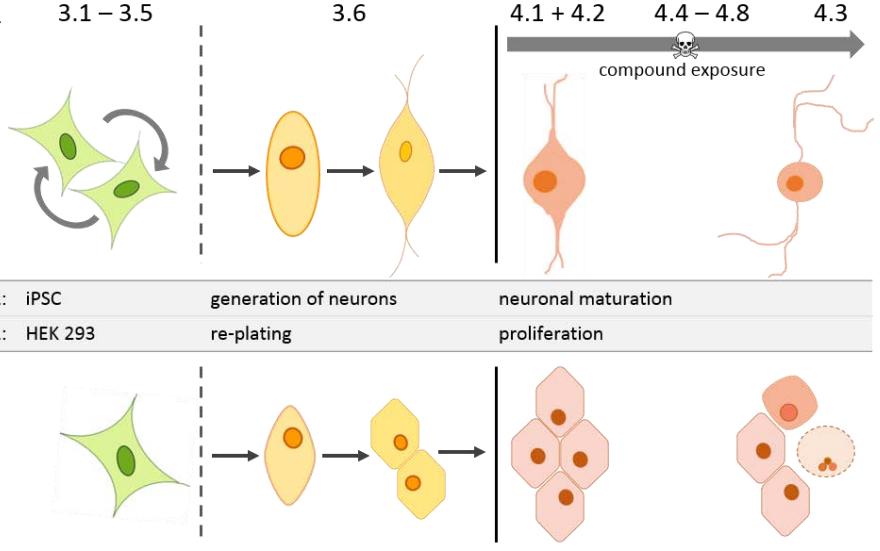


Figure 15: Overview of test system stages and where to find / deposit corresponding information in this document.

Note: Refer to overview figure to connect question numbers and cell culture stages.

3.1 Supply of source cells

The LUHMES cells have been brought from Lund (Sweden) to Konstanz in 2006 (Lotharius et al., 2005) and a masterstock has been frozen. The cells used for the test method are continuously generated by cell culture. Approximately every 4-5 years, a vial from the cell masterstock is thawed and extensively expanded. These cells are then frozen as working stock. A new cell batch is thawed every 4 weeks from this working stock; cells are cultured until passage 20-25.

3.2 Overview of cell source component(s)

LUHMES cell line

LUHMES cells originate from the ventral mesencephalon of an 8 week old human, female fetus. They exhibit the same characteristics as MESC2.10 cells.

LUHMES cells can be differentiated into morphologically and biochemically mature **dopamine-like neurons** following exposure to tetracycline, GDNF (glial cell line-derived neurotrophic factor), and db-cAMP for 6 days.

They are usually cultured in a **2D monolayer**, but have also been shown to grow into 3D structures (Smirnova et al., 2015, Brull et al., 2020). They can be co-cultured with mouse astrocytes.

3.3 Characterization and definition of source cells

- **ATCC number**: LUHMES ATCC® CRL-2927™; LUHMES cells used at the University of Konstanz differ from LUHMES cells deposited and distributed by ATCC. The assay described in here is based on UKN (University of Konstanz) LUHMES cells.
- **Origin**: mesencephalon of an 8 week old human fetus, subclone of the tetracycline-controlled, v-myc-overexpressing human mesencephalic-derived cell line MESC2.10. MESC2.10 has been conditionally immortalized with a LINX v-myc retroviral vector with a tet-off system.
- **Gender**: female
- **Morphology**: upon differentiation, cells form neuronal network, see figure below
- **Doubling time:** approx. 14 -20 h (depending on passage)
- **Phenotype:** Dopaminergic (DA)-phenotype; tyrosine hydroxylase (TH) expression depends on the presence of cAMP in the culture medium. Under differentiation conditions, cells extend neurites after 24h. They display extensive growth cones, but synapse formation is not clear. After 6 days of differentiation, cells express voltage-dependent ion channels and show **electrical activity**.
- **Expression:** Cells express α-synuclein (Parkinson's disease) (Schildknecht et al., 2013), all standard synaptic proteins and all proteins required for AD (Alzheimer's disease) pathology (Aβ-formation, Tau hyperphosphorylation)(Scholz et al., 2013; Scholz et al., 2018).

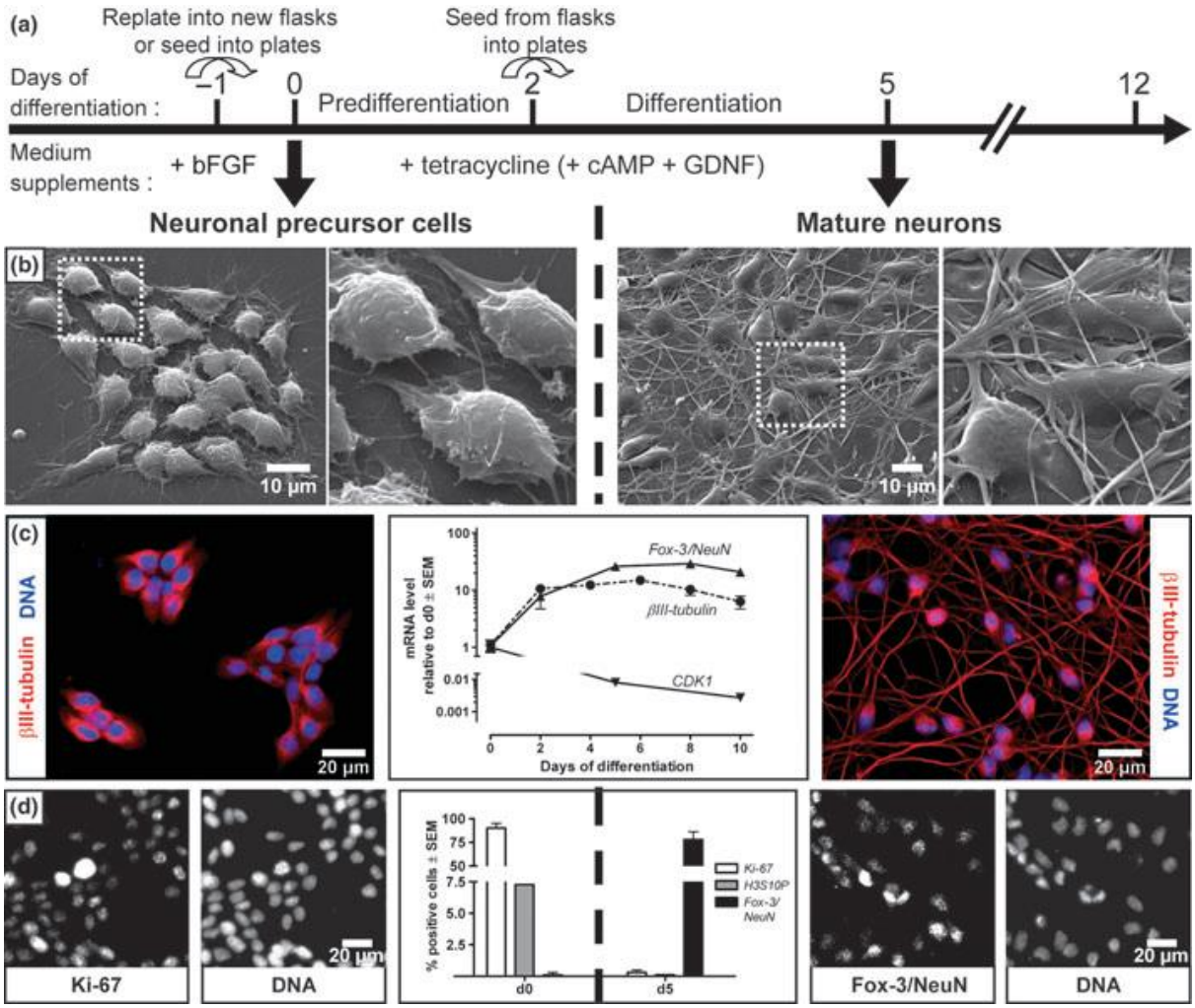


Fig. 2: Conversion of proliferating LUHMES cells into post-mitotic neurons. LUHMES were grown and differentiated either on glass cover slips or in multi-well plates. Cells were either fixed for microscopy or lysed for RNA extraction at different stages between day 0 and day 10 (d0–d10). **(a)** Schematic representation of the 2-step differentiation procedure, initiated by the absence of the cytokine basic fibroblast growth factor (bFGF) and addition of tetracycline. Unless mentioned otherwise, dibutyryl cAMP (cAMP) and glial cell derived neurotrophic factor (GDNF) were present throughout the differentiation. **(b)** Representative scanning electron microscopy (SEM) images of undifferentiated (d0) and differentiated (d5) LUHMES with marked squares shown at higher magnification. **(c)** LUHMES were immunostained on d0 and d5 for bIII-tubulin and nuclei were labeled by DNA staining with H-33341 dye. The mRNA expression levels of bIII-tubulin, Fox-3/NeuN and cyclin-dependent kinase 1 (CDK1) were determined after different days of maturation by RT-qPCR. **(d)** The proliferative status of d0 and d5 cells was quantified by immunostaining of Ki-67, H3S10P and Fox-3/NeuN. It is indicated as percentage of positive nuclei relative to all nuclei, as identified by DNA staining with H-33342. Quantitative data are expressed as means ± SEM from three independent differentiations.
Scholz et al., 2011

Short Tandem Repeat signature: (lab internal data)

	ATCC website	UKN LUHMES	ATCC LUHMES
Amelogenin	X	X	X
CSF1PO	13,14	13,14	13,14

D5S818	11,13	11,13	11,13
D13S317	9,11	9,11	9,11
D7S820	11,13	11,13	11,13
D16S539	11,12	11,12	11,12
vWA	14,17	14,17	14,17
THO1	7,9.3	7,9.3	7,9.3
TPOX	8	8	8
Penta D		12,13	12,13
D8S1179		12,13	12,13
FGA		19,21	19,21
D3S1358		17,18	17,18
D21S11		30,31	30,31
D18S51		12	12
Penta E		11,13	11,13

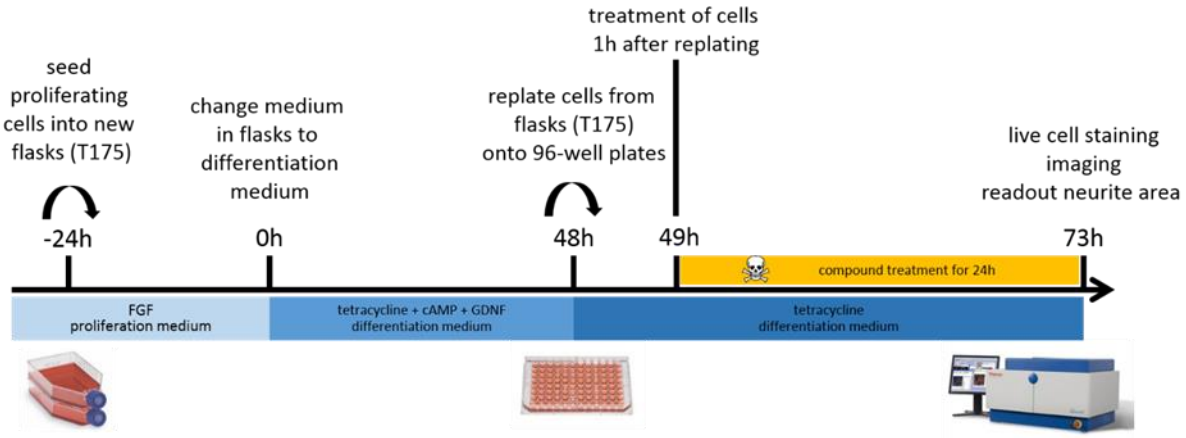
3.4 Acceptance criteria for source cell population

The cells have to be pathogen-free to be used in further experiments (regular testing for mycoplasma after thawing). Cells of the working stock are checked for marker gene expression once the stock is created, but not every thawed cell batch is checked. There only visual inspection is done. The cells should be of oval morphology, and should not form extensive aggregates and clustering. Approx. 5-10% of cell death (dead cells floating in flask) is considered normal, increased cell death one passage after thawing justifies to discard of the cell batch. Culture medium should not be yellow. Proliferating cells from working stock batches should only be used up to passage 18.

3.5 Variability and troubleshooting of source cells

- Plastic coating is critical for even cell distribution; problems with coating often leads to cell clumping and aggregation
- Cells maintained and distributed by ATCC might behave differently than cells from the stock at University of Konstanz (Gutbier et al., 2018)
- Passage number of working stock cells might influence cellular behavior
- Problems with plastic ware obtained from other suppliers than indicated have occurred in the past
- Too little or too high cell density can cause problems (confluency should never exceed 40-90%)

3.6 Differentiation towards the final test system



day -1: Proliferating LUHMES cells are seeded in proliferation medium
day 0: Medium is changed from proliferation medium to differentiation medium
day 2: Replating to 96-well plates, toxicant treatment
day 3: Readout

Differentiation is initiated by the exposure to tetracycline, GDNF (glial cell line-derived neurotrophic factor), and db-cAMP (tetracycline shuts down v-myc expression).

Coating of flasks and plates
principle:
plates and flasks are coated with a poly-L-orthinine (PLO) and fibronectin solution in ion-exchanged and purified MilliQ-H_2O overnight in the cell culture incubator. The next day the coated plastic ware is washed with MilliQ water once and dried under sterile conditions in the cell culture hood. The coated plastic can be stored at 4°C up to 4 weeks.

Differentiation
principle:
For differentiation, the following growth factors are added: GDNF (glial cell-derived neurotrophic factor), dibutyryl-cAMP (Cyclic adenosine monophosphate) and tetracycline.
If differentiation medium without cAMP and GDNF is required, the volume of those components is replaced with DMEM/F12 Advanced.

