EXTERNAL SCIENTIFIC REPORT

APPROVED: 13 April 2018

doi:10.2903/sp.efsa.2018.EN-1410

Literature review and appraisal on alternative neurotoxicity testing methods

[1]Stefan Masjosthusmann, [1]Marta Barenys, [1]Mohamed El-Gamal, [3]Lieve Geerts, [2]Laura Gerosa, [2]Adriana Gorreja, [1]Britta Kühne, [2]Natalia Marchetti, [1]Julia Tigges, [2]Barbara Viviani, [3]Hilda Witters, [1]Ellen Fritsche

[1]IUF - Leibniz Research Institute for Environmental Medicine, Auf'm Hennekamp 50, 40225 Düsseldorf, Germany

[2]Università degli Studi di Milano (UMIL), Dipartimento di Scienze Farmacologiche e Biomolecolari - DiSFeB, Via Balzaretti 9, 20133 Milan, Italy

[3]Flemish Institute for Technological Research (VITO), Environmental Risk & Health, Boeretang 200, B-2400 Mol, Belgium

ABSTRACT

The goal of this review was the evaluation of information on assessment methods in the field of alternative neurotoxicity (NT) testing. We therefore performed a systematic and comprehensive collection of scientific literature (in English) from the past 27 years until mid of 2017 on state of the art alternative testing methods including in vitro test methods, *in silico* methods and alternative non-mammalian models. This review identified a variety of test methods that have the ability to predict NT of chemicals based on predefined key NT endpoint categories (27). Those endpoint categories were derived from the Mode of Action (MoA) of known human neurotoxicants. Pre-evaluated MoAs of human neurotoxicants allowed the identification of performance characteristics with regard to the ability of a test system to correctly predict a chemical effect on an endpoint category. The most predictive in vitro model that covers a large variety of endpoint categories are primary rodent cells or tissues. Human based systems derived from induced pluripotent stem cells (iPSC) are promising and warrant human relevance. There is however not yet sufficient data on these models to demonstrate their suitability to reliably substitute primary rodent cells for NT testing purposes. Test methods for glia toxicity are rare and glia endpoint categories are clearly underrepresented. Therefore, a focus for future method development should be placed on glia, astrocytes, oligodendrocytes and microglia based models, preferably in a co-culture se up. The review on *in silico* methods, resulted into 54 QSARs publications, relevant for NT, of which 39 on blood brain barrier (BBB) permeation. The QSARs available in the publications were developed from data on drugs and chemicals, but there appears a limited set of experimental data for chemicals and pesticides on blood-brain barrier passage. The evaluation of NT methods using alternative whole organism approaches demonstrated a majority of data for *C. elegans* (nematode species), represented with high true prediction (96%). The main endpoint category was inhibition of cholinergic transmission, with specific endpoints for AChE activity and motor activity, the latter confirming the added value of a whole organism approach among alternative models. Though *D. rerio*, the zebrafish model appeared a

promising model for DNT studies with numerous advantages, it was poorly evaluated for NT endpoints. Next to the need for standardized protocols using *C. elegans* as a test organism, the zebrafish model needs further exploration for NT relevant endpoints. In conclusion, a NT alternative test battery covering identified and relevant MoA for NT is recommended. Therefore, test methods with relevant controls and standard operation procedures have to be set up for covering most important MoA. To link the human in vitro testing to rodent *in vivo* studies and validate the stem cell-derived systems, it is advised to include rodent primary cultures into the studies. For more complex, behavioural readout, effects in alternative organisms should be combined with electrophysiological assessments in vitro.

KEY WORDS: literature review, neurotoxicity, in vitro, in silico, alternative organism, mode of action

Question number: EFSA-Q- 2015-00822

Correspondence: pesticides.ppr@efsa.europa.eu

Suggested citation: Masjosthusmann S, Barenys M, El-Gamal M, Geerts L, Gerosa L, Gorreja A, Kühne B, Marchetti N, Tigges J, Viviani B, Witters H, Fritsche H, 2018. Literature review and appraisal on alternative Neurotoxicity testing methods. EFSA supporting publication 2018:EN-1410. 125 pp. doi:10.2903/sp.efsa.2018.EN-1410

ISSN: 2397-8325

Table of contents

BACKGROUND AS PROVIDED BY EFSA

Neurotoxicity refers to any adverse effect of exposure to chemical, biological or physical agents on the structure or functional integrity of the developing or adult nervous system (Faqi, 2013). Many common substances are neurotoxic, including lead, methylmercury and pesticides (AltTox.org).

Neurotoxicity studies are indicated in the list of "Toxicological and metabolism studies" under Part A, Section 5 of data requirements for active substances listed in Commission Regulation (EU) No 283/2013, in accordance with regulation (EC) No 1107/2009 concerning the placing of plant protection products on the market. Such studies should be performed in rodents in the case of regulatory applications of active substances having structures similar or related to those capable of inducing neurotoxicity, with specific indications of potential neurotoxicity and/or with a neurotoxic mode of pesticidal action. In addition, neurotoxicity studies "should provide sufficient data to evaluate the potential neurotoxicity of the active substance (neurobehavioral and neuropathological effects) after single and repeated exposure" (acute toxicity studies point 5.2, short-term toxicity studies point 5.3, long term toxicity and carcinogenicity studies point 5.5 and reproductive toxicity studies point 5.6).

The recognised test methods for the evaluation of the neurotoxicity potential of chemicals (including pesticides) are the OECD Guideline 424 (Neurotoxicity studies in rodents) and 426 (Developmental Neurotoxicity Studies). However, both these methods use complex *in vivo* tests which are often too laborious and expensive and might also not well reflect the human situation because of inter-species variation (Leist *et al*, 2013). On the other hand, the data requirements mentioned above clearly indicate that "tests on vertebrate animals shall be undertaken only where no other validated methods are available. Alternative methods to be considered shall include in vitro methods and *in silico* methods. Reduction and refinement methods for *in vivo* testing shall also be encouraged to keep the number of animals used in testing to a minimum." It is now recognised that the future of chemical safety assessment must move away from animal tests towards a combination of complementary approaches that address functional mechanistic endpoints tied to adverse outcomes of regulatory concern. The Adverse Outcome Pathway (AOP) concept can assist in the selection of the most important tests to use in integrated testing strategy (ITS), which are expected to efficiently combine different information sources in a quantifiable fashion for regulatory risk assessment (Bal-Price *et al*, 2015a).

The present Call is based on EFSA's draft 2016 Work Programme for grants and procurements in science. This call is launched without prejudice to the approval of the 2016 Work Programme and the 2016 Budget by the EFSA Management Board on 03/12/2015.

TERMS OF REFERENCE AS PROVIDED BY EFSA

This contract/grant was awarded by EFSA to: IUF – Leibniz Research Institute for Environmental Medicine

Contractor/Beneficiary: IUF – Leibniz Research Institute for Environmental Medicine, with subcontracting VITO – Flemish Institute for Technological Research and UMIL – University of Milano.

Contract/grant title: Literature review and appraisal on alternative Neurotoxicity (NT) testing methods

Contract/grant number: OC/EFSA/PRAS/2015/07

1. Project Summary

1.1. Background

Pesticides are globally used substances for controlling undesirable pests such as insects, weeds, fungi and rodents. Most pesticides are indiscriminate, implying toxicity to non-targeted species, including humans. As most pesticides', especially insecticides' targets involve the nervous system it is not surprising that a number of these compounds can cause neurotoxicity in mammals. This family of chemicals includes organophosphates, carbamates, pyrethroids, organochlorines, neonicotinoids, and other compounds. In addition to insecticides, some herbicides and fungicides as well as other non-pesticide related compounds like metals, industrial chemicals, solvents, natural toxins, pharmaceutical drugs and drugs of abuse also possess neurotoxic properties. The effects of pesticides on the nervous system may be involved in their acute toxicity, as in case of most insecticides, or are suspected to contribute to chronic neurodegenerative disorders, most notably Parkinson's disease. This was comprehensively reviewed in Costa et al., 2008.

Socioeconomic costs of overall neurotoxicity –including acute, chronic, central and peripheral nervous system effects- are difficult to estimate. However, there is sufficient indication in the literature that environmental exposure towards chemicals contributes to neurodegenerative diseases amongst those Alzheimer's (AD) and Parkinson's Disease (PD) as well as parkinsonian's syndromes or other degenerative motor syndromes as a form of chronic neurotoxicity(Landrigan *et al*, 2005; Tanner *et al*, 2014). A cross-European evaluation has revealed that on average each patient suffering from neurodegenerative disease costs € 28.000/year (Jönnson & Wim, 2009). This number is similar to the cost for PD in the United States (US) with $ 24.000/patient/year (Landrigan *et al*, 2005)[7]. In the US, just the treatment cost for PD range from 12 to 25 billion US $. Considering the increasing prevalence of neurodegenerative diseases in developed countries and the so far qualitatively confirmed (Costa *et al*, 2008), but quantitatively unknown contribution of chemicals towards such diseases, it seems necessary to prevent chemical-induced chronic neurotoxicity as well as the other forms of neurotoxicity as indicated above.

Animal experiments are currently the gold standard for NT testing (OECD TG 424, TG 418, TG 419 and TG 426). An iterative assessment/testing strategy is recommended and the first animal data for NT assessment are most often provided by standard single dose (OECD TG 402, TG 403, TG 420, TG 423 and TG 425) or repeated dose toxicity studies (OECD TG 407, TG 408). These studies include clinical observations and morphological examinations, which can reveal adverse effects on the nervous system. If evidence of a direct effect on the nervous system is provided by these standard, single or repeated dose toxicity studies, NT testing by the specific NT guidelines indicated above may be conducted.