3.7 Reference / link to maintenance culture protocol

External document is available

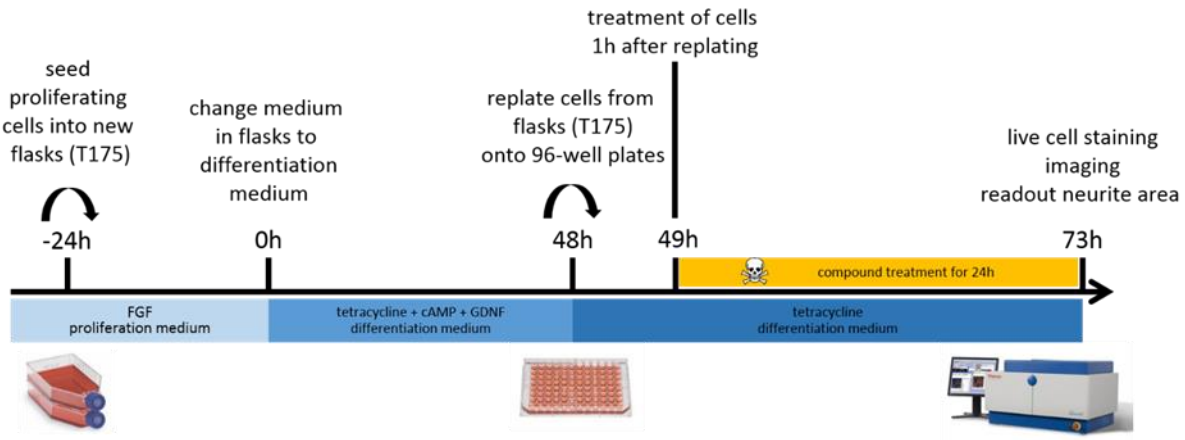
Mainentance
principle:
The growth factor bFGF is added to the medium of proliferating cultures. Cells are passaged every 2-3 days (Monday – Wednesday – Friday). Minimum confluence for LUHMES cells should not fall below 40%. A confluence of 85% should not be exceeded.
For splitting, cells are counted every time and seeded in the according cell numbers:
For 2 days: 3 Mio per T75 flask

6 Mio per T175 flask
For 3 days: 1 Mio per T75 flask
2.7 Mio per T175 flask

4. Definition of the test system as used in the method

Note: *this section refers to the stage of the test system, which is then used for the test method. See scheme for illustration. If you have cells that do not need prior differentiation, give their basic characteristics here. If your test system undergoes significant changes between the maintenance culture and the use for testing, please also fill in section 3.*

4.1 Principles of the culture protocol



day -1: Proliferating LUHMES cells are seeded in proliferation medium
day 0: Medium is changed from proliferation medium to differentiation medium
day 2: Cell number reaches about 30-40 Mio per T175 flask.
Cells are trypsinized and replated onto 96-well plates (30'000 cells/well in 90 µl) in differentiation medium without cAMP and GDNF.
At about 30 min - 2 h after seeding, when cells have attached, the compounds are added (10 µl of each dilution; total volume 100 µl)
day 3: 23.5 h after toxicant treatment, cells are live-stained with H-33342 and calcein-AM and incubated for 30 min. After 24 h of treatment (including staining), the cells are imaged using a high-content microscope (Cellomics).

4.2 Acceptance criteria for assessing the test system at its start

- Cells should have neurites approx. as long as their cell body when they are trypsinized for re-plating
- They must not be contaminated
- They should have a confluency of about 60%
- The moment the cells are being treated, cells are not attached to the plate completely yet and still appear round when seen through a microscope

→ Criteria are not quantified and are mainly based on visual inspection

4.3 Acceptance criteria for the test system at the end of compound exposure

After compound treatment, the negative controls should fulfil the following:
- Neurites should be at least as long as cell bodies
- Medium should not be orange/yellow
- Appropriate cell density
  → Criteria are not quantified and are mainly based on visual inspection

4.4 Variability of the test system and troubleshooting

Causes of variability:
- Problems with coating (different PLO/fibronectin batches; problems with water quality):
  → If problems occur: Wash twice after coating
  Wash with PBS instead of MilliQ-H_2O
  Don't store plates and flasks in the fridge, use them immediately
- Different cell passages:
  → Cells have different morphology and behavior the older they get; thawing a new batch might be useful
- Lots of different plates/flasks:
  → Plastic might be different, if the manufacturer delivers from a different/new lot
- Differences between the vials of one cell "lot"
- Different lots of medium and supplements
- Cells too confluent
  → impaired metabolism
  → too slow differentiation (autocrine proliferation stimulation)
  → discard cells in that case

Cell lot = cells that have all been frozen at the same time. Usually 10-15 flasks (T175) are frozen in numerous vials.
Cell batch = vial with approx. 3 Mio cells frozen
Cell passage = cells of each vial thawn are passaged up to 17 times

4.5 Metabolic capacity of the test system

Dopamine transporter is expressed and used for MPP^+ transport.

4.6 Omics characterization of the test system

Microarray analysis has been used to compare differentiated LUHMES on day 3 (end of UKN4) and day 6 to undifferentiated LUHMES cells (Data source: https://kops.uni-konstanz.de/bitstream/handle/123456789/28842/ Weng_288423.pdf;sequence=1).

Epigenetic modifiers have been extensively characterized (Weng et al., 2014).
Genes relevant to AD have been extensively characterized (Scholz et al., 2011).
Genes relevant for neuronal receptor composition have been extensively characterized (Loser et al., 2020; Loser et al., 2021). Genes triggered by mitochondrial toxicants have been identified (Delp et al., 2021).
There is a complete transcriptome data set on LUHMES differentiation from day 2 – day 10: It is deposited at the EBI data base under accession-ID S-TOXR1833 (public as of March 2022).

4.7 Features of the test system that reflect the in vivo tissue

- As neurons of the central nervous system they express the dopamine transporter DAT
- Cells grow axons and neurites in course of their differentiation
- They express various neuronal transporters and receptors (e.g. purinergic receptors, nicotinic acetylcholine receptors (Loser et al., 2020)
- They are electrophysiologically active and excitable, produce action potentials (Loser et al., 2020)

4.8 Commercial and intellectual property rights aspects of cells

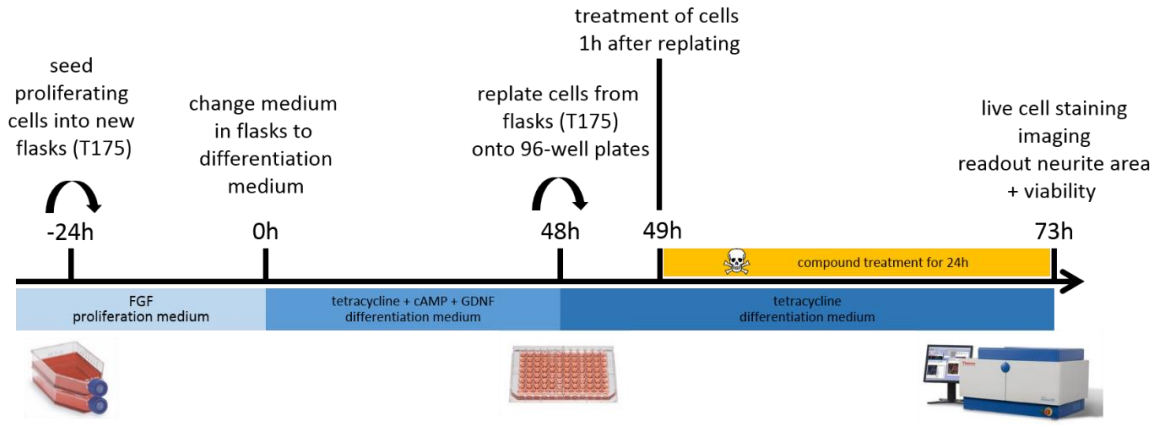
The cells are not protected by patents or any other licences.

4.9 Reference / link to the culture protocol

Chapter 3 has been answered.
The maintenance is described in the DB-ALM SOP °200 available at:
https://ecvam-dbalm.jrc.ec.europa.eu/methods-and-protocols/protocol/ukn4-assay-to-test-compound-derived-impairment-in-neurite-outgrowth-in-differentiating-human-dopaminergic-neurons-protocol-no.-200/key/p_1602
Another lab-internal SOP is also available upon request to the Leist-lab.

5. Test method exposure scheme and endpoints

5.1 Exposure scheme for toxicity testing



day -1: LUHMES cells are seeded in proliferation medium
day 0: Medium is changed from proliferation medium to differentiation medium
day 2: Cells number reaches about 30-40 Mio per T175 flask
Cells are trypsinized and replated onto 96-well plates (30'000 cells/well in 90 µl)
At about 30 min - 2 h after seeding when cells have attached, the compounds are added (10 µl of each dilution). **Toxicant exposure for 24 h from day 2 to day 3 of differentiation.**
day 3: 23.5 h after toxicant treatment, cells are stained with H-33342 and calcein-AM and incubated for 30 min. After 24 h the cells are imaged using a high-content microscope (Cellomics).

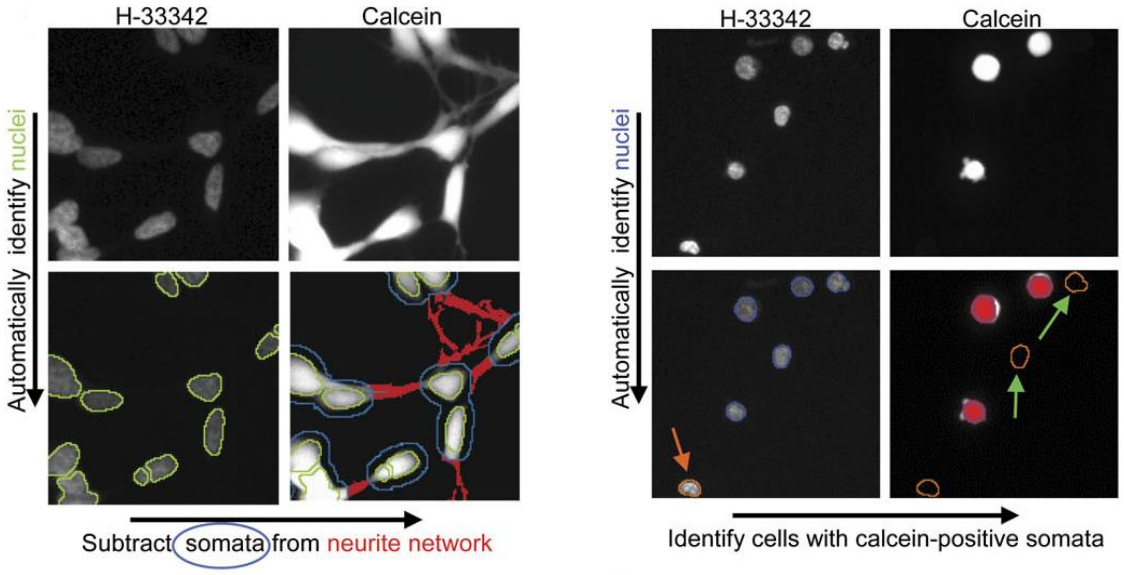
5.2 Endpoint(s) of the test method

Test endpoints: 1) neurite area (main endpoint)
2) cell number
3) % of viable cells (reference endpoint)

5.3 Overview of analytical method(s) to assess test endpoint(s)

Cells are stained with calcein-AM to mark viable cells. Co-staining with Hoechst H-33342 allows the identification of any cell.
Cells are stained for 30 min at 37°C and 5% CO_2 in the incubator.
The cell staining is imaged in a Cellomics Array Scan VTI HCS reader.
Hoechst H-33342 staining is imaged in channel 1 (UV-Hoechst); calcein staining is imaged in channel 2 (Green-FITC). Exposure times are set manually.

To measure the neurite area, the software acquires the Hoechst signal in channel 1 to identify the cells as objects (via identification of the nuclei), and the calcein-AM signal in channel 2 to measure neurite area. Double positive cells are counted as viable.



5.4 Technical details (of e.g. endpoint measurements)

Quantification of neurite outgrowth

An automated microplate reading microscope (Array-ScanII HCS Reader, Cellomics, PA) equipped with a Hamamatsu ORCA-ER camera (resolution 1024 x 1024; run at 2 x 2 binning) was used for image acquisition. Ten fields per well were imaged. Images were recorded in 2 channels using a 20x objective and excitation/emission wavelengths of 365 ± 50/535 ± 45 to detect H-33342 in channel 1 and 474 ± 40/535 ± 45 to detect calcein in channel 2. In both channels, a fixed exposure time and an intensity histogram-derived threshold were used for object identification. Neurite pixels were identified using the following image analysis algorithm: nuclei were identified as objects in channel 1 according to their size, area, shape, and intensity which were predefined on untreated cells using a machine-based learning algorithm, and manual selection of nuclei to be classified as intact. The nuclear outlines were expanded by 3.2 µm in each direction, to define a virtual cell soma area (VCSA) based on the following procedure: The average width of the cytoplasm ring (distance nucleus - cell

membrane) of LUHMES cells was experimentally determined to be 2.3 µm. Size irregularities were not always due to growing neurites, as determined by combined F-actin/tubulin beta-III staining. To avoid scoring of false positive neurite areas, the exclusion ring (VCSA) was made bigger than the average cell size. Then, we used two control compounds (U0126 and bisindolylmaleimid I) to vary the expanded outlines from 0.6 to 4 µm. We found 3.2 µm to be optimal both to detect neurite growth over time and to identify reduced neurite growth with high sensitivity. All calcein-positive pixels of the field (beyond a given intensity threshold) were defined as viable cellular structures (VCSs). The threshold was dynamically determined for each field after flat field and background correction and intensity normalization to 512 gray values and was set to 12% of the maximal brightness (channel 63 of 512). The VCS defines the sum of all somata and neurites without their assignment to individual cells. In an automatic calculation, the VCSAs, defined in the H-33342 channel, were used as filter in the calcein channel and subtracted from the VCS. The remaining pixels (VCS - VCSA) in the calcein channel were defined as neurite area.

Quantification of individual viable cells by imaging

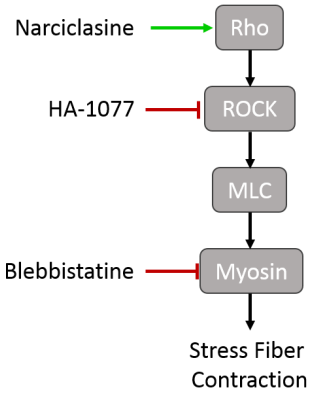
For a quantitative assessment of viable cells, the same images used to assess neurite area were analysed using another image analysis algorithm: nuclei were identified in channel 1 as objects according to their size, area, shape, and intensity. Nuclei of apoptotic cells with increased fluorescence were excluded. A VCSA was defined around each nucleus by expanding it by 0.3 µm into each direction. Calcein-AM staining, labelling live cells, was detected in channel 2. The algorithm quantified the calcein intensity in the VCSA areas. Cells having an average calcein signal intensity in the VCSAs below a predefined threshold were classified by the program as “not viable”. Valid nuclei with a positive calcein signal in their cognate VCSA were counted as viable cells. A positive calcein signal was based on measurements of the average intensity (normal cells: 1300 ± 115, threshold: < 50) and the total integrated intensity (normal cells: 186,000 ± 23,600, threshold < 1000) of cells.

5.5 Endpoint-specific controls / mechanistic control compounds (MCC)

Positive control for neurite growth inhibition:
Narciclasine: activates Rho

Positive control for neurite growth enhancement:
HA-1077: Rho-associated kinase inhibitor
Blebbistatine: inhibits myosin II

Rho/ROCK/LIM kinase/cofilin pathway:
induces actin polymerization, key regulator of the cytoskeleton and cell polarity



5.6 Positive controls

Positive control: narciclasine (50 nM final concentration)

5.7 Negative and unspecific controls

Negative control: solvent (0.1% DMSO final concentration), mannitol, paracetamol, ASS, galloflavin.

5.8 Features relevant for cytotoxicity testing*

Cells are highly sensitive to toxicants (Tong ZB 2016). Cell death is easily quantified, LDH release always shows very high baseline activity.

5.9 Acceptance criteria for the test method

Positive control narciclasine:
Neurite area ≤ 75% of DMSO control
Viability ≥ 90% of DMSO control (or not significantly changed)

Negative control DMSO:
Neurite area ≥ 35’000

5.10 Throughput estimate*

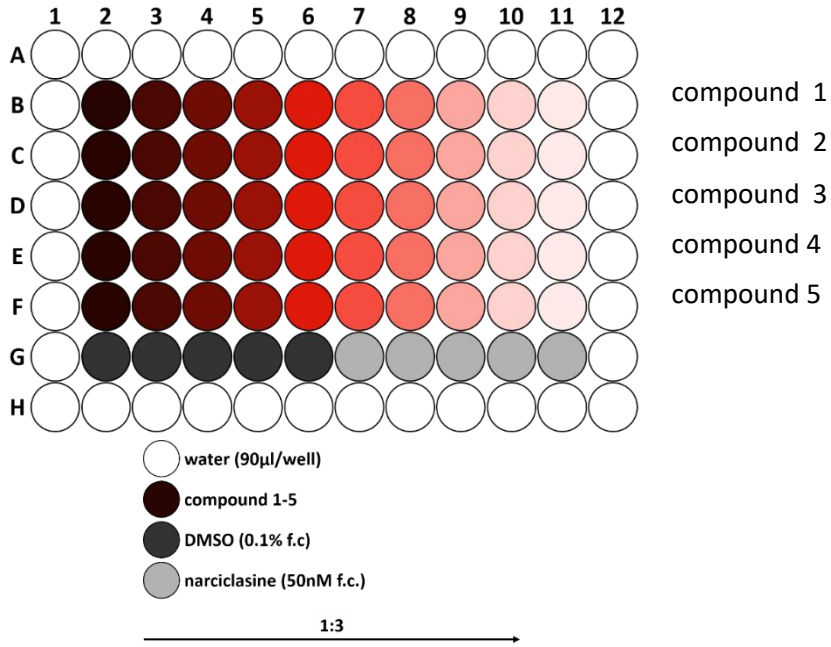
Data point = one biological replicate (→ usually 3 technical replicates); each concentration/condition of a compound counts as data point

1200 data points per month

5 compounds per plate, 10 different concentrations of each compound per plate (see figure) → 50 data points (3 plates)
one plate correlates to one technical replicate → 3 plates for 3 technical replicates

6 plates can be done per day (correlates to 10 compounds → 100 data points)
→ 3 days of readout per week → 300 data points per week; 4 weeks per month → 1200 data points per month.

Typical plate layout:



Typical weekly work schedule:

Montag	Dienstag	Mittwoch	Donnerstag	Freitag	Samstag	Sonntag
Diff (direkt)		Seed + treat	Cellomics readout			
Prediff	Diff (change medium)		Seed + treat	Cellomics readout		
				Prediff	Diff (change medium)	
Seed + treat	Cellomics readout					

6. Handling details of the test method

6.1 Preparation / addition of test compounds

- Compounds are stored according to the manufacturer's instructions (e.g. 4°C, room temperature, -20°C).

- Preferable solvent is DMSO. The used DMSO is stored in a lightproof, air-tight bottle at room temperature.

- Final DMSO concentration on the cells is 0.1%

- After dissolving the compounds, which are delivered in a solid/powder form, all compound solutions are aliquoted into volumes sufficient for one experiment (i.e. one biological replicate). In this way repeated freezing and thawing and therefore damaging the compound's stability and efficiency can be avoided.

- For conducting an experiment, a compound aliquot is thawn and diluted with 'DMEM/F12 Advanced' without supplement in a separate deepwell-plate.

- All compound dilutions in the deepwell plate contain 1% DMSO, so that a final concentration of 0.1% DMSO is reached on the cells. The highest compound concentration is diluted with medium

1:100 without DMSO as 1% is already reached with the DMSO the compound is solved in, the serial dilution is done with DMEM/F12 Advanced without supplement and 1% DMSO.

- The compound dilutions (10 µl each) are added to the cells using a multichannel-multistepper pipette, 6 filter tips at a time, dispense mode, speed for uptake is set to medium, speed for output is set to high. 40 µl are taken in and 10 µl are released on each plate.

6.2 Day-to-day documentation of test execution

Plate maps are defined prior to the experiment and documented in the lab book and files (Excel files) are stored on the work group server.
Concentrations and compound dilutions are calculated prior to the experiment.
Experimental procedures are noted manually in a paper lab book.

6.3 Practical phase of test compound exposure

The experimenter plans the experiment according to Cellomics microscope availability (has to be booked in advance) and availability of a sufficient number of cells.

Pipetting errors are marked directly on the plate maps and are documented in the lab book.

The paper lab book is taken to cell culture rooms and errors are documented in there right away.

The technical replicates were pipetted from left to right. Pipetting starts with the highest concentration at the left column.

6.4 Concentration settings

5 compounds per plate

As default a serial dilution 1:3 is used, i.e. a concentration range of 19683-fold is covered (e.g. from 20µM → 1nM).

Serial dilutions of compounds are prepared in a separate deepwell-plate, from which 10 µl are transferred to the according plates with attached cells using a multichannel-multistepper pipette. DMEM/F12 Advanced cell medium without supplements is used for dilution.

Dilution steps can be adapted to be more narrow (e.g. 1:1.5)

6.5 Uncertainties and troubleshooting*

- Compound solubility in stock and during dilution is too low (stock solved in 100% DMSO, final concentration of the solvent on the cells is 0.1% DMSO)
- Some compounds show autofluorescence and interfere with the detection of calcein-AM or H-33342.
- To prevent negative edge effects, only the inner 60 wells of a 96-well plate are used and the edge wells were filled with PBS/MilliQ water.
- Focusing failure of Array Scan VTI HCS Reader (Cellomics, PA) can be a problem that produces outliers; as well as imaging only one channel.
- Highly trained/automated handling with multichannel and multistepper pipette is necessary to achieve little variance.

- Operators can get trained within 2-4 weeks. Cell seeding and medium change should be performed as fast as possible to keep cells as short as possible at room temperature. The more practice an operator has, the faster the critical steps can be performed.
- Substances are added when pipette tips are touching the wall of the wells right above the medium surface. When the substance solution is pipetted too high above the medium surface, the droplet may just stick to the wall of the well without flowing down into the medium.

6.6 Detailed protocol (SOP)

Protocol n^{o} 200 in DB-ALM data base

See the SOP (Appendix L in Masjosthusmann et al. 2020).

6.7 Special instrumentation

The method requires a Cellomics Array Scan VTI HCS high content reader that may not be present in the standard lab.

6.8 Possible variations*

a) further additional endpoints:
- metabolic activity (resazurin reduction)
- glutathione levels
- staining of tubulin
- analysis of differentiation markers by qPCR or immunostaining

b) other analytical endpoints:
cell viability by:
- fluorimetric measurement of resazurin conversion
- measurement of extracellular LDH
- measurement of luminescence indicating ATP content

c) other exposure:
- compound can be washed out → acquisition on day 4
- longer exposure is possible
- later exposure is possible (from day 5 on) in order to measure more mature neurite networks
- the medium can be changed to contain galactose instead of glucose. This increases the sensitivity of the cells to inhibitors of mitochondrial respiration (Delp et al., 2019)

d) variants for recording of neurite growth:
- neurite growth by GFP-labelled cells (Schildknecht S 2013)

6.9 Cross-reference to related test methods*

UKN3a assay to test compound-derived neurite integrity impairment in human mature dopaminergic neurons after long-term compound exposure - Protocol no 202 in DB-ALM

UKN3b assay to test compound-derived neurite integrity impairment in human mature dopaminergic neurons - Protocol no 196 in DB-ALM

7. Data management

7.1 Raw data format

Raw data is extracted by copy-paste in Excel files (example file available).

Data from all technical replicates are collected in one file.

7.2 Outliers

4. Mathematical procedures to define outliers have not been defined. Data points that 'look' very far off are discarded. Biological outliers do practically not exist, most far data points are the result of technical problems (focus not found, only one channel imaged, etc.)
5. All raw data (incl. outliers) are stored.
6. Technical outliers make up 1-0.1%.

7.3 Raw data processing to summary data

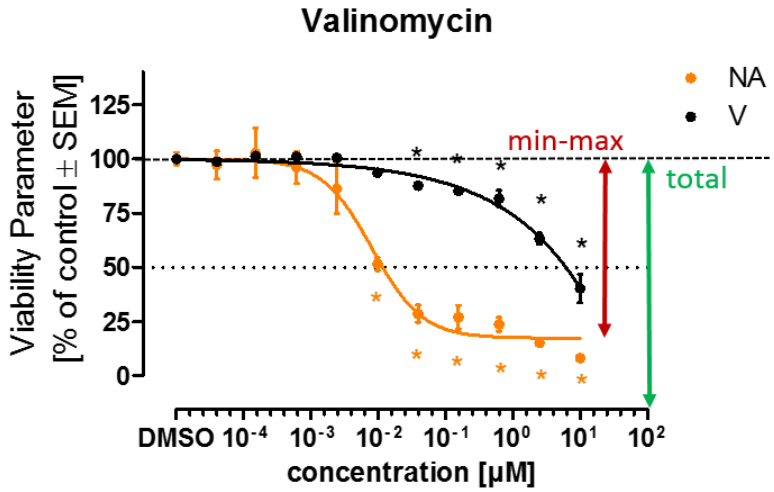
- Array Scan VTI HCS Reader (Cellomics, PA) takes images (optionally bitmap or tiff-format; 512 x 512 pixels, 8bit or 16bit)

- Images are locally analyzed using the Array Scan software, algorithms quantify neurite area and cell count (nuclei)

- data are copy-pasted into an Excel sheet, further analysis is done with Excel + KNIME + GraphPad Prism and BMCeasy.

7.4 Curve fitting

The data are analyzed with Excel and represented with GraphPad Prism.
For the concentration curve, a nonlinear regression fit is calculated. The fitting method is least squares. If a non-linear curve fit is not possible, a linear curve fit is performed. The curve deriving from the fit is a 4-parameter log function. To calculate the EC50 value, this log-function is solved for y=50% of the total scale, not for 50% of the min-max scale (see example below). Treated concentrations are analyzed for deviation from control. Sometimes it is analyzed whether the deviation of neurite growth is different from the deviation of viability. This is done by two-way ANOVA + Tukey-Kramer post hoc testing. Statistics applied are one-way ANOVA (and nonparametric) with Dunnett's post test.



EC50(NA)= 0.01 µM
EC50(V)= 6.16 µM

BMC values with their upper and lower confidence intervals (BMCU and BMCL) are calculated via the publically available online software:
http://invitrotox.uni-konstanz.de/BMCeasy/

7.5 Internal data storage*

The data are firstly stored on the microscope computer and then exported to other servers (lab group server and university server), which are back-upped regularly.

7.6 Metadata

The metadata are documented, stored and exported as text document (log)-files to the according scheme: (local PC)_descriptor(date and time)_XXX.log:

The following metadata are stored:

- cellinsight-pc_160429130003_AutomationControllerIni
- cellinsight-pc_160429130003_kineticprotocol
- cellinsight-pc_160429130003_protocol
- cellinsight-pc_160429130003_scan
- cellinsight-pc_160429130003_ScanIni
- cellinsight-pc_160429130003_spooling
- cellinsight-pc_160429130003.spooled

7.7 Metadata file format

Metadata files are available.

8. Prediction model and toxicological application

8.1 Scientific principle, test purpose and relevance

LUHMES cells used in this test method represent cells of the central nervous system with a dopaminergic phenotype.

The UKN4 test method models neurite outgrowth as a biological process and assesses viability of the cells in parallel. The cells are used in an early developmental stage (day 2 of differentiation) and chemical exposure occurs during this development. Therefore, the UKN4 test method assesses hazards for developmental neurotoxicity. It can be integrated into adverse outcome pathways as an important key event to predict potential adverse outcomes in humans.

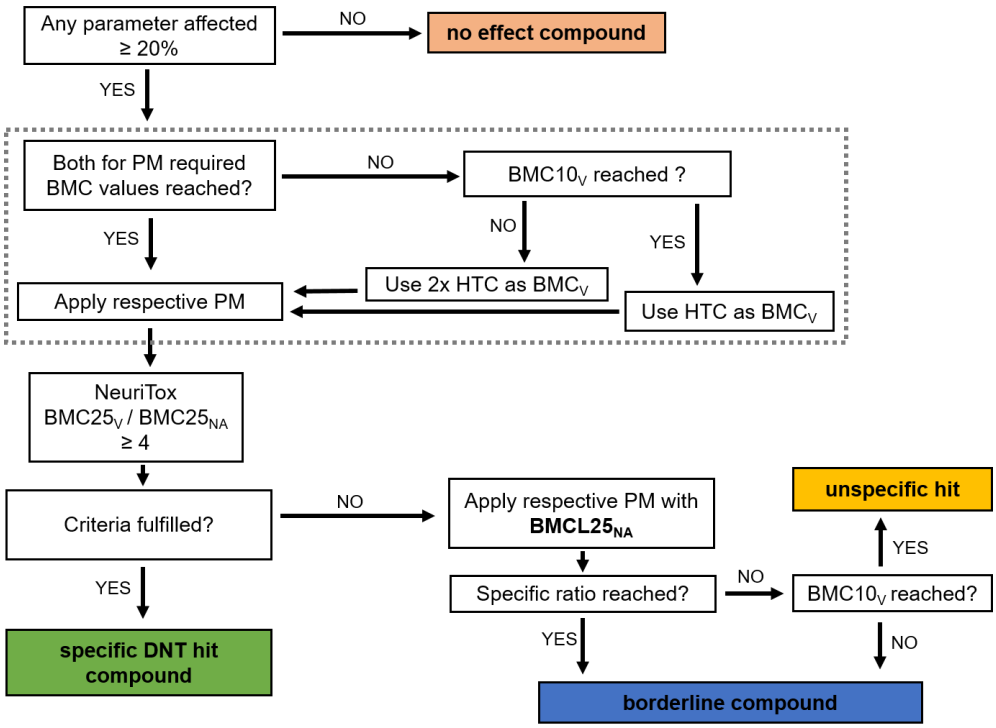
8.2 Prediction model

Three different models are used:

1. prediction model for screening:
hit = decrease/increase in neurite area while viability is not changed (compare to narciclasine positive control:
Neurite area ≤ 80% of DMSO control
Viability ≥ 90% of DMSO control

2. prediction model for compound hazard evaluation:
hit confirmation testing; BMC25 Viability (V) / BMC25 Neurite Area (NA) ≥ 4 → specifically neurotoxic

3. prediction model for borderline compounds:
A ratio of BMC25 Viability (V) / BMC**L**25 Neurite Area (NA) ≥ 4 is considered a borderline hit. In some scenarios the viability does not reach the BMC25 necessary for the ratio calculations. In this case the highest tested concentration (HTC) was used. Schematic representation of the complete prediction model is shown in scheme below.