Such presently practiced animal tests for NT are problematic in several ways: (i) they are ethically questionable; (ii) they are very expensive, laborious and time requiring since identification and detection of every possible change of the nervous system cannot be ensured by one single method due to the multiplicity of possible effects; (iii) they require highly trained and competent parties because of the complex array of behavioural, neurological, neurochemical, histopathological and morphological approaches needed to acquire neurotoxicity data. Furthermore, while such guideline studies are currently necessary for consumer safety, it is already known

that these animal tests might have limited prediction for human neurotoxicity of some compounds. One example is provided by paraquat and rotenone. While the guideline using rodent studies for these compounds did not reveal hazards for Parkinsonism, a recent Scientific Opinion by EFSA on 'Investigating experimental toxicological properties of plant protection products, having a potential link to Parkinson's disease and childhood leukaemia' (Ntzani *et al*, 2013) disclosed epidemiological evidence for these substances as neurodegenerative hazards for humans. It is to note that regulatory studies are intended to explore for any potential hazard but they are not specifically designed to inform on specific and complex human health outcomes. On the other hand, observational studies on the effects of pesticides for the induction of neurodegenerative diseases have weaknesses in providing causal exposure-effect relationships. Thus, experimental and mechanistic data focussing on specific neurologic pathways is needed to support epidemiological human data or to assess human hazard of new substances. This need is strongly supported by observations in pharmacological research where also due to lack of translation from animals to humans new drugs, e.g. for treatment of cerebral ischemia, fail to enter the clinics (Leist & Hartung, 2013; Perel *et al*, 2007; Matthews, 2008).

New data type and methods, like *in vitro* or alternative organism (AO) *in vivo* methods and test strategies, might be more effective in hazard identification. However, any alternative method can never be a 'stand-alone' test for neurotoxicity testing as structure and function of the brain are very complex and especially intellectual output is very difficult to measure in any model other than humans. However, to even be able to successfully use cell culture data and other alternative approaches for human risk assessment, one needs a framework where such alternative testing results based on cell free assays, cellular models and non-mammalian *in vivo* studies or *in silico* approaches can be embedded. A framework like this could for example be an Adverse Outcome Pathway (AOP)-based IATA (Integrated Approach for Testing and Assessment (Tollefsen *et al*, 2014). Such an approach is currently applied for the assessment of skin sensitizers (Patlewicz *et al*, 2014).

The AOP is a framework helping to organize existing scientific knowledge on a chemically triggered initiating event leading to an adverse outcome, e.g. human disease (OECD, 2013). It can be utilised to assess either biological plausibility or to instigate actual risk assessment. As EFSA is recognizing the value of the AOP concept for protection of human health, the EFSA panel on Plant Protection Products and their Residues (PPR) organized in 2014-2016 a Working Group on 'Experimental toxicology data of pesticides and their potential link to Parkinson's disease and childhood leukaemia', where all available information on paraquat and rotenone with regards to the induction of Parkinsonism is collected within the AOP framework and re-evaluated for biological plausibility as well as its usage for risk assessment (Ockleford *et al*, 2017). One outcome of generating AOP-based IATAs is a testing strategy for human hazard assessment. As a testing result is only as good as the model used for generating such result it is of outmost importance to test key events (KE), which are hallmarks of the causal chain from initiating event to adverse outcome, identified with the AOP concept in biologically relevant models. In addition, it is favourable to employ as little experimental models as necessary because also non-rodent testing produces significant costs. Therefore, a thorough evaluation of cell-free and cell-based *in vitro* as well as *in vivo* systems using alternative organisms, supported by theoretical modelling approaches (e.g. QSAR, grouping of structural similar compounds) seems a prerequisite for building alternative, AOP-based IATAs for neurotoxicity testing. Here, one focus of the model evaluations should lie on currently known KE for acute, chronic, central and peripheral neurotoxicity of

reference compounds. As a start, KE identified in Bal-Price et al. 2015 (Bal-Price *et al*, 2015b) will be included into the search strategy and supplemented by common modes-of-action for NT.

1.2. Review Question and Objectives

The **review question** is summarized as **"Which alternative test methods or approaches are available to evaluate endpoints (KE) for neurotoxicity with or without exposure to predefined neurotoxic compounds?"**

With this question the main goal of the project is a literature search, analysis and appraisal on state of the art of alternative NT testing methods (excluding developmental neurotoxicity, DNT) currently available or under development in order to support the peer review of active substances under Reg. 1107/2009.

In particular, the **1st objective** is to perform **an exhaustive and comprehensive literature search, collection and appraisal of all relevant information** in English for the period 1990-2017 on the state of the art of respectively 1) alternative *in vivo* neurotoxicity testing models (e.g. non-mammalian animal models), 2) *in vitro* cell-free and cell based neurotoxicity test methods that allow testing of a large number of chemicals (medium-and high-throughput screening), 3) *in silico* methods, 4) read across and 5) combination of testing methods in test batteries that incorporate different neurotoxicity-relevant endpoints including high-throughput systems. Besides validated methods, also methods currently under validation and methods at a research stage will be considered. The search will include publicly available peer reviewed research publications and publicly available 'grey literature' (government reports and official institutions documents).

These results will feed into work package 2 with the **objective** to make an **overall evaluation of the suitability of selected methods or combination of methods to support and complement current international regulatory requirements for NT testing**. To accomplish this, all relevant publications will be analysed for suitable test methods based on regulatory requirements. These methods will be listed and described in detail with information on e.g. the test system (alternative organism*, in vitro, in silico*), the protocol used and the quality of the data. This information will feed into a thorough evaluation of general performance characteristics (sensitivity, specificity), the potential to predict human NT and the strengths and weaknesses of each assay or a combination of assays.

This comprehensive literature search on state of the art of NT test methodologies and critical analysis for suitability to support regulatory assessments will be made from a Plant Protection Products regulatory point of view, with the focus to the needs and possible strategies for future research and risk assessment. Specifically, an emphasis will be put onto the AOP concept. AOPs concerning NT published either in the peer-reviewed literature, e.g. Bal-Price et al. (2015), or submitted to the AOP Wiki (https://aopwiki.org/) by the start of the project, will be studied for key events (KE; including molecular initiating event, MIE, which is defined as a specific KE (Villeneuve *et al*, 2014), preferably common KE, relevant for NT modes-of-action. These identified KE will receive special attention in the data collection, as those are probable candidates for tested endpoints within an AOP-based IATA as was exampled for skin sensitization earlier (Leist

& Hartung, 2013). To ensure relevance of the KE measured in the methods included in the final recommendation, we will perform a pre-screen on mode-of-actions (MoAs) of the identified neurotoxic compounds. Only the compounds with a proposed cellular and/or molecular MoA will be pursued in the further method evaluation because for a compound with unknown MoA one cannot judge if the *in vitro* system predicts NT correctly.

The 3rd objective will provide a report summarising the strategy, results and methods evaluation of the literature review. In addition, founded on AOP-based KE and possibly additional end-points identified as KE in AOPs yet to be created, a recommendation for a possible IATA for NT testing will be made.

2. Methodology

2.1. Compound selection and mode-of-action analyses

We exclusively consider methods that are used with compounds known to be neurotoxic in humans that have a known mode-of-actions (MoA). The reason for this decision was that by this procedure we can easily identify if a method correctly predicts an endpoint relevant for the compound and thus reduce uncertainty in method evaluation. Exceptions are computational models ('in-silico-related'), studies that focus on method development/characterization ('method-related') and methods that model the blood brain barrier ('BBB-related') because these studies are not a priori compound-related.

Primary selection of the compounds was based on two publications by Grandjean & Landrigan, 2006, 2014, an EFSA supporting publication by Choi *et al*, (2016) and chapter 16-22 of the 'Handbook of Neurotoxicology' (Chang, 1995). The list of compounds within the (Grandjean & Landrigan, 2006, 2014) publications contains all environmental chemicals that have been identified as toxic to the human brain until 2013, based on information from the Hazardous Substance Data Bank (HSDB) of the US National Library of Medicine. The compounds within Choi et al. (2016) include additional chemicals that induce Parkinsonian's symptoms (Appendix A). Natural neurotoxins were identified from the 'Handbook of Neurotoxicology'. In this book all compounds with a described MoA were selected. Additionally we screened the list of all natural neurotoxins named in the 'Handbook of Neurotoxicity' and identified compounds of clinical relevance which were then included based on expert judgment (Appendix B). For these seven compounds we performed a separate MoA analysis.

In the next step, MoA analyses for neurotoxicity were performed for 248 individual compounds, 23 compound classes and 212 natural neurotoxins (Appendix A+B) assembled in the four sources described above. Therefore, the compounds were distributed among the experts and each compound was reviewed by one expert. In case of uncertainty, e.g. due to multiple MoA or insufficient information in the literature, a second expert reviewed the compound. In the MoA analyses we focused on the different levels of organization: molecular, cellular, organ and organism and searched for causality between the different levels. Such causality between some, but not all levels was then called a 'partial MoA', while complete causality across all levels was needed for a 'full MoA'. As a first information source we used the Toxnet Database 'Hazardous Substance Data Bank' (HSDB). Within HSDB we performed a search using the Chemical Ab-

stracts Service (CAS) number of each compound. Retrieved data on compounds was then screened for neurotoxicity MoA as described above. If information from HSDB was not sufficient for defining at least a partial MoA, we performed a search in PubMed combining the key word 'neurotoxicity' with search terms related to MoA (e.g. mode of action, mechanism, key event) here primarily focusing on review articles. This 'MoA search' was then combined with a 'compound search' of the respective compound. The 'MoA search' was done in title and abstract and the 'compound search' in the title of the articles (Appendix E). Depending on the amount of articles retrieved we extended the search by performing the search for each compound in title and abstract or by directly combining the 'neurotoxicity search' with the 'compound search'. The selection of articles screened for MoA analyses was then performed by expert judgement. For compound classes, we used the following approach: Databases of regulatory organisations or governmental organisations or registry dealing with toxicology were screened manually in the following order: EFSA > World Health Organization: Environmental Health Criteria (EHC), a series of monographs prepared by the International Program on Chemical Safety (IPCS) (IPCS/WHO reports, used by EFSA for example to define the cumulative assessment group-CAG) > HSDB. In case an EFSA document was not available, websites from EPA, JECFA, COT (UK) and ATSDR were searched. By this procedure, we prepared a list containing the available species of a variety of compound classes (Appendix A) and gathered information on neurotoxic MoA of compounds of each class including e.g. metabolites common to a class, a released metal or only certain species. We then developed the MoA depending on the outcome of these analyses and according to the strategy described above. For compound classes without information on single species we performed the search for the individual compounds that are listed in Appendix A and for the most common names of each compound class.

Information on MoA of a compound/compound class was collected in the 'MoA analysis sheet'. Here we collected information on the effects described for each compound and classified them based on the different levels of biological organization (molecular, cellular, organ, organism), the species and the method types (*in vitro, in vivo*) that were used to identify the effect (Appendix C shows the layout of the MoA analysis sheet). Based on the collected information each expert concluded if the data on causality between at least two levels of organization was sufficient to call it a MoA as described above. In case a MoA was considered as 'partial', the compound was labelled as 'Yes - Partial' in the MoA sheet or if it was a full MoA, the label was 'Yes' in the sheet. For a brief overview, we summarized the positive results of the MoA analyses for all single compounds and compound classes (Appendix A), as well as for all natural neurotoxins (Appendix B) naming the name of the compound/compound class, CAS-number and indicated if a (partial) MoA was identified. Compounds with no identified MoA were excluded for the next step, the search strategy (2.).

With regards to the upcoming data collection of the full texts, which is the step after searching for articles containing compounds with identified MoA, we sorted the identified compounds by MoA groups called 'MoA/endpoint categories'. In these categories, MoA are collected that share common key events (KE). These categories associated with neurotransmission (cholinergic, GABAergic, glycinergic, glutamatergic, adrenergic, serotonergic, dopaminergic, neurotransmission in general), ion channels/receptors (sodium channels, potassium channels, calcium channels, chloride channels, other receptors), and cellular endpoints (mitochondrial dysfunction/oxidative stress/apoptosis, redox cycling, altered calcium signaling, cytoskeletal alterations, neuroinflammation, axonopathies, myelin toxicity, delayed neuropathy, enzyme inhibition) and

other (Appendix D). These MoA categories will be used as 'Endpoint Categories' in the data collection sheet.

2.2. Search strategy

The search procedure consisted of four independent search strategies for searches related to compounds, *in silico* methods, new *in vitro* methods and methods for mimicking the BBB, which are described in detail below. The corresponding four search strings are listed in Appendix E.

1st search strategy ('compound-related'):

The first strategy consists of search strings that are designed to find articles that use alternative methods (cellular, cell free and alternative organisms) for assessing neurotoxicity of compounds and compound classes with an identified MoA as described under 2.1. This search combines the 'neurotoxicity search' with the 'alternative method search 1' and the 'compound search' by the Boolean operator 'AND'. The 'neurotoxicity search' is designed to identify all articles that are related to any adverse effects to the nervous system. The 'alternative method search 1' is designed to identify all articles that use alternative methods and combines one search string for *in vitro* models, one for cell free models, one for alternative organism based models and one with keywords associated to alternative methods in general (e.g. alternative approach, test method, testing battery). For the compound search all synonyms as listed in the ChemIDplus (https://chem.sis.nlm.nih.gov/chemidplus/)[1] database together with the CAS number were combined by the Boolean operator 'OR'. Synonyms with less than four letters were excluded because of the high probability of false positive search results. This first search strategy was performed for each compound/compound class and search results were distributed to the individual experts who identified the respective compounds' MoA.

For some compounds the number of retrieved articles was too high for a manual screening and most of these articles did not meet our inclusion criteria. Therefore for compounds/compound class with >500 search results we performed an additional alternative method related search ('alternative method search 2') to retrieve only those articles that focus on the use and development of alternative methods for neurotoxicity testing.

2nd search strategy ('*in silico*-related'):

The second strategy was designed to retrieve all *in silico* models/methods useful for the assessment of neurotoxicity. For this search the 'neurotoxicity search' was combined with the '*in silico* search' that consists of key words associated with *in silico* models/methods such as QSAR, *in silico* or read across by the Boolean operator 'AND'.

3rd search strategy ('method-related'):

The third search strategy was designed to retrieve all publications on well-characterised alternative models for NT testing that do NOT study the effect of our preselected compounds, but are

[1] https://chem.sis.nlm.nih.gov/chemidplus/

dedicated to basic neurotoxicity model description/development. This search combines the 'neurotoxicity search' with an adapted version of the 'alternative method search 1' and the 'alternative method search 2'. We adapted the 'alternative method search 1' by removing all general alternative methods-related keywords and by searching for '*in vitro*' instead of '*in vitro*' combined (AND) with keywords such as 'model', 'assay' or 'test' as in the 'Alternative method search 1' used for the 1st search strategy. The 'alternative method search 2' consists of keywords related to model development and characterization such as 'assay validation', 'assay development', 'test assay', or 'screening method'.

4th search strategy ('BBB-related'):

The fourth search strategy was designed to retrieve publications that focus on the use or development of BBB models. This search combines a 'BBB search' with the 'alternative method search 1 for BBB search' and the 'alternative method search 2 for BBB search'. The 'BBB search' was designed to retrieve all studies with a focus on BBB or BBB models. The 'alternative method search 1 for BBB search' combines keywords from the search strings for *in vitro* models, cell free models, alternative organism based models and the *in silico* search. The 'alternative method search 2 for BBB search' adds keywords like 'fabrication', 'microtechnology', or 'engineering' to the 'alternative method search 2'.

The information collected covered a time span from the 1st of January 1990 to 2017 (between April and July, depending on when the searches were performed). The searches were performed in the two major bibliographic databases, Web of Science® (WoS; http://apps.webofknowledge.com) and PubMed (http://www.ncbi.nlm.nih.gov/pubmed). The searches in PubMed were performed in 'Title/Abstract' which includes a search in title, abstract and author keywords of the article. The search in WoS was performed in the search field 'Topic' which includes a search in title, abstract, author keywords and keywords plus® of the article. The only exception were the keywords 'BBB', 'brain barrier' and 'brain blood barrier' of the 'BBB search' which were only searched for in the title in both databases. Publications from the search in WoS were limited to English language articles and the document types 'review' and 'meeting abstract' were excluded.

For the 1st search strategy we performed one search for each compound/compound class. For the other search strategies it was one search for each strategy. After each search all duplicate articles (based on exact match of title and year were deleted from the search. All articles were exported from EndNote, imported into a shared Mendely database and distributed to the experts. Because Mendeley has an automatic identification of duplicate articles, additional duplicates that were not recognized by Endnote were automatically deleted after import to Mendeley. Here, it was not possible to reconstruct which articles were deleted. Therefore there are slight discrepancies between the amount of articles in Endnote and the amount of articles in Mendeley. Appendix A gives detailed information on the number of articles retrieved after each search.

2.2.1. 'Grey' literature search

Data mining for 'grey literature' was performed by a different strategy than scientifically published information involving a peer-review process.

We searched all websites listed in Appendix F in the 'Search' function of the respective websites with the keywords 'neurotoxicity' or, if a search function was not available or only provided a google or PubMed search, scanned the website for information on alternative neurotoxicity methods. All information on promising methods from the Website search were collected and sorted with regard to the specific topics: *in vitro*, *in vivo* alternative organisms, *in silico* and *biological barrier model* (Appendix F).

Additionally, personal contacts were utilized to contact organizations/consortiums currently involved in neurotoxicity method development. Therefore, we designed an email that contained the question of unpublished methods for neurotoxicity testing, which was sent to the personal contacts by each of the partners. Specifically,

- VITO contacted: University of Amsterdam (J. Legradi), EPA (Stefanie Padilla)
- IUF contacted: Coordinator of ESNATS (Jürgen Hescheler), Coordinator of EU-ToxRisk (Bob van de Water), EURL-ECVAM (Anna Price), EPA (Kevin Crofton), ZEBET (Gilbert Schönfelder), HESI (Michelle Embry), CAAT-US (Helena Hogberg), CAAT-Europe (Marcel Leist)
- UMIL contacted: ALTTOX (community blog), Swiss Centre for applied human toxicology (Florianne Tschudi-Monnet), Department of Environmental and Occupational Health Sciences, University of Washington (Lucio Costa/Gennaro Giordano), Institute for Risk Assessment Sciences, Utrecht University (Remco Westerink), Department of Neuroscience, Karolinska Institute (Sandra Ceccatelli), Istituto Superiore di Sanità (Luisa Minghetti)

Emails were sent to the organizations/consortiums, all information on promising methods was sorted with regard to the specific topics: *in vitro*, *in vivo* alternative organisms, *in silico, blood brain barrier model.*

The language restrictions for 'grey' literature are in concordance with the languages spoken by the consortium members (English, German, Dutch, Spanish and Italian).