8.3 Prediction model setup

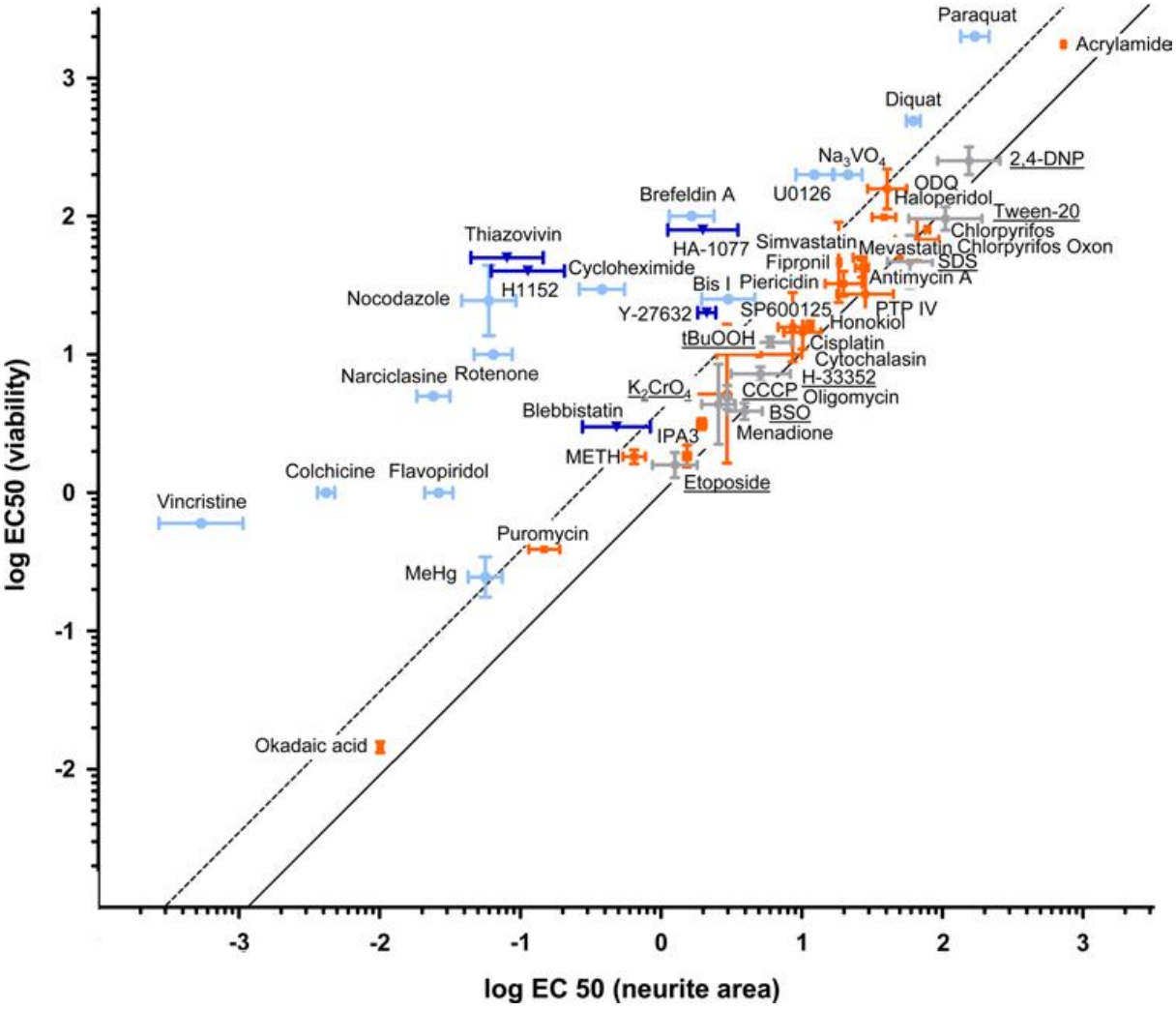
a) The prediction model was established using the following compounds (Stiegler 2011; Krug 2013):
-colchicine
-vincristine
-nocodazole
➔ Positive controls
-etoposide
-buthionine sulfoximine (BSO)
➔ Unspecific toxicants affecting general viability
-cycloheximide
-paraquat
➔ Rules of assay interpretation, criteria to define a positive test result
-CCCP
-2,4-DNP
-SDS
-tween-20
-K2CrO4
-H-33352
-tertiary butyl hydroperoxide (tBuOOH)
➔ Unspecific toxicants

b) The prediction model has been applied to screen 80 compound library of NTP (Delp et al., 2018). The prediction model including the borderline classification has been applied to screen a 120 compound library (Masjosthusmann et al., 2020).

c) The process is documented in Krug et al., 2013 and Stiegler et al., 2011

d) Sensitivity/ specificity have not been defined due to a lack of reference compounds. Below, a set of compounds that triggers specific inhibition of neurite growth is shown (light blue). Other compounds are cytotoxic without specific effects on neurites (orange).

The use of the updated prediction model
BMC25 Viability (V) / BMC25 Neurite Area (NA) ≥ 4
was validated by comparing classifications derived by the initial and the updated prediction model. The reason for updating the prediction model was that a decrease of 50% in neurite area and viability cannot always be achieved with our range of test concentrations. An effect of 25% is more often observed and came to similar results when applied (Delp et al., 2018).



8.4 Test performance

Some background on the test performance is given in chapters 8.2/8.3 (prediction model).

Several performance parameters for the test were obtained in several separate evaluation rounds.

A first evaluation was done during the first publication of the model and its applications (Stiegler et al. 2011, Krueg et al. 2013). Here, a panel of well-selected positive and negative controls have been

tested. Accordingly, the specificity was 100% and the sensitivity was > 90 %. In dedicated experiments, S/N ratios of > 20 and a z' of > 0.5 have been determined. The compound narciclasine (run on each plate as positive control) was tested across 36 different test plates and 12 independent assay runs. The neurite area relative to the solvent control varied between 40% and 75% across all plates with a viability value constantly >90% (Delp et al. 2018).

The test has been used in screening campaigns, and real-live performance data under broader screen conditions have been obtained. The different performance data need to be considered, when a compound is a hit in a screen, or whether it has been specifically evaluated in a hit follow-up or a mechanistic project.

A first screen application has been the NTP80 screen (80 compounds provided by the US NTP). Data are published Delp et al. 2018.

A second screen application has been the cross systems case study of the EU-ToxRisk project. The baseline variation is indicated in Krebs et al., 2020. Moreover, an overview is given for 19 compounds on the BMC/BMCL ratio as measure of readout certainty.

A third screen was performed in the context of the EFSA DNT test battery evaluation with 120 compounds (Masjosthusmann et al. 2020). From this screen the following performance indicators were obtained:

A: **Sensitivity: 100%**

→ With NeuriTox as standalone assay in 17 `tool negatives´ tested (Masjosthusmann et al. 2020).

B: **Specificity: 82.7%**

→ With NeuriTox combined in a full DNT battery and a selected set of 27 positive compounds with evidence for DNT (Masjosthusmann et al. 2020).

C: **Baseline variation (intra-experimental)**

Neurite outgrowth: 11.7 ± 4.5%

Cell viability: 3.8 ± 3.5%

D: **Baseline variation (inter-experimental)**

Neurite outgrowth: 15.3%

Cell viability: 4.3%

E: **Variation of a positive control run on each (inter-experimental)**

Neurite outgrowth: 38.5%

Definition of values C-E

C: **Baseline variation (intra-experimental)** is the mean CV±SD of the CV of all replicates of the solvent control from each experiment across n>200 experiments.

D: **Baseline variation (inter-experimental)** is the variability across all independent experiments (n>200) after normalization based on the response of the lowest test concentration. It was assumed that the lowest test concentration does not affect any of the endpoints measured.

E: **Variation of a positive control run on each (inter-experimental)** is the variability of the positive control across all independent experiments (n>40) after normalization. Example for a positive control that on average reduced the specific endpoint down to 40% (relative to solvent control) and a calculated variability of 50%: 0.5 x 40% = +/- 20% → The positive control with mean of 40% varies from 20% to 60%.

8.5 *In vitro – in vivo extrapolation (IVIVE)*

1. Estimated lipid content and albumin concentration in in vitro test media and human plasma:

Medium	Lipid content (mg/l)	Albumin concentration (µM)
UKN4	2.9	5.8
Human plasma	6000	600

this information can be used to calculate from nominal to free concentrations of compounds tested (https://doi.org/10.1007/s00204-020-02802-6))

2. The test has not been used extensively for IVIVE. However, data from the cell model (but different exposure scheme) have been used for IVIVE modelling (Loser et al., 2021, b). The test has also been used in projects with potency estimates and dose estimates (Klose et al., 2021; van der Stel et al., 2021).
3. No special considerations known.

8.6 *Applicability of test method**

Test is sensitive to cytoskeletal toxicants, some signaling modifiers and flame retardants. Polycyclic aromatic hydrocarbons (PAH) and HDAC inhibitors have no effect.

8.7 *Incorporation in test battery**

a) Strengths:

- Medium to high throughput
- Automated microscopy

b) compared to UKN5 (which quantifies neurite outgrowth of peripheral neurons), UKN4 measures neurite outgrowth specifically of CNS dopaminergic neurons. This was shown by treatment with MPP^+, which is transported by the dopamine transporter (DAT) and had an effect in UKN4, but not UKN5 peripheral neurons, which lack the DAT transporter.

c) specific effects on central nervous system dopaminergic neurons. Implementation in a DNT battery was investigated (Masjosthusman et al., 2020).

d) Preferential use in first tier, no complementary assays required

9. Publication / validation status

9.1 Availability of key publications

"Progressive degeneration of human mesencephalic neuron-derived cells triggered by dopamine-dependent oxidative stress is dependent on the mixed-lineage kinase pathway."
Lotharius J, Falsig J, van Beek J, Payne S, Dringen R, Brundin P, Leist M.
J Neurosci. 2005 Jul 6;25(27):6329-42. DOI: 10.1523/JNEUROSCI.1746-05.2005. PMID: 16000623

"Rapid, complete and large-scale generation of post-mitotic neurons from the human LUHMES cell line."
Scholz D, Pöltl D, Genewsky A, Weng M, Waldmann T, Schildknecht S, Leist M.
J Neurochem. 2011 Dec;119(5):957-71. DOI: 10.1111/j.1471-4159.2011.07255.x. PMID: 21434924

"Assessment of chemical-induced impairment of human neurite outgrowth by multiparametric live cell imaging in high-density cultures."
Stiegler NV, Krug AK, Matt F, Leist M.
Toxicol Sci. 2011 May;121(1):73-87. DOI: 10.1093/toxsci/kfr034. PMID: 21342877

"Control of Abeta release from human neurons by differentiation status and RET signaling."
Scholz, D., Y. Chernyshova and M. Leist
Neurobiol Aging. 2013 Jan;34(1):184-99. DOI: 10.1016/j.neurobiolaging.2012.03.012. PMID: 22534065

"Evaluation of a human neurite growth assay as specific screen for developmental neurotoxicants."
Krug AK, Balmer NV, Matt F, Schönenberger F, Merhof D, Leist M.
Arch Toxicol. 2013 Dec;87(12):2215-31. DOI: 10.1007/s00204-013-1072-y. PMID: 23670202

"Generation of genetically-modified human differentiated cells for toxicological tests and the study of neurodegenerative diseases."
Schildknecht S, Karreman C, Pöltl D, Efrémova L, Kullmann C, Gutbier S, Krug A, Scholz D, Gerding HR, Leist M.
ALTEX. 2013;30(4):427-44. DOI: 10.14573/altex.2013.4.427. PMID: 24173167

"A LUHMES 3D dopaminergic neuronal model for neurotoxicity testing allowing long-term exposure and cellular resilience analysis."
Smirnova L, Harris G, Delp J, Valadares M, Pamies D, Hogberg HT, Waldmann T, Leist M, Hartung T.
Arch Toxicol. 2015 Dec 8. DOI: 10.1007/s00204-015-1637-z. PMID: 26647301

"Reduced Abeta secretion by human neurons under conditions of strongly increased BACE activity."
Scholz, D., Y. Chernyshova, A. K. Uckert and M. Leist
J Neurochem. 2018 Oct;147(2):256-274. DOI: 10.1111/jnc.14467. PMID: 29804308

"Major changes of cell function and toxicant sensitivity in cultured cells undergoing mild, quasi-natural genetic drift."
Gutbier S, May P, Berthelot S, Krishna A, Trefzer T, Behbehani M, Efremova L, Delp J, Gstraunthaler G, Waldmann T, Leist M.
Arch Toxicol. 2018 Dec;92(12):3487-3503. DOI: 10.1007/s00204-018-2326-5. PMID: 30298209

“A high-throughput approach to identify specific neurotoxicants/ developmental toxicants in human neuronal cell function assays.”
Delp J, Gutbier S, Klima S, Hoelting L, Pinto-Gil K, Hsieh JH, Aichem M, Klein K, Schreiber F, Tice RR, Pastor M, Behl M, Leist M.
ALTEX. 2018;35(2):235-253. DOI: 10.14573/altex.1712182. Erratum in: ALTEX. 2019;36(3):505. PMID: 29423527

“Stage-specific metabolic features of differentiating neurons: Implications for toxicant sensitivity.”
Delp J, Gutbier S, Cerff M, Zasada C, Niedenführ S, Zhao L, Smirnova L, Hartung T, Borlinghaus H, Schreiber F, Bergemann J, Gätgens J, Beyss M, Azzouzi S, Waldmann T, Kempa S, Nöh K, Leist M.
Toxicol Appl Pharmacol. (b) 2018 Sep 1;354:64-80. DOI: 10.1016/j.taap.2017.12.013 Erratum in: Toxicol Appl Pharmacol. 2019 Jun 1;372:70. PMID: 29278688

"Major changes of cell function and toxicant sensitivity in cultured cells undergoing mild, quasi-natural genetic drift."
Gutbier, S., P. May, S. Berthelot, A. Krishna, T. Trefzer, M. Behbehani, L. Efremova, J. Delp, G. Gstraunthaler, T. Waldmann and M. Leist
Arch Toxicol. 2018 Dec;92(12):3487-3503. DOI: 10.1007/s00204-018-2326-5. PMID: 30298209

"Development of a neurotoxicity assay that is tuned to detect mitochondrial toxicants."
Delp, J., M. Funke, F. Rudolf, A. Cediel, S. H. Bennekou, W. van der Stel, G. Carta, P. Jennings, C. Toma, I. Gardner, B. van de Water, A. Forsby and M. Leist
Arch Toxicol. 2019 Jun;93(6):1585-1608. DOI: 10.1007/s00204-019-02473-y. PMID: 31190196

"Incorporation of stem cell-derived astrocytes into neuronal organoids to allow neuro-glial interactions in toxicological studies."
Brull, M., A. S. Spreng, S. Gutbier, D. Loser, A. Krebs, M. Reich, U. Kraushaar, M. Britschgi, C. Patsch and M. Leist
ALTEX. 2020 Mar; 37(3): 409-428. DOI:10.14573/altex.1911111. PMID: 32150624

"Establishment of an a priori protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of developmental neurotoxicity."
Masjosthusmann, S., J. Blum, K. Bartmann, X. Dolde, A.-K. Holzer, L.-C. Stürzl, E. H. Keßel, N. Förster, A. Dönmez, J. Klose, M. Pahl, T. Waldmann, F. Bendt, J. Kisitu, I. Suciu, U. Hübenthal, A. Mosig, M. Leist and E. Fritsche
EFSA Supporting Publications. 2020; 17(10): 1938E.
https://doi.org/10.2903/sp.efsa.2020.EN-1938

"Human neuronal signaling and communication assays to assess functional neurotoxicity."
Loser, D., J. Schaefer, T. Danker, C. Moller, M. Brull, I. Suciu, A. K. Uckert, S. Klima, M. Leist and U. Kraushaar
Arch Toxicol. 2021 Jan;95(1):229-252. DOI: 10.1007/s00204-020-02956-3. PMID: 33269408