Grey literature search for *in silico*:
The result of the first screening of websites offering *in silico* models or tools is presented in Appendix F. This selection included publicly available and commercial websites. These websites were further explored and if needed, contacted via the online contact point to obtain more information.
During the full paper text screening, additional *in silico* models came up, as some of the published algorithms have been developed to a web-based or a stand-alone QSAR model.

2.3. Selection

2.3.1. Selection based on title and abstract

The manual study selection process was performed with the Mendeley reference managing software using a Mendeley Institutional Edition account as this software made it possible for all experts to share the same database.

The selection process based on title and abstract was an unmasked assessment in which each expert screened titles and abstracts of a set of pre-selected articles and made a decision on inclusion or exclusion according to the pre-defined selection criteria (see list below). If there was an uncertainty in the decision, the article was tagged as 'un' (uncertain) and a second expert reviewed the same article. Both experts discussed and agreed on a final decision. For documentation of this selection, the ID of the article (PMID, DOI, or ISBN) and the first author of each publication were transferred to an Excel sheet. In this sheet, 'yes' for included or 'no' for excluded was documented together with the reason for exclusion (Appendix G).

The decision process in the study selection for title abstract screening was based on the following exclusion criteria:

For all articles from 1st search strategy: (compound-related)

1. Study does not deal with neurotoxicity
2. Secondary literature (review, meeting abstract, etc.)
3. Duplicate
4. No compound with known MoA tested
5. DNT study
6. Wrong species (not human, mouse, rat, chicken, *C. elegans*, sea urchin, zebra fish, xenopus, *drosophila*)
7. Test method not in agreement with MoA
8. Test method not able to measure NT endpoint
9. Only mixture tested
10. *In vivo* study
11. Not possible to study this endpoint with brain cells *in vitro*
12. Article language is not English
13. Retracted publication

For the exclusion criteria 4, 7 and 9 we had the following exceptions:

Articles that focus on the development and characterization of promising alternative models for neurotoxicity testing or biological barriers that would be excluded based on these criteria could be included if the expert judges these methods as valuable for the projects objectives. These articles are tagged by 'om' (only method) or 'bb' (biological barrier model) in the Mendeley library.

Articles that focus on *in silico* models should not be excluded based on exclusion criteria 4, 7 or 9. These articles are tagged by 'comp' (computational model) in the Mendeley library.

For all articles from 2nd search strategy (in-silico-related)

1. Study does not deal with neurotoxicity
2. Secondary literature (review, meeting abstract, etc.)
3. Duplicate
4. No compound with known MoA tested (not for computational models)
5. DNT study
6. Wrong species (not human, mouse, rat, chicken, *C. elegans*, sea urchin, zebra fish, xenopus, *drosophila*; not for computational models)
7. Test method not in agreement with MoA (not for computational models)
8. Test method not able to measure NT endpoint
9. Only mixture tested, or other test items than chemical compounds (not for computational models)
10. *In vivo* study (not for computational models)
11. Not possible to study this endpoint with brain cells *in vitro* (not for computational models)
12. Articles language is not English
13. Artificial neural network (unless for neurotoxicity)
14. General considerations on AOP for NT
15. General considerations on computational tools

For the exclusion criteria 4, 7 and 9 exist the same exceptions as mentioned above.

For all articles from 3rd search strategy (method-related)

1. Study does not deal with nervous system
2. Secondary literature (review, meeting abstract, etc.)
3. Duplicate
4. DNT study
5. Wrong species (not human, mouse, rat, chicken, *C. elegans*, sea urchin, zebra fish, xenopus, *drosophila*)
6. Test method not able to measure NT endpoint
7. *In vivo* study
8. Not possible to study this endpoint with brain cells *in vitro*
9. Articles language is not English
10. Study does not focus on development/characterization of a valuable alternative test method/model
11. No method for NT testing

For all articles from 4th search strategy (BBB-related):

1. Secondary literature (review, meeting abstract, etc.)
2. Duplicate

3. Wrong species (not human, mouse, rat, chicken, *C. elegans*, sea urchin, zebra fish, xenopus, *drosophila*)
4. *In vivo* study
5. Articles language is not English
6. Study does not focus on development/characterization or use of alternative test method/model
7. No valuable alternative method

Application of all these criteria led to a total of 1803 articles out of 9066 studies that qualified for full text screening.

For the full-text screening, full-text copies of selected references were obtained by a manual search. Therefore a search for the PMID, DOI or the whole title of the articles was performed in PubMed, WoS, or google scholar by using the 'search for PDF'-function or in the local library. In case the articles were not retrieved we contacted the corresponding author of the study by email. In case of non-resonance, the study was not evaluated.

Each PDF with the full text was attached to the respective Mendeley reference.

2.3.2. Selection based on full text

The selection process based on full text was an unmasked assessment in which each expert screened the full text of a set of pre-selected articles and made a decision on inclusion or exclusion according to the pre-defined selection criteria (see list below). If there was an uncertainty in the decision, the article was tagged as 'un' (uncertain) and a second expert reviewed the same article. Both experts discussed and agreed on a final decision. For documentation of this selection, the ID (PMID, DOI or ISBN) and the first author of each publication were transferred to an Excel sheet (the Data collection sheet). In this sheet, 'yes' for included or 'no' for excluded was documented together with the reason for exclusion (Appendix H).

The decision process in the study selection for full text screening was based on all selection criteria for the title abstract screening and the following additional full text exclusion criteria:

<u>For all articles from 1st search strategy: (compound-related)</u>

1. The study does not give sufficient experimental detail to assess the described methods
2. Study on neuroprotection
3. Only single dose
4. Manipulated test system
5. No quantification of endpoints
6. No full text
7. Tumor cells
8. Life stage = not alternative (e.g. zebrafish beyond 5 days post-fertilisation)
9. Not whole organism method (e.g. expression of sodium channels in Xenopus oocytes→ *in vitro* method)

Dose-response increases the likelihood of a causal association, for this reason evaluations based on single dose only were not taken into account, with the exception of those studies considered particularly relevant by the expert judgement. In this case, comments to justify have been added to the DCS. In general, studies based on at least 3 doses were considered, not to be too restrictive.

Studies on neuroprotection were excluded when based on single dose only of the tested compound.

Tumor cells have genetic differences from neuronal cells and physiologically may diverge from normal cells in various respects. The analysis therefore mainly focused on:
- stem/progenitor cells, which also allows for the possibility of generating human CNS neurons with normal properties
- primary cells, which accurately represent mature neurons
- immortalized cells

Manipulated test systems were excluded since they physiologically may diverge from normal cells in various respects.

Quantification of the result allows comparison to other experimental group and statistical analysis necessary to support the plausibility of the study design and data.

For all articles from 2nd search strategy (in-silico-related)

No additional full-text criteria for *in silico*.

For all articles from 3rd search strategy (method-related)

1. The study does not give sufficient experimental detail to assess the described methods
2. Study on neuroprotection
3. Manipulated test system
4. No full text
5. Tumor cells
6. No particularly new/relevant test method

For all articles from 4th search strategy (BBB-related):

1. The study does not give sufficient experimental detail to assess the described methods
2. Manipulated test system
3. No full text
4. study does not focus on development/characterization of BBB alternative model
5. published before 2016
6. No innovative BBB model

For studies that focus on the use or development of BBB models we identified four reviews that give an overview of established *in vitro* BBB models up until 2016 (Banerjee *et al*, 2016a; Helms *et al*, 2016a; Palmer *et al*, 2013; Wolff *et al*, 2015b). Starting from these reviews the full text screening was only performed for studies that were published in 2016 and 2017. From these we selected only those studies that use or develop innovative BBB models.

For all studies that were included based on full text selection we performed an assessment of the methodological quality (see below).

2.3.3. Selection of 'grey' literature

In general the consortium agreed that 'work in progress', e.g. granted projects on method validation, ring trials, etc. are important for EFSA to recognize. However, methods, which are developed and presented in various media, but have not yet undergone scientific peer review e.g. through a standard scientific publication process should not be taken forward in the process on the same grounds as publications from the peer-reviewed literature since such data can be preliminary and open to final adjustments. Due to this fact, the partners used additional criteria for the selection of relevant unpublished methods.

Selection criteria:

- Information on granted projects currently being performed
- Information containing preliminary results of on-going projects not yet finished

2.4. Assessment of methodological quality

The assessment of methodological quality was performed according to the description in the project outline. In short, collected all necessary information on:

I: Test substance identification
II: Test system/organism characterisation
III: Study design description
IV: Study results and data analysis documentation
V: Plausibility of study design and results

using the ToxRTool. This publicly available tool to assess the reliability of toxicological data is based on the approach of Klimisch *et al*, (1997)[2] and will assign a quality score to each study by the use of a QA sheet for each publication selected based on the criteria for full text evaluation. According to this score, the study is assigned to the following categories:

1. Reliable without restriction (15-18 for *in vitro*; 18-21 for *in vivo* alternative organisms)
2. Reliable with restrictions (11-14 for *in vitro*; 13-17 for *in vivo* alternative organisms)

3. Not reliable (<11 for *in vitro*; <13 for *in vivo*; or not all red criteria met)

The ToxRTool and the instructions were distributed among all experts. Each expert saved the individual QA sheet (Excel-format) for each study including the title of the study, authors and year, the score given in each criterion, the final category assigned and possible comments for documentation. Studies that obtained a score <11 (*in vitro*) or < 13 (*in vivo*), or zero points for one of the following criteria:

1. Was the test substance identified?
2. Is the test system described?
3. Are doses administered or concentrations in application media given?
4. Are frequency and duration of exposure as well as time-points of observations explained?
5. Were negative controls included (give also point, if not necessary, see explanations)?
6. Were positive controls included (give also point, if not necessary, see explanations)?
7. Is the study design chosen appropriate for obtaining the substance-specific data aimed at (see explanations for details)?

were excluded from the data collection process and method evaluation, unless the expert judges the study as valuable. In this case there will be an argumentation why the study should not be excluded in the comment section of the data collection sheet. The result of the category assigned as well as the total points received per study were documented in the data collection sheet (Appendix D).