"Neurotoxicity and underlying cellular changes of 21 mitochondrial respiratory chain inhibitors."
Delp, J., A. Cediel-Ulloa, I. Suciu, P. Kranaster, B. M. van Vugt-Lussenburg, V. Munic Kos, W. van der Stel, G. Carta, S. H. Bennekou, P. Jennings, B. van de Water, A. Forsby and M. Leist
Arch Toxicol. 2021 Feb;95(2):591-615. DOI: 10.1007/s00204-020-02970-5.PMID: 33512557

"Neurodevelopmental toxicity assessment of flame retardants using a human DNT in vitro testing battery."
Klose, J., M. Pahl, K. Bartmann, F. Bendt, J. Blum, X. Dolde, N. Forster, A. K. Holzer, U. Hubenthal, H. E. Kessel, K. Koch, S. Masjosthusmann, S. Schneider, L. C. Sturzl, S. Woeste, A. Rossi, A. Covaci, M. Behl, M. Leist, J. Tigges and E. Fritsche
Cell Biol Toxicol. 2021 May 10. DOI: 10.1007/s10565-021-09603-2. PMID: 33969458

"Functional alterations by a subgroup of neonicotinoid pesticides in human dopaminergic neurons."
Loser, D., M. G. Hinojosa, J. Blum, J. Schaefer, M. Brüll, Y. Johansson, I. Suciu, K. Grillberger, T. Danker, C. Möller, I. Gardner, G. F. Ecker, S. H. Bennekou, A. Forsby, U. Kraushaar and M. Leist
Arch Toxicol. (b) 2021 Jun;95(6):2081-2107. DOI: 10.1007/s00204-021-03031-1. PMID: 33778899

"New approach methods (NAMs) supporting read-across: Two neurotoxicity AOP-based IATA case studies."
Van der Stel, W., G. Carta, J. Eakins, J. Delp, I. Suciu, A. Forsby, A. Cediel-Ulloa, K. Attoff, F. Troger, H. Kamp, I. Gardner, B. Zdrazil, M. J. Mone, G. F. Ecker, M. Pastor, J. C. Gomez-Tamayo, A. White, E. H. J. Danen, M. Leist, P. Walker, P. Jennings, S. Hougaard Bennekou and B. Van de Water
ALTEX. 2021;38(4):615-635. DOI: 10.14573/altex.2103051. PMID: 34114044

*9.2 (Potential) linkage to AOPs**

Test method could be potentially linked to the following AOPs in AOPwiki:

- AOP 48 : Binding of agonists to ionotropic glutamate receptors in adult brain causes excitotoxicity that mediates neuronal cell death, contributing to learning and memory impairment.
  → Organ effects: Neurodegeneration, decreased neuronal network function
  → Organism effects: Impairment of learning and memory
- AOP 13: Chronic binding of antagonist to N-methyl-D-aspartate receptors (NMDARs) during brain development induces impairment of learning and memory abilities.
  → Organ effects: Decreased neuronal network function
  → Organism effects: Impairment of learning and memory
- AOP 3: Inhibition of the mitochondrial complex I of nigra-striatal neurons leads to parkinsonian motor deficits.
  → Organ effects: degeneration of DA neurons of nigrostriatal pathway
- AOP 42: Inhibition of Thyroperoxidase and Subsequent Adverse Neurodevelopmental Outcomes in Mammals
  → Adverse Outcome: Cognitive function decreased
- AOP 54: Inhibition of Na+/I- symporter (NIS) decreases TH synthesis leading to learning and memory deficits in children.
  → Organ effects: decreased neuronal network function
  → Organism effects: learning and memory deficits
- AOP 8: Upregulation of Thyroid Hormone Catabolism via Activation of Hepatic Nuclear Receptors, and Subsequent Adverse Neurodevelopmental Outcomes in Mammals.
  → Adverse Outcome: Altered neurodevelopment
- AOP 134: Sodium Iodide Symporter (NIS) Inhibition and Subsequent Adverse Neurodevelopmental Outcomes in Mammals.

→ Adverse Outcome: Cognitive function decreased

9.3 Steps towards mechanistic validation*

a) LUHMES are dopaminergic, express DAT, TH. Are of human origin, form network

b) Tubulin plays a major role in neurite outgrowth and if inhibited by colchicine/vincristine/nocodazole neurite outgrowth is reduced. If the Rho/Rock pathway is activated, neurite outgrowth is enhanced.
c) A formal mechanistic validation has not been performed. Reversibility and protection/counterregulation by mechanistic compounds have been shown (Stiegler 2011; Krug AK).

d) The test covers a fundamental neurodevelopmental process. In some contexts, this might be seen as a key event of an AOP (Smirnova 2014, Bal-Price 2015 (ISTNET))

9.4 Pre-validation or validation*

To date, 143 unique compounds (as defined by unique DTXSIDs) have been tested successfully in this assay.

No formal OECD 34 validation study has been done (eg., ring trials with a standard set of known positive and negative controls).

In total, >200 different compounds were tested in the NeuriTox assay. The test method was developed using a compound training set (Krug et al. 2013). It was used for an 80 compound screening library from the US National Toxicology Program (NTP) (Delp et al. 2018). The test method was part of a DNT hazard assessment for 120 compounds in a DNT testing battery. The later compound set includes potential DNT positive and DNT negative compounds (Masjosthusmann 2020).

9.5 Linkage to (e.g. OECD) guidelines / regulatory use*

Test is not linked to regulatory guidelines.

10. Test method transferability*

10.1 Operator training*

Experiences are required in:
- LUHMES cell culture
- multichannel/multistep pipetting
- handling of Array Scan VTI HCS Reader (Cellomics, PA) and its software
- Microsoft Excel
- KNIME
- GraphPad Prism

Operator is trained and guided by a highly experienced instructor. Approximately 4 weeks will be needed for a smooth assay performance.

10.2 Transfer*

Test system (UKN LUHMES cells) has been transferred and established to numerous other labs. The NeuriTox assay has been successfully transferred to one other lab for tool compound testing and data comparison.

11. Safety, ethics and specific requirements

11.1 Specific hazards; issues of waste disposal

No specific requirements.

11.2 Safety data sheet (SDS)

SDS are available in the university DaMaRIS database (**Da**ngerous **Ma**terials **R**egistry **I**nformation **S**ystem).

11.3 Specific facilities / licenses

Work requires S1 cell culture laboratories (genetically modified cells).
No specific facilities are required.
No specific ethical approval is required.

11.4 Commercial aspects / intellectual property of material / procedures*

To our best knowledge, no elements needed to conduct the experimental part of the test method are protected. Programs used to conduct the analysis of the data (Microsoft Excel and GraphPad Prism) need to be purchased or obtained by license agreement, however data analysis and plotting can be done with other, freely available tools.

Annex V – ToxTemp UKN5 (PeriTox) assay

1. Overview

1.1. Descriptive full-text title

Assay to test compound-derived impairment in neurite outgrowth in human iPSC-derived immature dorsal root ganglia (iDRG) neurons (PeriTox; UKN5) – V2.0

1.2 Abstract

This in vitro test method is based on human iPSC-derived immature dorsal root ganglia (iDRG) neurons at a stage of neurite growth. It assesses (a) disturbances in the development of the (peripheral) nervous system, and (b) direct damage to the peripheral nervous system, by exposure to toxicants. The neurite area (which serves as indirect measurement of neuronal interconnectivity) of stained differentiating neurons, as well as cellular viability are measured simultaneously using high content imaging. The processes of neurite outgrowth and cell death are measured. The data of this method are meant to predict (a) developmental disorders in children caused by compound exposure during fetal development, and (b) damage to the developed nervous system, in particular to the peripheral nervous system. The method has not undergone formal validation and has not been part of a ring trial. It predicts some aspects of neurotoxicity, but not all aspects covered by an in vivo neurotoxicity study (TG424). It has been used in the screening of medium-sized compound libraries, has undergone some mechanistic evaluation, and has been linked to AOP-279 (AOPwiki ID) / ETR09N (EU-ToxRisk AOP task ID) (Peripheral neuropathy caused by microtubule interacting drugs). According to the readiness criteria as published by Bal-Price et al. (2018) the PeriTox assay obtained the readiness score A-.

2. General information

2.1 Name of test method

PeriTox test, UKN5

2.2 Version number and date of deposition

This is Version 2.0 of the protocol “Assay to test compound-derived impairment in neurite outgrowth in human iPSC-derived immature dorsal root ganglia (iDRG) neurons (PeriTox; UKN5) – V2.0)”. It was assembled and deposited in February 2022. A previous version was assembled in 2019 in the context of the EU-ToxRisk project (see publication Krebs et al., 2020).

2.3 Summary of introduced changes in comparison to previous version(s)

Changes compared to V1 refer mainly to the generation of the test system and the cell line used. Test procedures and parameters remain unchanged.

2.4 Assigned data base name

Normal text names often do not uniquely define the method. Therefore, each method should be assigned a clearly and uniquely defined data base name.

Here this is exemplified by the names generated in the EU-ToxRisk project:

UKN1a_DART_NPC_Diff_6D_02

UKN1b_DART_NPC_Diff_4D_01

UKN2a_DART_NC_Migr_24h_04

The name is assembled (in more generic terms) from the following elements:

Axa_B_C_D_E

Axa: mandatory part of the identifier allowing unambiguous identification

A: Abbreviation / acronym of the partner depositing the assay

x: Consecutive number (referring to the partner's assay number)

a: Sub-specifier (for variants, i.e. very similar assays, but e.g. different readout or medium); not mandatory, but 'Axa' must be specific (i.e. clearly identifying) for each assay variant

B: Indication of the main intended use (max 5 letters), e.g. DART, Neuro, Liver, Lung, Renal, Redox, Stress...

C: Specifier for test system, e.g. cell type, e.g. NPC (neural precursor cells), NC (neural crest), Hep (liver cells), REN (kidney cells), PUL (lung cells) (max. 4 letters)

D: Identification of test endpoint, e.g. Diff_6D = Differentiation for 6 days; exp_24 h = exposure for 24 hours; RNA_6h = transcriptome after 6 hours; and so on (use max. 15 signs altogether; if desired in 2-3 blocks), name (and acronym) of the project partner home organisation.

E: version number.

UKN5_DART_iDRG_24h_02

2.5 Name and acronym of the test depositor

University of Konstanz (UKN), Germany

2.6 Name and email of contact person

Prof. Dr. Marcel Leist
marcel.leist@uni-konstanz.de
Tel: +49-7531885037

2.7 Name of further persons involved

Anna-Katharina Holzer (PhD student, experimenter)
anna-katharina.holzer@uni-konstanz.de

Jonathan Blum (PhD student)
jonathan.blum@uni-konstanz.de

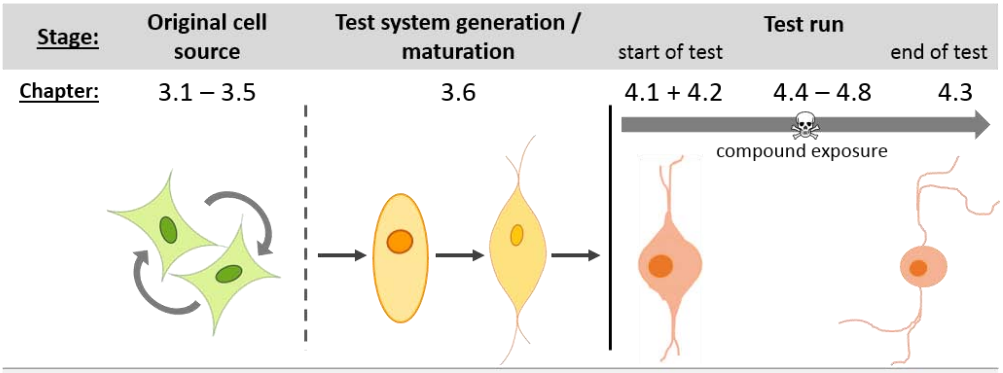
2.8 Reference to additional files of relevance

- An important reference is the DB-ALM Protocol n° 218 → The original iPSC are in the meantime cultured feeder-free (see 3.2)
- Raw data file
- Data processing file

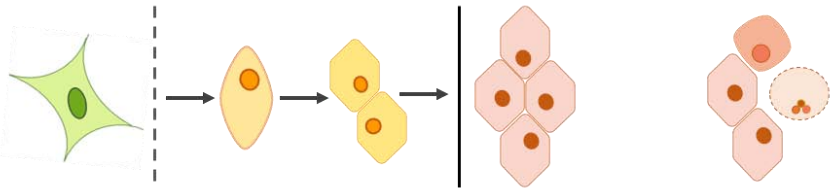
3. Description of general features of the test system source*

Note: *this section might be redundant with section 4. It is meant to describe the procedure of generating the cells, which are eventually used in the test method. This applies e.g. for stem cells, which have to be differentiated towards the cell type with which the method is conducted (e.g. neurons or hepatocytes derived from iPSC). See scheme for illustration. If this is not applicable to your test system, go directly to section 4.*

Overview of cell culture stages relevant to their use in an *in vitro* test method



Example 1:	iPSC	generation of neurons	neuronal maturation
Example 2:	HEK 293	re-plating	proliferation



Example 3:	primary hepatocytes	spreading, partial de-differentiation, bile duct formation	spontaneous cell death, de-differentiation
Example 4:	primary astrocytes	re-plating, secondary culture	spontaneous activation, metabolic shifts
Processes:	maintenance, isolation, thawing	differentiation, proliferation, aggregation, self-organization, re-plating, maturation, de-differentiation	maturation, proliferation, cell death, de-differentiation / differentiation

Figure 16: Overview of test system stages and where to find / deposit corresponding information in this document.

Note: Refer to overview figure to connect question numbers and cell culture stages.

3.1 Supply of source cells

The human induced pluripotent stem cell (hiPSC) line EPITHELIAL-1 has been bought from Sigma-Aldrich, Germany in 2018 and a masterstock has been frozen. From the masterstock several working stocks have been prepared. The working stocks are regularly thawed and can be continously maintained due to self-renewal and pluripotency capabilities of the cells. The cells are maintained up to 8 passages before a new vial of the working stock is thawed.

3.2 Overview of cell source component(s)

The human induced pluripotent stem cell line iPSC EPITHELIAL-1 (Cat# IPSC0028) is purchased from Sigma-Aldrich, Taufkirchen, Germany as a frozen suspension of single cells. iPSC EPITHELIAL-1 cells are produced via reprogramming of epithelial cells from a Caucasian female (24 years) using OSKM retrovirus. Pluripotency was certified by gene and protein expression of pluripotency markers. The maintenance culture is usually cultured in colonies under feeder-free conditions on Laminin-521 coating in Essential 8 (E8) medium. The cells are split weekly.

3.3 Characterization and definition of source cells

- **Tissue**: epithelium
- **Gender**: female
- **Culture properties:** adherent
- **Disease:** no disease was diagnosed
- **Age:** 24 year old
- **Ethnicity:** Caucasian
- **Expression:** iPSC EPITHELIAL-1 express all expected pluripotency markers, such as OCT4, NANOG, SSEA4 and SOX1.
- **STR analysis:** conformes, at least 80% loci homology of observed:

  TH01:9,9.3

  D5S818:12,13

  D13S317:8,12

  D7S820:8,11

  D16S539:11,12

  CSF1P0:10,12

  AMEL:X,X

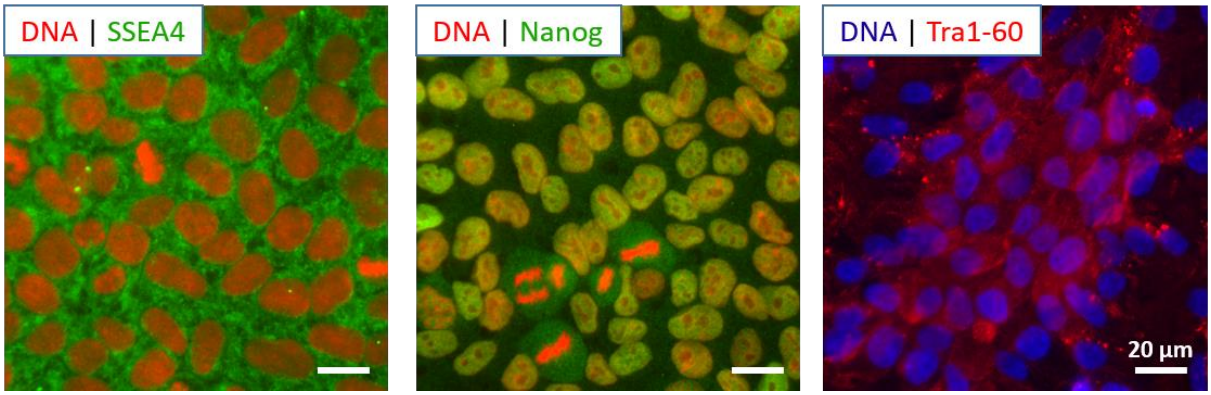
  vWA:16,16

  TPOX:8,9

3.4 Acceptance criteria for source cell population

The cells have to be pathogen-free to be used in further experiments (regular testing for mycoplasma).

The iPSC maintenance is regularly checked for expression of pluripotency markers (Oct4, Nanog, Tra-1-60) by immunocytochemistry.

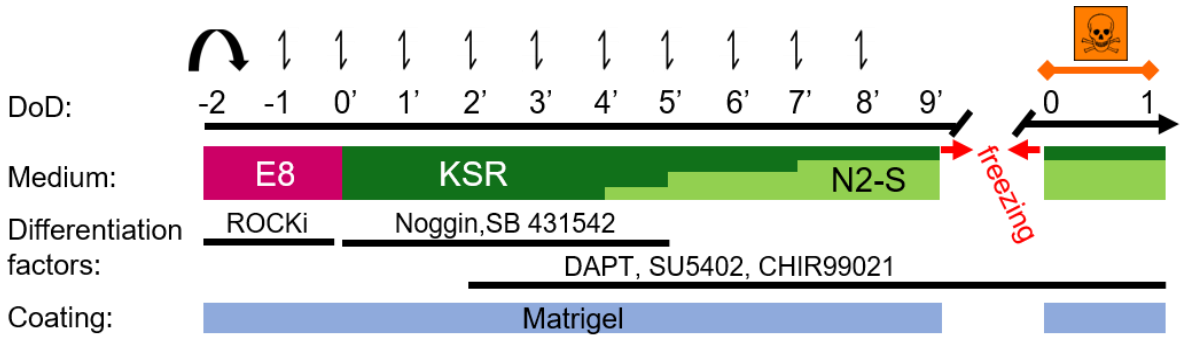


The cells should grow in colonies with sharp edges, no spontaneously differentiated cells should be visible. Stem cells are split every 5-7 days whenever they reach >80% confluency. The cells can be used for differentiation from passage 2 on until passage 8.

3.5 Variability and troubleshooting of source cells

- hiPSC can be maintained up to 8 passages, high passage number might influence performance of cells
- Too little or too high cell density leads to detachment of cells or spontaneous differentiation
- If cells start to differentiate, cells should be discarded immediately
- Cells have to be maintained as colonies and not as single cells. Therefore splitting should be performed as fast as possible, iPSC have to be detached and seeded as clumps. Avoid single cells.
- Plastic coating is critical for even cell distribution; problems with coating often leads to cell detachment, especially at the edges of culture dishes
- Batch effects of critical additives (e.g. holo-transferrin, TGF-β) can lead to differentiation of cells at low passage number

3.6 Differentiation towards the final test system



Culture is essentially as described in Hoelting et al., 2016. PMID: 26933043, with minor changes detailed in Holzer et al., 2022a, PMID: 35689659 and Holzer et al., 2022b, PMID: 35409095.

External SOP document is published in Holzer et al., 2022a.

Neural differentiation:

The human pluripotent stem cell line EPITHELIAL-1 is prepared for neural differentiation on day of differentiation minus 2 (DoD -2) by replating the pluripotent stem cells in a single cell suspension onto Matrigel coated plates in Essential 8 (E8) medium. This E8 is freshly supplemented with 10 ng/ml Rock inhibitor Y-27632.

On DoD0', neural differentiation is started by adding neural differentiation medium KSR and the combination of 4 small molecule pathway inhibitors. From DoD0'-5', Noggin (17.5 ng/ml) and SB-431642 (10 µM) are added and CHIR99021 (1.5 µM), SU5402 (5 µM) and DAPT (γ-Secretase inhibitor IX, 5 µM) are added on DoD2'-9'. From DoD4' onwards, the KSR medium is gradually replaced by N2-S medium.

On DoD9' the cells are cryopreserved in FCS/10% DMSO.

After thawing, cells are cultured in 25% KSR and 75% N2-S supplemented with CHIR99021 (1.5 µM), SU5402 (5 µM) and DAPT (5 µM). Cells are seeded on 96-well-plates in a density of 100.000 cells / cm^2. One hour after seeding, cells have attached to the plate and compounds for the treatment can be added.

Coating of plates:

Frozen matrigel is resolved and diluted 1:40 in cold DMEM/F12 medium. Plates are coated with diluted matrigel (6-well plate: 1 ml/well, 96-well plate: 50 µl/well) and incubated for 30 min at 37°C.

3.7 Reference / link to maintenance culture protocol

Mainentance

principle:

The iPSC line EPITHELIAL-1 is cultured in Essential 8 (E8) medium under feeder-free conditions on Laminin-521 coated plastic dishes. Cells are passaged every 5-7 days, or as soon as the cells reach >80% confluency.

For splitting, cells are detached as clumps using EDTA, diluted 1:35-50 in prewarmed medium (depending on culture confluency) and reseeded in E8 medium on Laminin-521 coated plastic dishes.

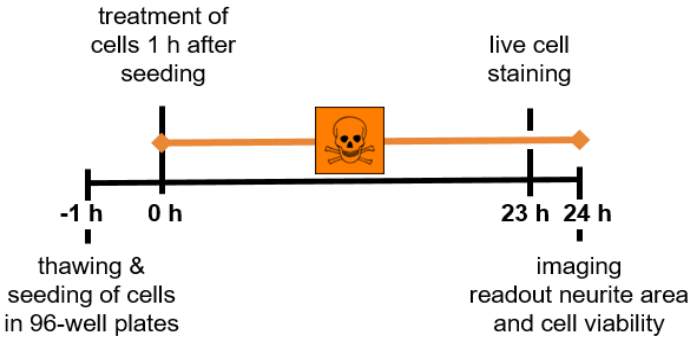
The cells are checked for basic stem cell morphological characteristics as cell growth in defined colonies, the expression of marker genes and proteins like Oct-4 and Nanog and the absence of spontaneously differentiated cells.

As soon as differentiated cells are spotted in the stem cell culture, a new batch of cells is thawed.

4. Definition of the test system as used in the method

Note: *this section refers to the stage of the test system, which is then used for the test method. See scheme for illustration. If you have cells that do not need prior differentiation, give their basic characteristics here. If your test system undergoes significant changes between the maintenance culture and the use for testing, please also fill in section 3.*

4.1 Principles of the culture protocol



The previously differentiated immature peripheral neurons are thawed and seeded on matrigel coated plates (1:40 diluted) in 75 µl medium composed of 75% N2-S medium and 25% KSR medium, supplemented with CHIR99021 (1.5 µM), SU5402 (5 µM) and DAPT (γ-Secretase inhibitor IX, 5 µM) at a density of 100.000 cells/cm^2.
One hour after seeding, treatment compounds are added to the cells in 25 µl of culture medium.
23 h after toxicant application, cells are live-stained with H-33342 and calcein-AM and incubated for 60 min. After 24 h of treatment (including staining), the cells are imaged using a high-content microscope (Cellomics VTI Array Scan).

4.2 Acceptance criteria for assessing the test system at its start

Cells should be attached to the plate (appear flattened at the edges) when the toxicant treatment is applied

> → as the cells are freshly thawed for the test run, there are no quantifiable criteria the culture can be checked for before toxicant treatment. However, neurite growth and appearance of the control cells are checked visually before live staining of the cells.

In general, cells are checked for the expression of the (sensory) neuronal markers Brn3A, Islet-1, peripherin and βIII tubulin (on DoD1, 4 and 7 after thawing).

4.3 Acceptance criteria for the test system at the end of compound exposure

After compound treatment, the negative controls should fulfil the following:
- control cells should have properly grown neurites, neurite area quantification (via Cellomics) has to be > 150.000 in the control wells.

4.4 Variability of the test system and troubleshooting

Causes of variability:
- different differentiations:
  → contaminating, non-neuronal cells might be present in some differentiations for unknown reasons during the differentiation process.
- lots of different plates/flasks:
  → plastic might be different, if the manufacturer delivers from a different/new lot
- differences between the vials of one cell "lot"
- different lots of medium and supplements

differentiation = cells that have all been differentiated from the one iPSC passage and frozen at the same time. Usually one to three 6-well plates are frozen in numerous vials with $8x10^6$ cells/vial.

4.5 Metabolic capacity of the test system

No specific information available.

4.6 Omics characterization of the test system

Transcriptomics data (unpublished) will become available from the originator lab (Leist) upon request.

4.7 Features of the test system that reflect the in vivo tissue

- As neurons of the peripheral nervous system they express peripherin, Brn3A, Islet-1
- They express various neuronal receptors and channels (e.g. purinergic receptors, TRP channels) and especially the tetrodotoxin-resistant voltage-gated sodium channel $Na_v1.8$ which is specifically expressed in dorsal root ganglia
- They are electrophysiologically active and excitable
- Cells do not exhibit the typical pseudo-unipolar morphology of dorsal root ganglion-neurons

4.8 Commercial and intellectual property rights aspects of cells

The cells are not protected by patents or any other licences.

4.9 Reference / link to the culture protocol

Chapter 3 has been answered.

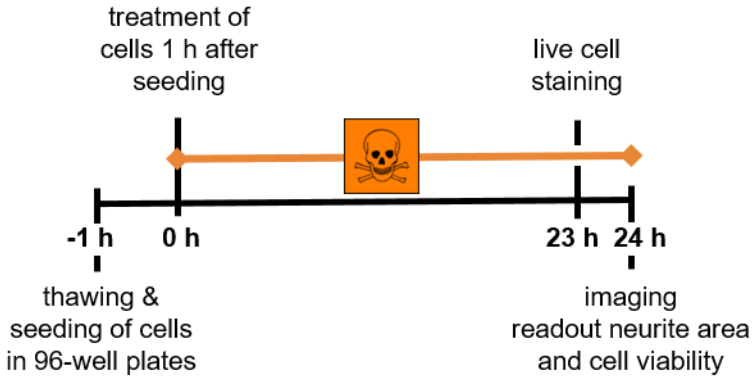
The maintenance is described in the SOP available in Holzer et al., 2022a, supplementary file 1:

https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/stcltm/PAP/10.1093_stcltm_szac031/1/szac031_suppl_supplementary_material_1.pdf?Expires=1658232010&Signature=VHxFr2GRD2-SPR0rJ9AlkxjukzjRH-Xs-2oGQ4S6Nuweb9CgpF82r50vdUdF9oi4lG9dyS1zqzhBbfQHhdiuvqNToX-IjDfJo1l8XMw~OJcTjgqy1MXIKJJJVbl3WppZOm8ZmqqjqgoCUDhTDendIKvn8PIn5maZy-h62YLtX84XWLvChc24UiiJS2nNmkaYfQSjE8MEnWP~M8zAr-KfUGKjizaevM9WbL1c1NReQzKYyBShvbCgregzEuUZHtwfpwqwFpW-FLamTHuUFeABgUc8EaBI1dhnllFXDKDbzJzEgxcXiBohMGC78V4S03sM9ROu3anwl-PQpjQJUsS7Yf7QA__&Key-Pair-Id=APKAIE5G5CRDK6RD3PGA

A lab-internal handling protocol is also available upon request to the Leist-lab.

5. Test method exposure scheme and endpoints

5.1 Exposure scheme for toxicity testing



The previously differentiated immature peripheral neurons are thawed and seeded on matrigel coated plates (1:40 diluted) in 75 µl medium composed of 75% N2-S medium and 25% KSR medium, supplemented with CHIR99021 (1.5 µM), SU5402 (5 µM) and DAPT (γ-Secretase inhibitor IX, 5 µM) at a density of 100.000 cells/cm^2.
One hour after seeding, treatment compounds are added to the cells in 25 µl of culture medium.
23 h after toxicant application, cells are live-stained with H-33342 and calcein-AM and incubated for 60 min.
After 24 h of toxicant treatment (including staining), the cells are imaged using a high-content microscope (Cellomics VTI Array Scan).

5.2 Endpoint(s) of the test method

Test endpoints: 1) neurite area (specific endpoint)
2) cell number
3) % of viable cells (reference endpoint)

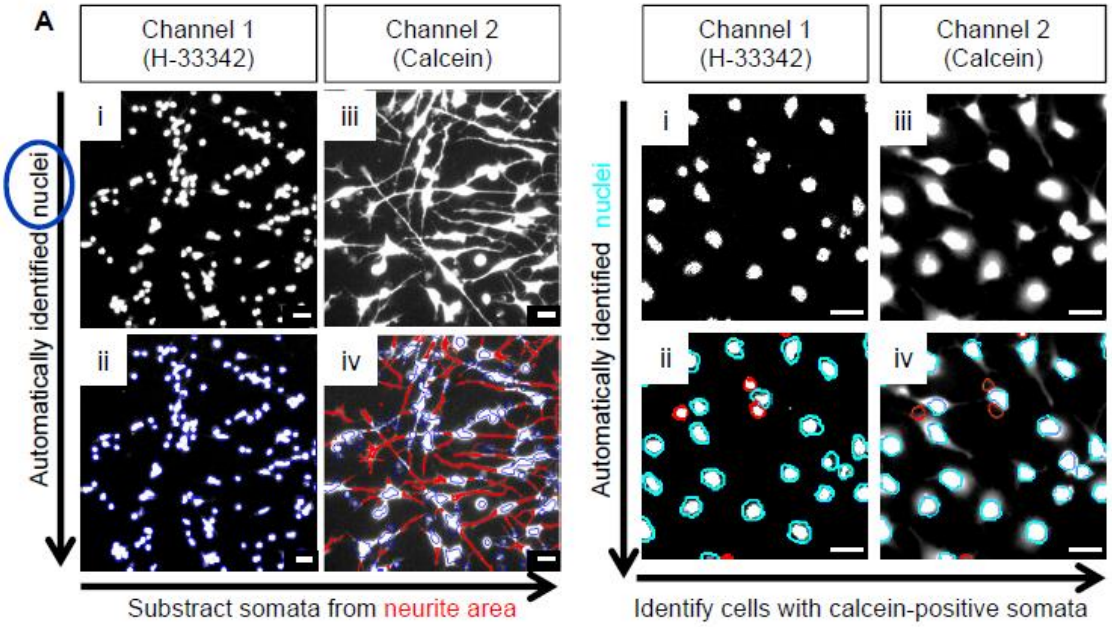
5.3 Overview of analytical method(s) to assess test endpoint(s)

Cells are stained with calcein-AM to mark viable cells. Co-staining with Hoechst H-33342 allows the identification of any cell.