Information from all included studies were collected in the different data collection sheet for either *in vitro* methods, for alternative organisms, or for *in silico* approaches. The number of all articles retrieved after each search and selection process is summarized in a flow chart in Figure 1.

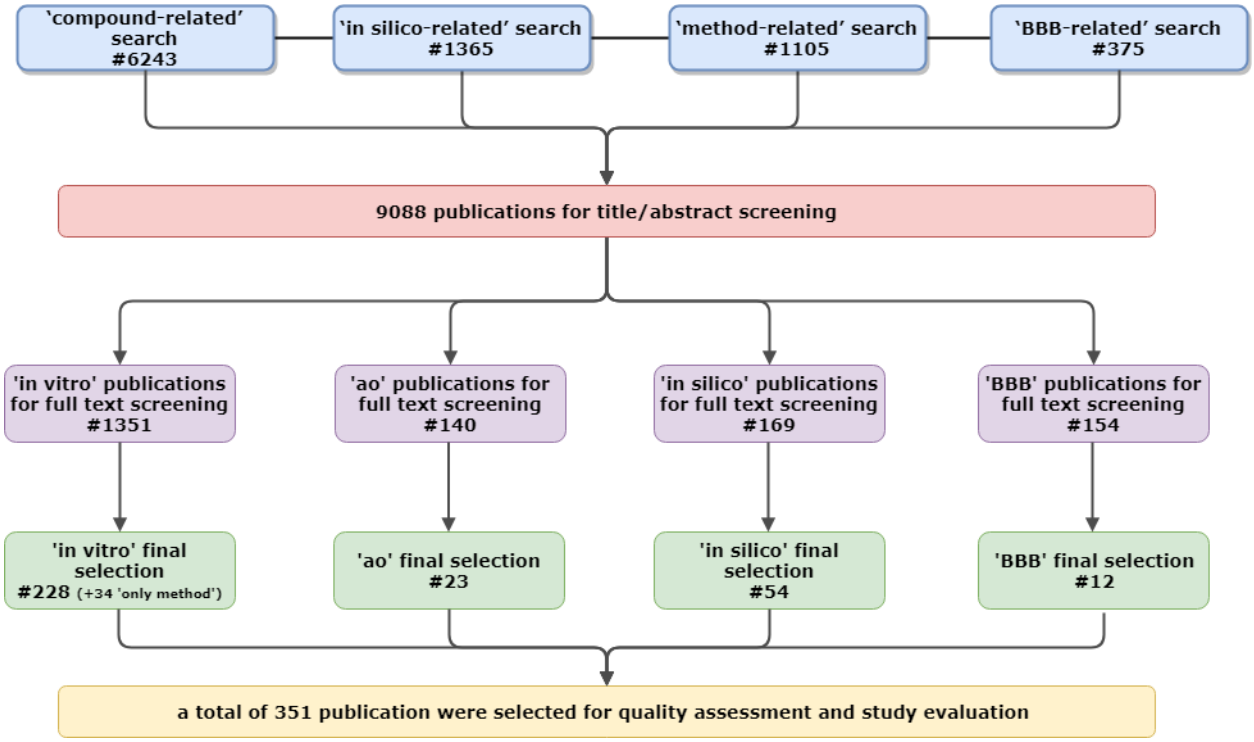


Figure 1 Graphical summary of publication selection.

2.5. Data Collection

The data collection sheet is designed to summarize all information from the selected scientific papers that are relevant for an evaluation and comparison of the different test systems with regard to the endpoints they are able to assess. For the 'compound-related' studies we only collected information on effects that are in accordance with the described MoA. In case of negative compounds all endpoint effects were collected.

Information that was collected from 'compound-related' scientific papers:

Publication ID, first author, journal and year.

- To identify each publication.

Selection based on full text (decision 1st reviewer, Argumentation for exclusion)

Decision 1st reviewer

- Decision based full text screening and quality assessment (yes/no).

Method type.

- To distinguish the different method types, *in vitro*, cell free, organelle, alternative organisms, *in silico* or a combination of these. In case one publication presents data on more than one method type the information is collected in the respective data collection sheet.

Test system classification/characterisation (test system, species, strain, age of cell source or age of alternative organism, cell type, brain region, culture age, serum use)

- To classify the test system that is used in each publication. Each test system will be entered in a separate row.

- Serum use can alter chemical properties *in vitro* and thus may change the biological response to a chemical. To identify the reason for different effects of test systems after chemical exposure it can be crucial to know if the system was treated in the presence of serum (yes/no).

Endpoint assessment.

Endpoint category.

- Endpoint categories were pre-defined by MoA analyses as outlined in Appendix D.

Endpoint.

- The biological or chemical process, response or effect assessed by a test method that is grouped into an endpoint category.

Multiple endpoints.

- To see if the test method was used to analyse multiple endpoints (yes/no). In case multiple endpoints were studied, each endpoint with each compound has to be entered in a separate row.

Analytical/test method.

- The process or procedure used to obtain information on the endpoint. The analytical method describes the method that is used to do the endpoint measurement.

Endpoint-specific controls.

- Control that selectively and reproducibly modulates the endpoint by a known MoA (yes/no).

Compound testing information (compound name, CAS number, use of multiple test compounds, compound classification a priori, concentration range, exposure duration).

- Each compound with each endpoint in one test system will be entered in a separate row.

Multiple test compounds.

- Is more than one compound used in the study (yes/no)? If yes, for each compound a new row has to be filled, yet only for those compounds that have a defined MoA.

Compound classification a priori

- Negative or positive compound?

Information on compound effect.

Effect analysis

- Describes how an effect on this endpoint is analysed (e.g. EC_{50}, concentration-response, induction/reduction).

Hazard.
- Indicates if the endpoint is affected (yes/no).

Classification.
- Gives a classification of the compound according to the prediction for this compound in the test method (true positive, true negative, false positive, false negative).

Effect concentration.
- Gives the concentration range in which the endpoint is affected.

Data analysis.
- Describes how the data was analysed (quantitative, qualitative, statistical).

n =
- Gives the number of biological replicates (one 'n' means one independent experiment).

Additional information.

Throughput.
- Indicates if a test system with this specific endpoint is adaptable for high/medium throughput (yes/no).

Reliability score.
- Gives the rating based on the ToxRtool for reliability, calculated after criteria evaluation using the QA sheet.

Reliability category.
- Gives the Klimish Category (1-4) based on the ToxRtool rating.

Comments.

For all 'method-related' and 'BBB-related' studies the data collection sheet was adapted to collect information with the focus on method type, test system classification/characterisation and endpoint assessment (Appendix H).

For the evaluation of 'in-silico-related' studies, a section on general information is compiled, followed by the quality criteria that are based on the five OECD criteria for QSAR validation (OECD report 69 on QSAR validation for regulatory use):
0. General information
1. Defining the endpoint
2. Defining the algorithm
3. Defining the applicability domain (AD)
4. Internal and external validation
5. Providing a mechanistic interpretation

3. Results and Discussion

3.1. In vitro

Data retrieved from 228 publications concerning neurotoxicity of chemicals with known MoA during the years 1990-2017 were collected in a data collection sheet and analyzed in multiple ways. Because most publications contain several endpoint evaluations measured with different methods, different endpoints, different test systems or different compounds, it was not effective for the evaluation to work with 'number of publications'. Therefore, we evaluated the numbers of times a test system was used for each endpoint evaluation and/or chemical during the rest of this report. E.g., when 5 compounds are evaluated in the same publication for two endpoints and one test system, this will result in 10 citations for this one test system. In the end, each of these citations is then evaluated with respect to the assay performance for each compound and each endpoint. Thus, the number of citations for each test system or cell type category is higher than the amount of total publications. For each of this test system/endpoint measure/compound evaluation one line in the 'Data Collection Sheet' was filled. This resulted in 977 citations within 228 publications, while 1123 publications were excluded according to the defined criteria.

First, we analyzed the total number of citations over time (Figure 2a). It is obvious, that citation numbers fluctuate from year to year irrespective of species or cell type. Next, the usage of cells from different species over time was studied. The graph (Figure 2b) shows that rat cells are the most cells used over time, followed by mouse and human cells and only few publications were included in the data analyses from chicken and Xenopus. Rat cells have been used since the 1990ies, and since 1993 mouse citations became more frequent. Usage of human cells in our selection of articles started in 2002, but really accelerated from the year 2013. Rat and mouse cells show fluctuations throughout the whole period. Publications with cells from chicken and Xenopus were sporadic with no new chicken publication since 2011. The Xenopus model was still used lately. With regards to the cell type (Figure 2c), we found citations for primary cells starting from the beginning of the analysis period in 1990. Tumor cells were found to be used starting from the early 1990ies. The latest development is the usage of stem-/progenitor cells for neurotoxicity evaluation in the early 2000s.