Cells are stained for 30 min at 37°C and 5% CO2 in the incubator.

The cell staining is imaged in a Cellomics Array Scan VTI HCS reader.
Hoechst H-33342 staining is imaged in channel 1 (UV-Hoechst); calcein staining is imaged in channel 2 (Green-FITC). Exposure times are set manually.

To measure the neurite area, the software acquires the Hoechst signal in channel 1 to identify the cells as objects (via identification of the nuclei), and the calcein-AM signal in channel 2 to measure neurite area. Double positive cells are counted as viable.



5.4 Technical details (of e.g. endpoint measurements)

Quantification of neurite outgrowth

An automated microplate reading microscope (Array-ScanII HCS Reader, Cellomics, PA) equipped with a Hamamatsu ORCA-ER camera (resolution 1024 x 1024; run at 2 x 2 binning) was used for image acquisition. Ten fields per well were imaged. Images were recorded in 2 channels using a 20x objective and excitation/emission wavelengths of 365 ± 50/535 ± 45 to detect H-33342 in channel 1 and 474 ± 40/535 ± 45 to detect calcein in channel 2. In both channels, a fixed exposure time and an intensity histogram-derived threshold were used for object identification. Neurite pixels were identified using the following image analysis algorithm: nuclei were identified as objects in channel 1 according to their size, area, shape, and intensity which were predefined on untreated cells using a machine-based learning algorithm, and manual selection of nuclei to be classified as intact. The nuclear outlines were expanded by 3.2 µm in each direction, to define a virtual cell soma area (VCSA) based on the following procedure: The average width of the cytoplasm ring (distance nucleus - cell membrane) of LUHMES cells was experimentally determined to be 2.3 µm. Size irregularities were not always due to growing neurites, as determined by combined F-actin/tubulin beta-III staining. To avoid scoring of false positive neurite areas, the exclusion ring (VCSA) was made bigger than the average cell size. Then, we used two control compounds (U0126 and bisindolylmaleimid I) to vary the expanded outlines from 0.6 to 4 µm. We found 3.2 µm to be optimal both to detect neurite growth over time and to identify reduced neurite growth with high sensitivity. All calcein-positive pixels of the field (beyond a given intensity threshold) were defined as viable cellular structures (VCSs). The threshold was dynamically determined for each field after flat field and background correction and intensity normalization to 512 gray values and was set to 12% of the maximal brightness (channel 63 of 512). The VCS defines the sum of all somata and neurites without their assignment to individual cells. In an automatic calculation, the VCSAs, defined in the H-33342 channel, were used as filter in the calcein channel and subtracted from the VCS. The remaining pixels (VCS - VCSA) in the calcein channel were defined as neurite area.

Quantification of individual viable cells by imaging

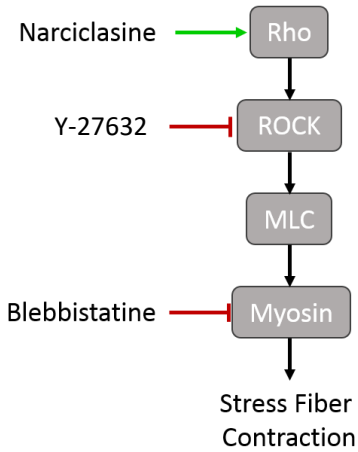
For a quantitative assessment of viable cells, the same images used to assess neurite area were analyzed using another image analysis algorithm: nuclei were identified in channel 1 as objects according to their size, area, shape, and intensity. Nuclei of apoptotic cells with increased fluorescence were excluded. A VCSA was defined around each nucleus by expanding it by 0.3 µm into each direction. Calcein-AM staining, labeling live cells, was detected in channel 2. The algorithm quantified the calcein intensity in the VCSA areas. Cells having an average calcein signal intensity in the VCSAs below a predefined threshold were classified by the program as “not viable”. Valid nuclei with a positive calcein signal in their cognate VCSA were counted as viable cells. A positive calcein signal was based on measurements of the average intensity (normal cells: 1300 ± 115, threshold: < 50) and the total integrated intensity (normal cells: 186,000 ± 23,600, threshold < 1000) of cells.

5.5 Endpoint-specific controls / mechanistic control compounds (MCC)

Endpoint-specific control for neurite growth inhibition:
Vincristine: microtubule toxicant
Colchicine: microtubule polymerization inhibitor
Cytochalasin D: actin polymerization inhibitor
Narciclasine: activates Rho

Endpoint-specific control for neurite growth enhancement:
Y-27632: ROCK inhibitor
Blebbistatine: inhibits myosin II

Rho/ROCK/LIM kinase/cofilin pathway:
induces actin polymerization, key regulator of the cytoskeleton and cell polarity



5.6 Positive controls

Positive control: narciclasine (50 nM final concentration)

5.7 Negative and unspecific controls

Solvent control: 0.1% DMSO final concentration (standard)

Up to 0.5% DMSO final concentration was tested in this test system and can be also used as solvent control

The concentration of the solvent control is aligned with the highest final DMSO concentration in wells treated with toxicants which is normally at 0.1% DMSO.

Further possible negative control compounds: e.g. mannitol, paracetamol

5.8 Features relevant for cytotoxicity testing*

Cell death can easily be quantified.

Distinguishing slowed proliferation from cell death is not an issue for this test system as the cells are mainly post-mitotic at the time point of toxicant exposure.

5.9 Acceptance criteria for the test method

Positive control narciclasine (50 nM):
Neurite area ≤ 75% of DMSO control
Viability ≥ 90% of DMSO control (or not significantly changed)

Negative control DMSO:
Neurite area ≥ 150.000

5.10 Throughput estimate*

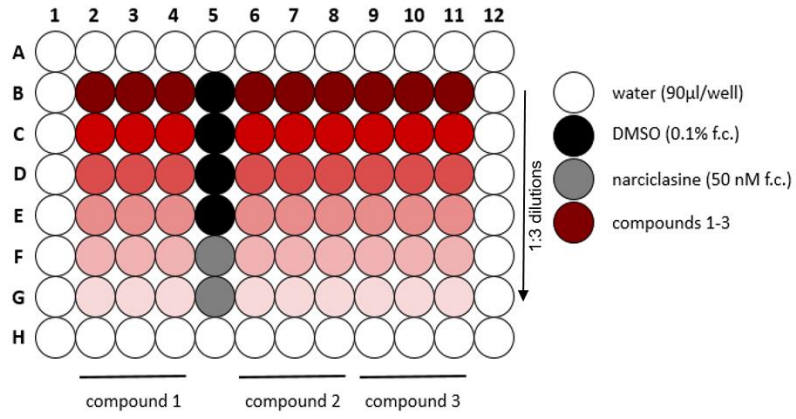
Data point = one biological replicate (→ usually 3 technical replicates); each concentration/condition of a compound counts as data point

1440 data points per month

3 compounds per plate, 6 different concentrations of each compound per plate, 3 technical replicates per plate (see figure) → 18 data points (1 plate)

5 plates can be done per day (correlates to 15 compounds → 90 data points)
→ 4 days of readout per week → 360 data points per week; 4 weeks per month → 1440 data points per month.

Typical plate layout:



6. Handling details of the test method

6.1 Preparation / addition of test compounds

- Compounds are stored according to the manufacturer's instructions (e.g. 4°C, room temperature, -20°C, -80°C).

- Preferable solvent is DMSO. The used DMSO is stored in a lightproof, air-tight bottle at room temperature.

- After dissolving the compounds which are delivered in a solid/powder form, all compound solutions are aliquoted into volumes sufficient for one experiment (i.e. one biological replicate). In this way repeated freezing and thawing and therefore damaging the compound's stability and efficiency can be avoided.

- final concentration of DMSO is 0.1%

- For conducting an experiment, a compound aliquot is thawn and diluted with culture medium (75% N2-S medium, 25% KSR medium, supplemented with CHIR99021 (1.5 µM), SU5402 (5 µM) and DAPT (5 µM)

- All compound dilutions in the master plate contain 0.4% DMSO, so that a final concentration of 0.1% DMSO is reached on the cells. The highest compound concentration is diluted with medium 1:250 without DMSO as 0.4% is already reached with the DMSO the compound is solved in. The serial dilution is done with culture medium and 1% DMSO.

- The compound dilutions (25 µl each) are added to the cells using a multichannel pipette, 6 filter tips at a time. Pipetting has to be performed slowly.

6.2 Day-to-day documentation of test execution

Plate maps are defined prior to the experiment and documented in the lab book and files (Excel files) are stored on the work group server.
Concentrations and compound dilutions are calculated prior to the experiment.
Experimental procedures are noted manually in a paper lab book.

6.3 Practical phase of test compound exposure

The experimenter plans the experiment according to Cellomics microscope availability (has to be booked in advance) and availability of a sufficient number of cells.

Pipetting errors are marked directly on the plate maps and are documented in the lab book.

The paper lab book is taken to cell culture rooms and errors are documented in there right away.

The technical replicates were pipetted from left to right.

6.4 Concentration settings

3 compounds per plate

As default a serial dilution 1:3 is used, i.e. a concentration range of 1024-fold is covered (e.g. from 100 µM → 100 nM).

Serial dilutions of compounds are prepared in a separate deepwell-plate, from which 25 µl are transferred to the according plate with attached cells using a multichannel pipette.

Dilution steps can be adapted (e.g. 1:1.5, 1:4)

6.5 Uncertainties and troubleshooting*

- Compound solubility in stock and during dilution is too low (stock solved in 100% DMSO, final concentration of the solvent on the cells is 0.1% DMSO)
- Some compounds show autofluorescence and interfere with the detection of calcein-AM or H-33342.
- To prevent negative edge effects, only the inner 60 wells of a 96-well plate are used and the edge wells were filled with PBS or water.
- Focusing failure of Array Scan VTI HCS Reader (Cellomics, PA) can be a problem that produces outliers; as well as imaging only one channel.
- Highly trained/automated handling with multichannel and multistepper pipette is necessary to achieve little variance.
- Operators can get trained within 2-4 weeks. Cell seeding and medium change should be performed as fast as possible to keep cells as short as possible at room temperature. The more practice an operator has, the faster the critical steps can be performed.
- Substances are added when pipette tips are touching the wall of the wells right above the medium surface. When the substance solution is pipetted too high above the medium surface, the droplet may just stick to the wall of the well without flowing down into the medium.

6.6 Detailed protocol (SOP)

Protocol n° 218 in DB-ALM data base

Holzer et al., 2022a, supplementary file 1 (p. 32-57)

https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/stcltm/PAP/10.1093_stcltm_szac031/1/szac031_suppl_supplementary_material_1.pdf?Expires=1658232010&Signature=VHxFr2GRD2-SPR0rJ9AlkxjukzjRH-Xs-2oGQ4S6Nuweb9CgpF82r50vdUdF9oi4lG9dyS1zqzhBbfQHhdiuvqNToX-IjDfJo1l8XMw~OJcTjgqy1MXIKJJJVbl3WppZOm8ZmqqjqgoCUDhTDendIKvn8PIn5maZy-h62YLtX84XWLvChc24UiiJS2nNmkaYfQSjE8MEnWP~M8zAr-KfUGKjizaevM9WbL1c1NReQzKYyBShvbCgregzEuUZHtwfpwqwFpW-FLamTHuUFeABgUc8EaBI1dhnllFXDKDbzJzEgxcXiBohMGC78V4S03sM9ROu3anwl-PQpjQJUsS7Yf7QA__&Key-Pair-Id=APKAIE5G5CRDK6RD3PGA

6.7 Special instrumentation

The method requires a Cellomics Array Scan VTI HCS high content reader that may not be present in the standard lab.

6.8 Possible variations*

a) further additional endpoints:
- metabolic activity (resazurin reduction)
- glutathione levels

- staining of tubulin
- analysis of differentiation markers by qPCR or immunostaining

b) other analytical endpoints:
cell viability by:
- fluorimetric measurement of resazurin conversion
- measurement of extracellular LDH
- measurement of luminescence indicating ATP content

c) other exposure:
- compound can be washed out → acquisition on day 2
- longer exposure is possible
- later exposure is possible (from day 3 on) in order to measure effects on more mature neurons

*6.9 Cross-reference to related test methods**

There is no related test.

7. Data management

7.1 Raw data format

Raw data is extracted by copy-paste in Excel files (example file available).

Data from all technical replicates are collected in one file.

7.2 Outliers.

1. Mathematical procedures to define outliers have not been defined. Data points that 'look' very far off are discarded. Biological outliers do practically not exist, most far data points are the result of technical problems (focus not found, only one channel imaged, etc.)
2. All raw data (incl. outliers) are stored.
3. Technical outliers make up 1-0.1%.

7.3 Raw data processing to summary data

- Array Scan VTI HCS Reader (Cellomics, PA) takes images (optionally bitmap or tiff-format; 512 x 512 pixels, 8bit or 16bit)

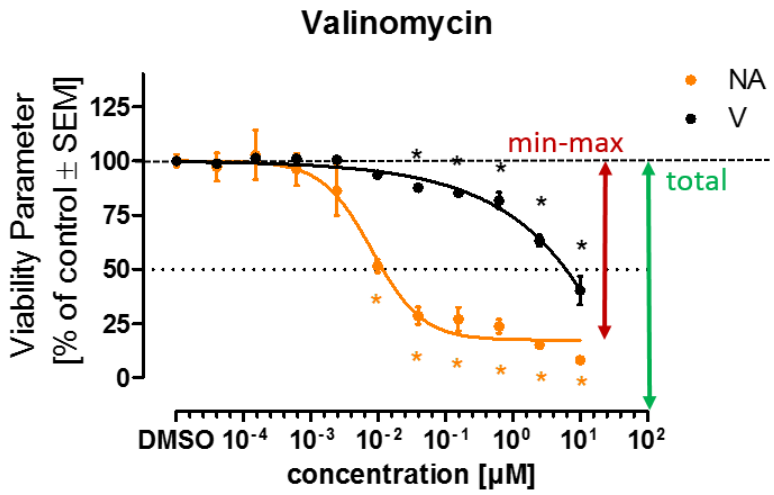
- Images are locally analyzed using the Array Scan software, algorithms to quantify neurite area, total cell count (nuclei) and viable cell count.

- data are copy-pasted into an Excel sheet, further analysis is done with Excel + GraphPad Prism

7.4 Curve fitting

The data are analyzed with Excel and represented with GraphPad Prism.

For the concentration curve, a nonlinear regression fit is calculated. The fitting method is least squares. If a non-linear curve fit is not possible, a linear curve fit is performed. The curve deriving from the fit is a 4-parameter log function. To calculate e.g. the EC50 value, this log-function is solved for y=50% of the total scale, not for 50% of the min-max scale (see example below). Treated concentrations are analyzed for deviation from control. Sometimes it is analyzed whether the deviation of neurite growth is different from the deviation of viability. This is done by two-way ANOVA + Tukey-Kramer post hoc testing. Statistics applied are one-way ANOVA (and nonparametric) with Dunnett's post test.



EC50(NA)= 0.01 µM
EC50(V)= 6.16 µM

BMC values with their upper and lower confidence intervals (BMCU and BMCL) are calculated via the publically available online software:
http://invitrotox.uni-konstanz.de/BMCeasy/

*7.5 Internal data storage**

The data are firstly stored on the microscope computer and then exported to other servers (lab group server and university server), which are back-upped regularly.

7.6 Metadata

The metadata are documented, stored and exported as text document (log)-files to the according scheme: (local PC)_descriptor(date and time)_XXX.log:

The following metadata are stored:

- cellinsight-pc_160429130003_AutomationControllerIni
- cellinsight-pc_160429130003_kineticprotocol
- cellinsight-pc_160429130003_protocol
- cellinsight-pc_160429130003_scan
- cellinsight-pc_160429130003_ScanIni
- cellinsight-pc_160429130003_spooling
- cellinsight-pc_160429130003.spooled

7.7 Metadata file format

Metadata files are available.

8. Prediction model and toxicological application

8.1 Scientific principle, test purpose and relevance

Immature peripheral neurons used in this test method represent the (developing) peripheral nervous system.

The test method therefore measures adverse effects on peripheral neurons that directly or subsequently affect neurite growth and integrity or the cell viability in general.

The test method not only predicts the hazard to induce developmental neurotoxicity but also to induce neurotoxicity in mature peripheral neurons, as these are highly dependent on an intact cytoskeleton due to their enormous length. Any adverse interference with the cytoskeleton in the state of developing neurons might therefor also present an adverse interference in mature peripheral neurons. Therefore, this test method can be related to adverse human outcomes like peripheral neuropathies.

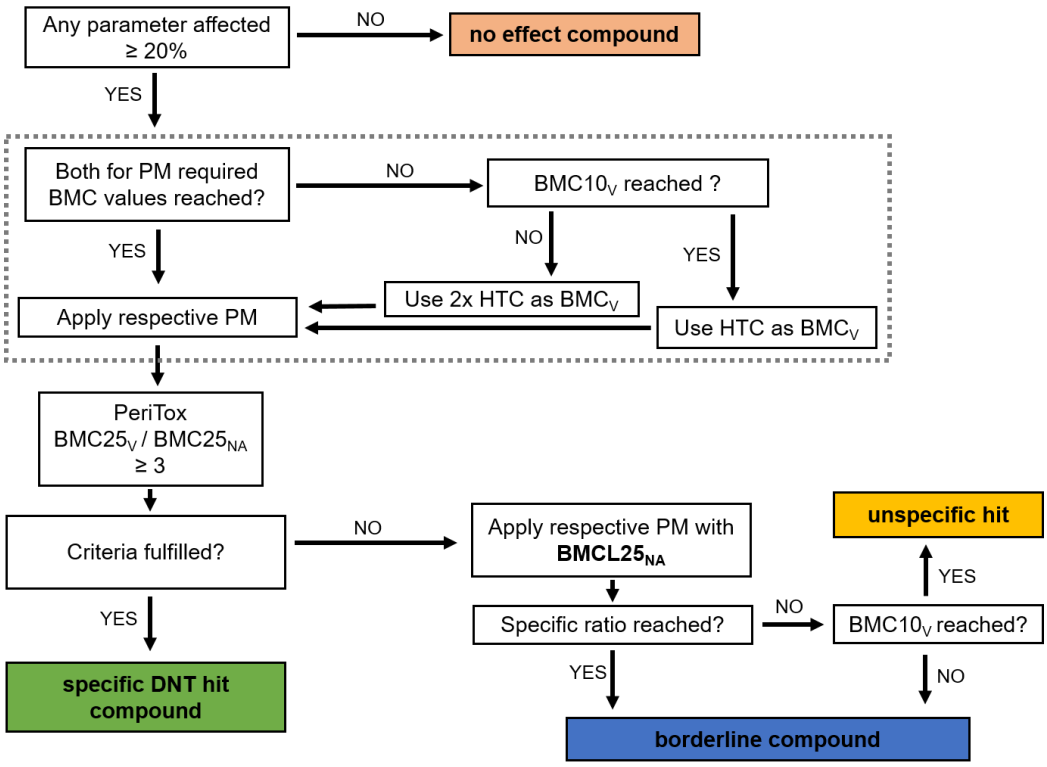
8.2 Prediction model

Three different models are used:

1. prediction model for screening:
hit = decrease/increase in neurite area while viability is not changed (compare to narciclasine positive control:
Neurite area ≤ 75% of DMSO control
Viability ≥ 90% of DMSO control

2. prediction model for compound hazard evalaution:
hit confirmation testing; BMC25 Viability (V) / BMC25 Neurite Area (NA) ≥ 3 → specifically neurotoxic

3. prediction model for borderline compounds:
A ratio of BMC25 Viability (V) / BMC25 Neurite Area (NA) ≥ 3 is considered a borderline hit. In some scenarios the viability does not reach the BMC25 necessary for the ratio calculations. In this case the highest tested concentration (HTC) was used. Schematic representation of the complete prediction model is shown in scheme below.



8.3 Prediction model setup

To design the initial prediction model for the PeriTox test, we took the following steps:
(a) use of the "ratio" of EC50 (viability)/EC50 (neurites) as the primary endpoint
(b) measurement of this value for "unspecific toxicants" (the uncoupler CCCP, SDS, Triton-X100, and the topoisomerase inhibitor etoposide; the average ratio was 1.376±0.39)
(c) definition of a "noise band" (4SD from the average of the ratios of these compounds)
(d) definition of compounds with a ratio outside the noise band (EC50 ratio of >3) as "neurite specific."
The use of the updated prediction model

BMC25 Viability (V) / BMC25 Neurite Area (NA) ≥ 3

was validated by comparing classifications derived by the initial and the updated prediction model. The reason for updating the prediction model was that a decrease of 50% in neurite area and viability cannot always be achieved with our range of test concentrations but to reach an effect of 25% is more reasonable. The initial prediction model therefore often used EC50 values that were only based on a pure mathematical curve fit. However, the now used BMC25 is more related to the data that was practically obtained.

The prediction model has been applied to screen 80 compound library of NTP (Delp et al., 2018). The prediction model including the borderline classification has been applied to screen a 120 compound library (Masjosthusmann et al., 2020).

8.4 Test performance

Some background on the test performance is given in chapters 8.2/8.3 (prediction model).

Several performance parameters for the test were obtained in several separate evaluation rounds.

A first evaluation was done during the first publication of the model and its applications (Hoelting et al. 2016). Here, a panel of well-selected positive and negative controls have been tested. Accordingly, the specificity was 100% and the sensitivity was > 90 %. In dedicated experiments, S/N ratios of > 20 and a z' of > 0.5 have been determined. Operator reproducibility was shown in Hoelting et al. 2016 (Supplementary 3) for the compound colchicine.

The test has been used in screening campaigns, and real-live performance data under broader screen conditions have been obtained. The different performance data need to be considered, when a compound is a hit in a screen, or whether it has been specifically evaluated in a hit follow-up or a mechanistic project.

A first screen application has been the NTP80 screen (80 compounds provided by the US NTP). Data are published Delp et al. 2018.

A second screen application has been the cross systems case study of the EU-ToxRisk project. The baseline variation is indicated in Krebs et al., 2020. Moreover, an overview is given for 19 compounds on the BMC/BMCL ratio as measure of readout certainty.

A third screen was performed in the context of the EFSA DNT test battery evaluation with 120 compounds (Masjosthusmann et al. 2020). From this screen the following performance indicators were obtained:

A: **Sensitivity: 100%**

→ With PeriTox as standalone assay in 17 `tool negatives´ tested (Masjosthusmann et al. 2020).

B: **Specificity: 82.7%**

→ With a selected set of 27 positive compounds with evidence for DNT (Masjosthusmann et al. 2020).

C: **Baseline variation (intra-experimental)**

Neurite outgrowth: 7.8 ± 4.3%

Cell viability: 5.5 ± 3.4%

D: **Baseline variation (inter-experimental)**

Neurite outgrowth: 21%

Cell viability: 15.7%

E: **Variation of a positive control run on each (inter-experimental)**

Neurite outgrowth: 17.5%

Definition of values C-E

C: **Baseline variation (intra-experimental)** is the mean CV±SD of the CV of all replicates of the solvent control from each experiment across n>200 experiments.

D: **Baseline variation (inter-experimental)** is the variability across all independent experiments (n>200) after normalization based on the response of the lowest test concentration. It was assumed that the lowest test concentration does not affect any of the endpoints measured.

E: **Variation of a positive control run on each (inter-experimental)** is the variability of the positive control across all independent experiments (n>40) after normalization. Example for a positive control that on average reduced the specific endpoint down to 40% (relative to solvent control) and a calculated variability of 50%: 0.5 x 40% = +/- 20% → The positive control with mean of 40% varies from 20% to 60%.

8.5 In vitro – in vivo extrapolation (IVIVE)

1. Lipid and Albumin content is not known. Medium used during toxicant treatment is as follows:

25% KSR:
Knockout DMEM with
15 % knockout serum replacement,
2 mM Glutamax,
0.1 mM MEM non-essential amino acids and
50 µM beta-mercaptoethanol

75% N2:
DMEM/F12 medium
1 % Glutamax
1.55 mg/ml glucose
0.1 mg/ml apotransferrin
25 µg/ml insulin
100 µM putrescine
30 nM selenium
20 nM progesterone)

2. The test has not been used for IVIVE or other use of potency information.
3. No special considerations known.

*8.6 Applicability of test method**

Test is sensitive to cytoskeletal toxicants, some signaling modifiers and flame retardants. Polycyclic aromatic hydrocarbons (PAH) and HDAC inhibitors have no effect.

8.7 Incorporation in test battery*

a) Strengths:

- Medium to high throughput
- Automated microscopy

b) compared to UKN4 (which quantifies neurite outgrowth of central neurons), UKN5 measures neurite outgrowth specifically of peripheral neurons. This was shown by treatment with MPP+ which is transported by the dopamine transporter (DAT) and had effect in UKN4, but not UKN5 as peripheral neurons which lack the DAT transporter. Furthermore, toxicants known to specifically induce peripheral neuropathies, like proteasome inhibitors (e.g. Bortezomib) or acrylamide were shown to have neurite specific effects in UKN5 but not in UKN4 (Hoelting et al. 2016)

c) specific effects peripheral neurons. The test method is currently used in the setup of a DNT test battery.

d) Preferential use in first tier, no complementary assays required for the assessment of chemical effects on the endpoints investigated by this test method.

9. Publication / validation status

9.1 Availability of key publications

Stem cell-derived immature human dorsal root ganglia neurons to identify peripheral neurotoxicants.
Hoelting, L. et al. Stem Cells Transl Med, 2016. PMID : **26933043**

A high-throughput approach to identify specific neurotoxicants / developmental toxicants in human neuronal cell function assays.
Delp, J. et al. Altex, 2018. PMID : **29423527**

The EU-ToxRisk method documentation, data processing and chemical testing pipeline for the regulatory use of new approach methods.
Krebs, A. et al. Arch. Toxicol., 2020. PMID: **32632539**

Establishment of an a priori protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of developmental neurotoxicity.
Masjosthusmann, S. et al. EFSA Supporting Publications. 2020; 17(10): 1938E.

Neurodevelopmental toxicity assessment of flame retardants using a human DNT in vitro testing battery.
Klose, J. et al. Cell Biol Toxicol. 2021; PMID: **33969458**

Generation of nociceptor-enriched sensory neurons for the study of pain-related dysfunctions.
Holzer, A-K. et al. Stem Cells Transl Med., 2022a; PMID: **35689659**

Specific attenuation of purinergic signaling during bortezomib-induced peripheral neuropathy in vitro.
Holzer, A-K. et al. Int J Mol Sci., 2022b; PMID: **35409095**

9.2 (Potential) linkage to AOPs*

Test method could be potentially linked to the following AOPs in AOPwiki:

- AOP 249 : Microtubule interacting drugs lead to peripheral neuropathy
  → Adverse Outcome: Sensory axonal peripheral neuropathy

9.3 Steps towards mechanistic validation*

a) Cells express typical sensory neuronal markers, are of human origin and form a network

b) Tubulin plays a major role in neurite outgrowth and if the dynamic instability of microtubules is inhibited by compounds like colchicine/vincristine/taxol neurite outgrowth is reduced. If the Rho/Rock pathway is activated, neurite outgrowth is enhanced.
c) A formal mechanistic validation has not been performed. Reversibility has been shown (Hoelting et al., 2016).

d) The test rather covers a fundamental neurodevelopmental process than a key event (Smirnova 2014, Bal-Price 2015 (ISTNET))

9.4 Pre-validation or validation*

To date, 145 unique compounds (as defined by unique DTXSIDs) have been tested successfully in this assay.

No formal OECD 34 validation study has been done (eg., ring trials with a standard set of known positive and negative controls).

In total, >200 different compounds were tested in the PeriTox assay. The test method was developed using a compound training set (Hoelting et al. 2016). It was used for an 80 compound screening library from the US National Toxicology Program (NTP) (Delp et al. 2018). The test method was part of a DNT hazard assessment for 120 compounds in a DNT testing battery. The later compound set includes potential DNT positive and DNT negative compounds (Masjosthusmann 2020).

9.5 Linkage to (e.g. OECD) guidelines / regulatory use*

Test is not linked to regulatory guidelines.

10 Test method transferability*

10.1 Operator training*

Experiences are required in:
- cell culture
- multichannel/multistep pipetting
- handling of Array Scan VTI HCS Reader (Cellomics, PA) and its software
- Microsoft Excel
- GraphPad Prism

Operator is trained and guided by a highly experienced instructor. Approximately 4 weeks will be needed for a smooth assay performance.
Learning iPSC culture and cell differentiation takes several months.

10.2 Transfer*

The assay hasn't been transferred or applied in other labs.

11. Safety, ethics and specific requirements

11.1 Specific hazards; issues of waste disposal

No specific requirements.

11.2 Safety data sheet (SDS)

SDS are available in the university DaMaRIS database (**Da**ngerous **Ma**terials **R**egistry **I**nformation **S**ystem).

11.3 Specific facilities / licenses

Work requires S1 cell culture laboratories (genetically modified cells).
No specific facilities are required.
No specific ethical approval is required.

11.4 Commercial aspects / intellectual property of material / procedures*

To our best knowledge, no elements needed to conduct the experimental part of the test method are protected. Programs used to conduct the analysis of the data (Microsoft Excel and GraphPad Prism) need to be purchased or obtained by license agreement, however data analysis and plotting can be done with other, freely available tools.