For the general strategy, publications containing *in vitro* studies for assessment of NT (n = 228) were grouped according to **species** (human, rat, mouse, chicken and Xenopus), **cell type** (stem/progenitor cells, primary cells, tumor/immortalized cells) and individual test systems (Appendix I1). Brain regions that primary cells were derived from are specified. These data are then evaluated with regards to their performances for certain MoA represented by endpoint categories as described in the methods section. A comprehensive list of the MoA analyses can be found in Appendix D. Questions addressed within the data evaluation are: 1. *In vitro* methods from which species addresses which endpoint categories? 2. What is the predictive capacity of each cell type and individual test system for each endpoint category? Which MoA/endpoint category can be correctly assessed by which test system?

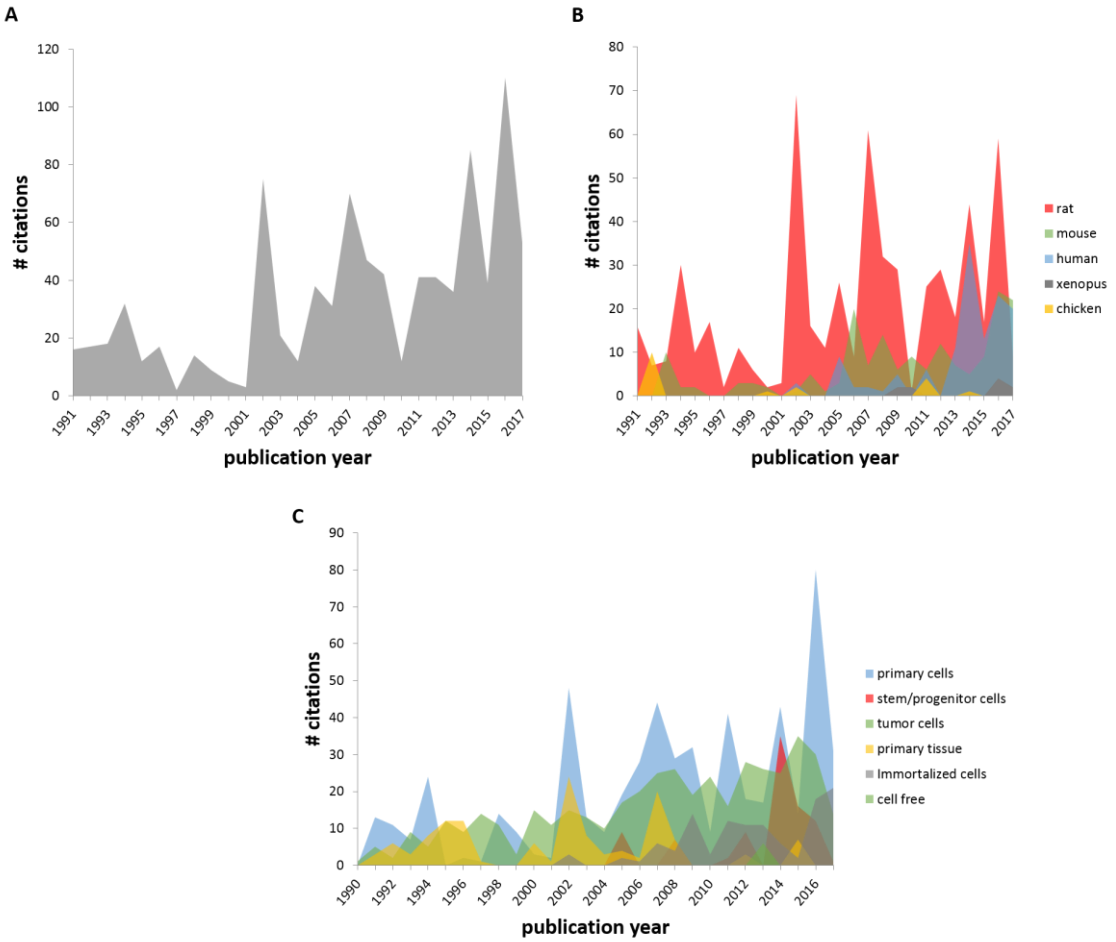


Figure 2 Number of NT citations between the years 1990 and 2017 in total (a) and subdivided for species (b) and for cell types (c) amongst the selected publications.

After grouping according to **species**, we counted 131 citations with human, 644 with rat, 174 with mouse, 18 with chicken and 10 with Xenopus cells (Figure 3). Of these, 92, 633 and 118 belonged to the category of stem/progenitor cells, primary cells and immortalized cells, respectively, for all species (Figure 5).

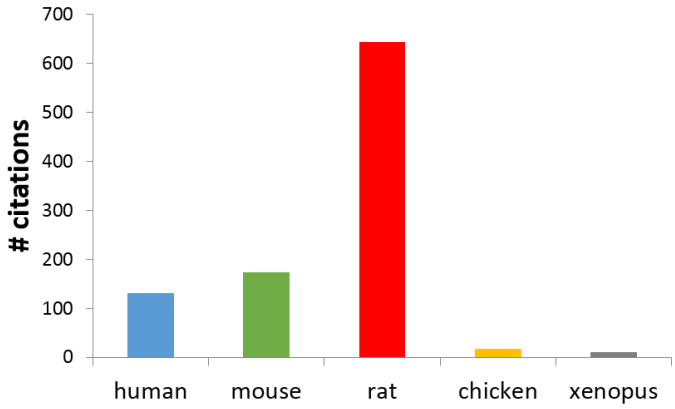


Figure 3 Within 228 original NT (selected) in vitro publications published since 1991, a total of 977 citations were found for the species human, rat, mouse, chicken and Xenopus. Thereby, some publications contain multiple species and are thus counted multiple times.

Species-specific analyses of the identified **cell types** revealed for human cells 66 (50%), 11 (8.5%) and 49 (37.7%) citations for stem/progenitor cells, primary cells and immortalized cells, respectively; for mouse cells 19 (10.9%), 117 (67,2%), and 27 (15.5%) citations for stem/progenitor cells, primary cells and immortalized cells, respectively; for rat cells 7 (1.1%), 479 (74.4%), and 42 (6.5%) citations for stem/progenitor cells, primary cells and immortalized cells, respectively; for chicken cells 16 (88.9%) and 2 (11.1%) citations for primary cells and primary tissue, respectively; and 10 (100%) citations for Xenopus primary cells (Figure 4). While due to obvious availability reasons human primary cells are hardly ever employed in *in vitro* NT studies (11; most of them NSC methods, one primary astrocyte method), most data is produced with rat primary cells (479) followed by mouse primary cells (117), human stem/progenitor (66) and tumor/immortalized cells (49), rat and mouse tumor/immortalized cells (42 and 27) and mouse stem/progenitor cells (19; Figure 4). The overrepresentation of primary rodent cells is due to historical reasons. However, within the recent years also the number of human stem/progenitor cell-based methods is clearly on the rise (Figure 4a).

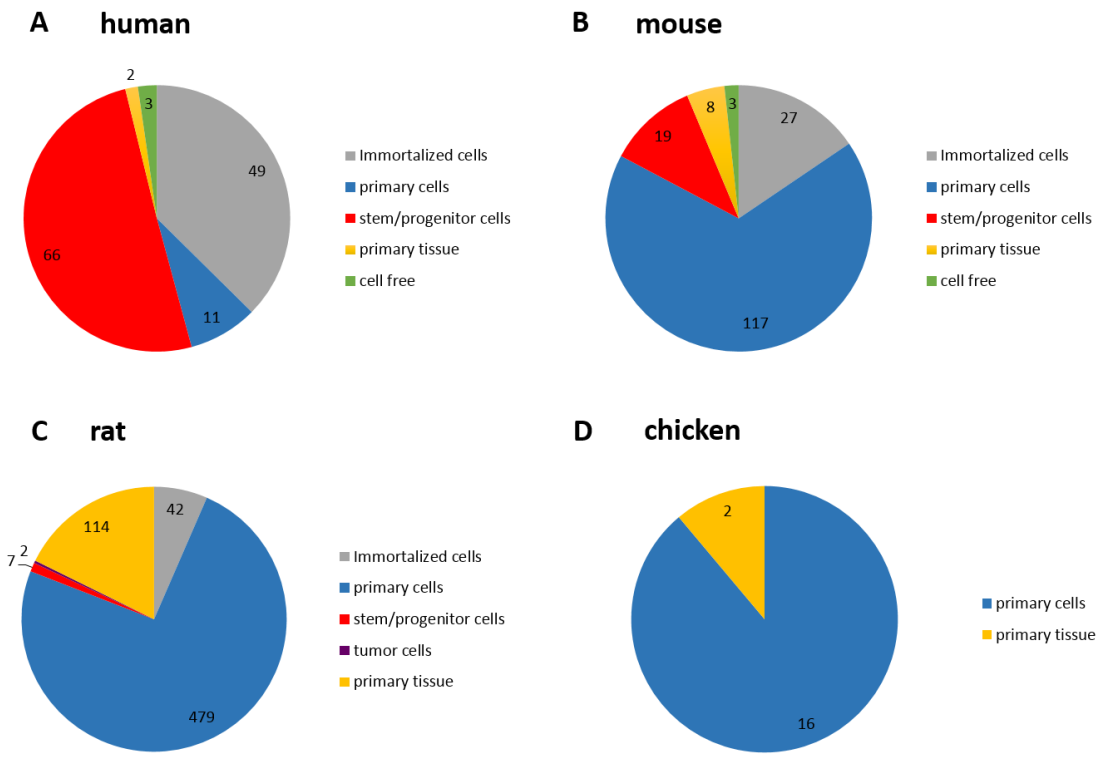


Figure 4 Within 228 original NT (selected) in vitro publications published since 1990, a total of 977 citations were found for the species human, rat, mouse and chicken. These are analyzed for usage of stem/progenitor cells, primary cells and immortalized cells in each species. Given are numbers of citations of each cell type.

Analyzing the **species** distribution within the different **cell types**, it is obvious that most data is generated with primary cells of rat 479 (75.7%) > mouse 117 (18.5%) > chicken 16 (2.5%) > human 11 (1.7%) > xenopus 10 (1.6%) origin. For the cell type stem-/progenitor cells, the ranking is different. Here, mostly human with 66 (71.7%) citations followed by mouse 19 (20.7%) and few rat 7 (7.6%) citations were counted. Similarly, immortalized cells were of human 49 (41.5%) > rat 42 (35.6%) < mouse 27 (22.9) origin. Also primary tissues from rat 114 (90.5%) > mouse 8 (6.3%) and a couple from human 2 (1.6%) and chicken 2 (1.6%) were utilized (Figure 4). Also the cell free methods were sparse; human and mouse studies 3 citations each (Figure 4). On the total scale, of the 977 citations, 633 (64.8%) citations were from primary cells, here 479 (49% from total) from rats. This illustrates impressively, that of the studies selected, we have the greatest part of information generated with primary rat cells.

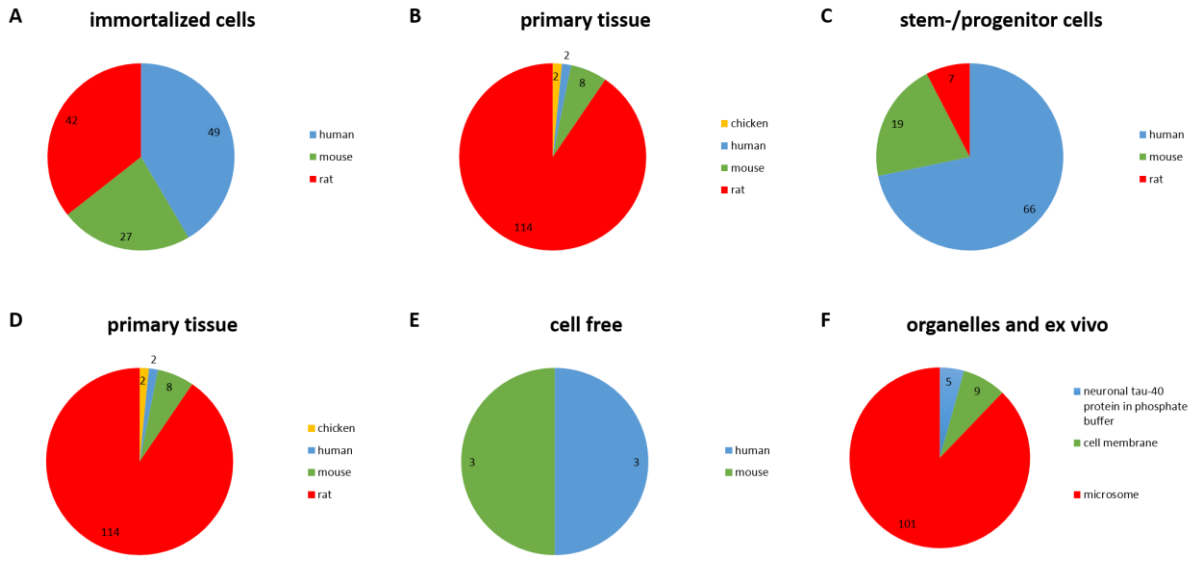


Figure 5 NT *in vitro* studies (228) employing stem/progenitor, primary, immortalized and tumor cells from human, rat, mouse, chicken and Xenopus. Given are the distributions of species within each cell type category with the numbers of citations.

In addition to the cell-based methods, there are a number of studies using cell free, organelle-based systems. Species distribution of organelles is shown in the pie graph (Figure 5). Here the rat dominance is obvious. For understanding the test systems behind, we provide Table 1, where all the different systems for the individual species are listed. While for human, recombinant proteins are the only test systems used, rodents employed cell membrane (28 citations), mitochondria (24 citations), microsomes, synaptosomes (rat) and also recombinant protein (mouse). In the rat, the highest number of citations was recorded for synaptosomes (41 citations). Studies with organelles from the other species were negligible.

Table 1: Overview over the types of organelles used from different species

	# citations
Human	**5**
neuronal tau-40 protein in phosphate buffer	2
recombinant AChE	3
Mouse	**9**
cell membrane	2
mitochondria	3
recombinant AChE	3
synaptosomes	1
Rat	**101**
cell membrane	26
microsome	8
mitochondria	18

mitochondrial suspension from liver	3
synaptosomal mitochondria	5
synaptosomes	41
Total	**113**

When looking at the tissue origin of the primary cells, they were derived from different brain regions. The regional distribution of cells across species is shown in Figure 6. In rat, across the 479 citations for primary cells most cultures were generated from cerebral cortex (241 citations) followed by hippocampus (50 citations), mesencephalon (55 citations), cerebellum (40 citations), whole brain (46 citations) and striatum (29 citations). In mouse, across the 117 citations, cerebral cortex (44 citations) and mesencephalon (36 citations) were the most widely used. In humans, only 2 citations use primary cells generated from whole brain. In chicken, forebrain was the most frequently used brain area (10 citations) followed by cerebellum (5 citations). In Xenopus, no brain region was used because isolated receptors were expressed in Xenopus oocytes.

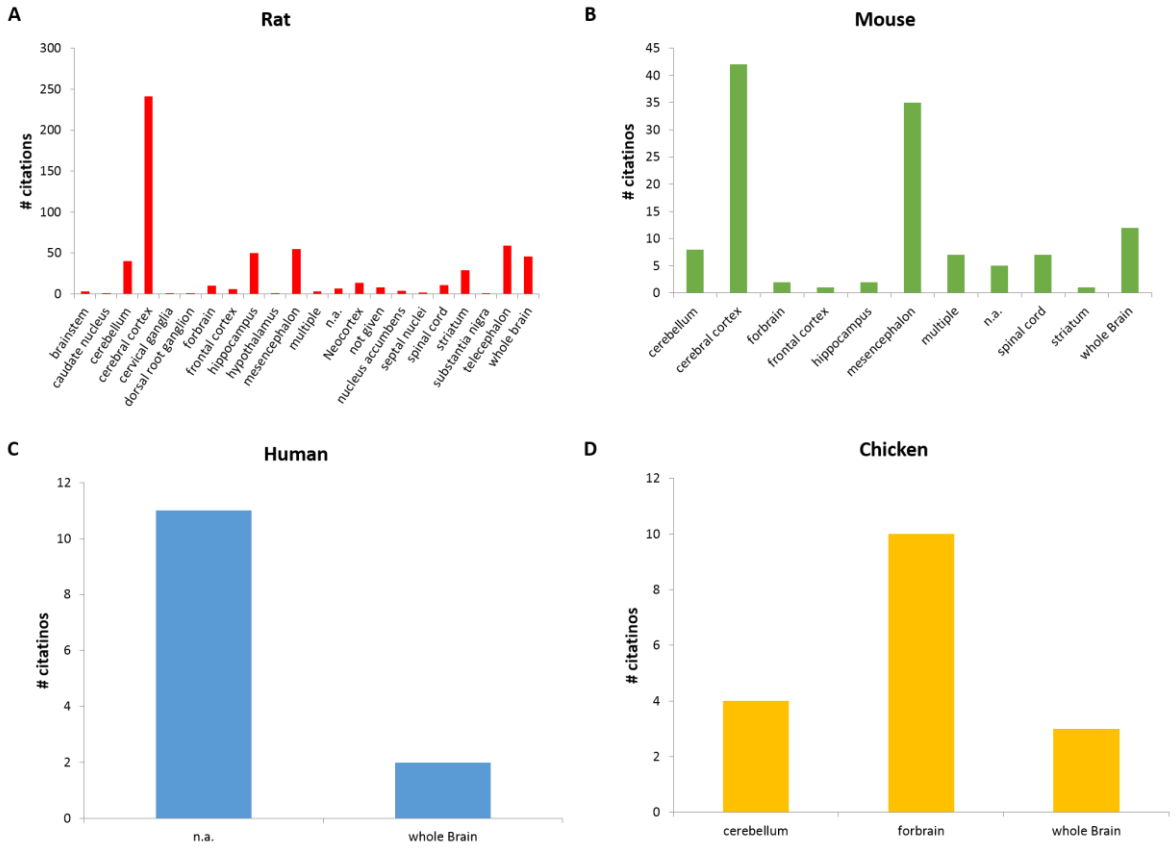


Figure 6 Brain region distribution of primary neural cells and primary tissues (see Figure 5) within the different species.

Next, we evaluated which **test systems** underlie the different **cell types**. Table 2 summarizes the Test Systems found across the different species. Amongst the rat primary cells, the by far most frequently used cell type within this investigation, primary neurons are the most frequent-

ly used test system (194 citations, 40.5% of rat primary cells). Also for mouse test systems, primary neurons were the ones with the highest number of citations (62, 53% of all mouse primary cells). For human cultures, hiPSC-derived neurons were highest cited (49 citations, 75.4% of all human stem-/progenitor cells). Although for chicken there were only 18 citations in total, also here the primary neuronal cultures dominate (10 citations, 62.5% of all primary chicken cells). Glia cells were not cited that often, here for human 2 (4.1%) primary glia cultures, for mouse 3 (2.6%) primary glia cultures and for rat 55 (11.5%) primary glia cultures were found. Only from the rat microglia cultures were identified (7 citations, 1.5% of all rat primary cells). Isolated glia is rarely published, however, mixed neuronal/glia cultures can be found more frequently. This makes sense because the interplay of neurons and glia often determines toxicity, thus a co-culture model of multiple cell types seems useful. Here, 8 citations (12.1% of stem-/progenitor cells) with hiPSC-derived mixed cultures, 15 (12.8% of mouse primary cells) mouse mixed neuron and glia cultures and 84 (17.5% of rat primary cells) of those cultures derived from rat were identified (Table 4). Although there are a large variety of test systems in this data set, it is obvious that for most test systems publications are few, and only for some publications are sufficient for deeper data analyses.

Table 2: Test systems from each cell type utilized for NT *in vitro* publications. The # citations indicates the frequency of model citation, including distinct endpoint (categories) and chemicals.

	# citations
Human	**130**
Immortalized cells	**49**
CHME-5 (microglia)	3
LUHMES	32
mesencephalic cells (MESC2.10)	9
ReNcell CX cells	5
primary cells	**11**
NSC derived culture	8
primary glia	2
sigmoid colon tissue	1
stem/progenitor cells	**65**
iPSC derived mixed culture (neurons+glia)	8
iPSC derived neurons	49
NPCs	1
undifferentiated neurospheres	8
primary tissue	**2**
neuronal tau-40 protein in phosphate buffer	2
cell free	**3**
recombinant AChE	3
Mouse	**173**
Immortalized cells	**27**
2.3D (neuroepithelial cells differentiated into astrocytes and neurons)	1

BV-2	15
CRL-2534, astrocyte type III	2
GT1-7 cells (hypothalamic cell line)	2
HT-22	5
N9 microglia	1
SN4741	1
primary cells	**116**
brain slices	1
cerebellar granule cell	3
cerebellar granule neurons	3
dorsal root ganglia/spinal cord cultures	2
isolated mouse hemidiaphragm muscles	1
mixed culture (dopaminergic neurons+astrocytes)	22
mixed culture (neurons+microglia)	1
mixed culture (spinal cord - skeletal muscle)	1
murine brain microvascular endothelial cells	1
neuron/astrocyte contact co-culture	1
primary glia	3
primary neurons	62
mixed neuron and glia cultures	15
Dorsal root ganglia	1
stem/progenitor cells	**19**
ESC	3
ESC derived glutamatergic neurons	9
ESC derived neurons	7
primary tissue	**8**
brain homogenate	2
cell membrane	2
Mitochondria	3
Synaptosomes	1
cell free	**3**
recombinant AChE	3
Rat	**567**
Immortalized cells	**38**
E18 neuroblast	2
HAPI	8
N27	27
RBE4	5
primary cells	**408**
astrocyte rich culture	11
brain slices	33
cell membrane	4
cerebellar granule cell	13
cerebellar granule neurons	9
dopaminergic neurons	1

hippocampal CA1 pyramidal neurons	1
oligodendrocyte progenitors	1
primary glia	55
primary microglia	7
primary neurons	194
primary oligodendrocytes	1
purkinje neurons	2
re-aggregating brain cell cultures	59
trigeminal ganglion neurons	1
mixed neuron and glia cultures	84
Dorsal root ganglia	3
stem/progenitor cells	**7**
differentiated mesencephalic NPCs	4
differentiated striatal neural NPCs	3
tumor cells	**2**
PC6-3	2
primary tissue	**112**
brain homogenate	17
cell membrane	22
Microsome	8
Mitochondria	18
mitochondrial suspension from liver	3
synaptosomal mitochondria	5
Synaptosomes	41
Chicken	**18**
primary cells	**16**
cerebellar bergmann glia	5
primary neurons	10
Spheroids	1
primary tissue	**2**
brain homogenate	2
Xenopus	**10**
primary cells	**10**
neurolemma in Xenopus Oocytes	6
sodium channels in Xenopus Oocytes	4
Total	**898**

In addition, cell free systems and organelles were used in our study collection. Here, the rat is the predominant species with 114 citations over 11 (mouse) and 5 (human; Table 2). In total, we recorded 132 citations with such systems.

Before the data collection, we assessed the **MoA** for 258 individual neurotoxic compounds and 23 compound classes (Appendix A, B and D). These MoA analyses were performed rather stringently, meaning that if a relation between levels of organization was not causally linked, the potential MoA was not included. This procedure might have resulted in elimination of valid studies but reduced the probability of inclusion of false-positive studies. According to these MoA, we defined **endpoint categories** where endpoints assessed in the studies were grouped into. The list of endpoints grouped into endpoint categories can be found in Appendix D. One endpoint category was 'negative'. This one was not retrieved from the MoA analyses, but resembles the negative NT compounds that we identified during the screening process. Because performance of test systems concerning negative compounds is crucial, we inserted them into the endpoint categories. In addition, we grouped studied endpoint categories by species (Figure 7).

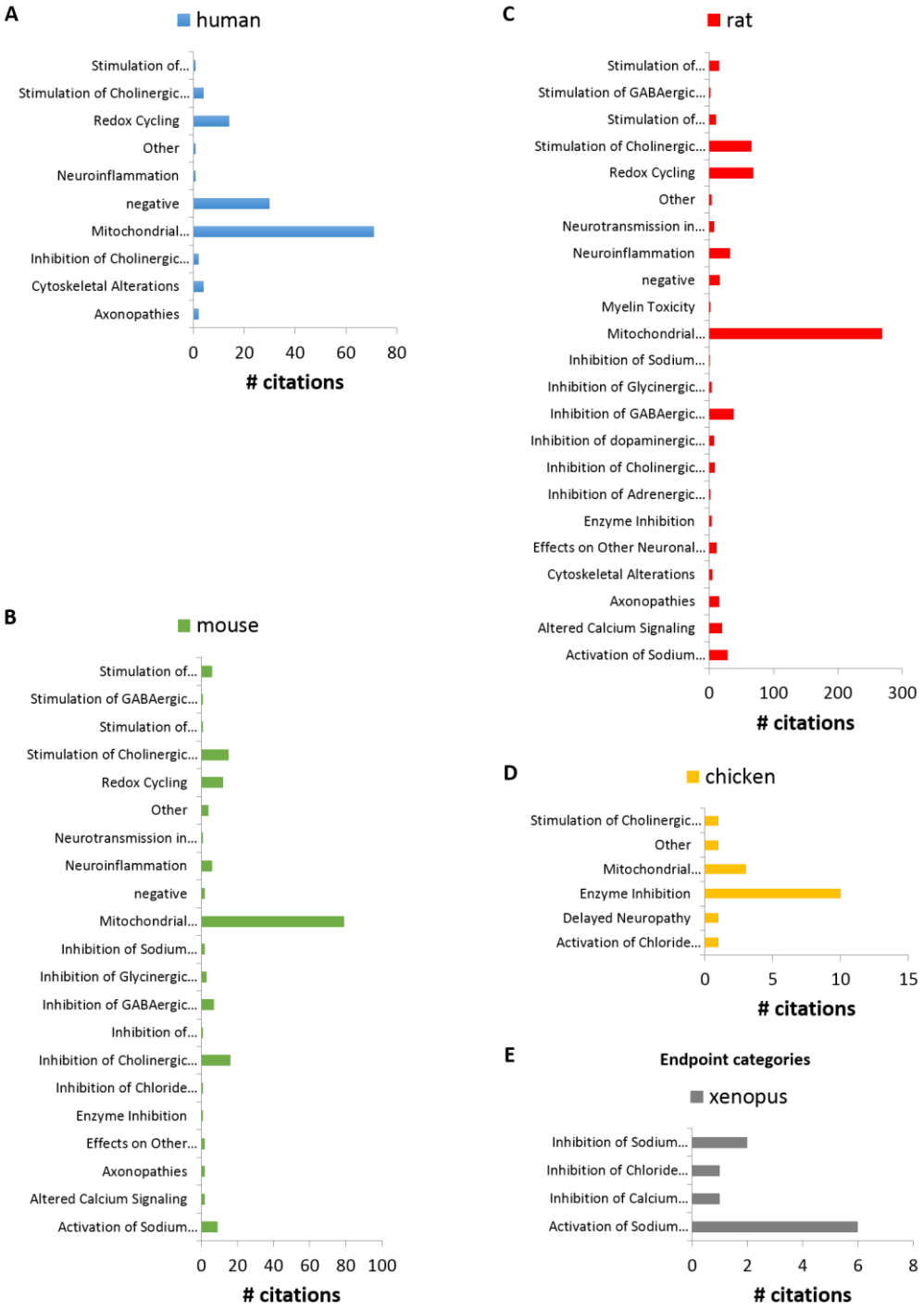


Figure 7 Distribution of # citations across endpoint categories within the different species. Endpoint categories were derived from the compounds' MoAs.

When evaluating the collected 1223 publications, only the studies that were in agreement with the identified MoA of the neurotoxic compounds were included leading to the 977 citations. This procedure guarantees, that an effect observed in the *in vitro* system is a 'true' effect and allows

the categorization into 'true positives' and 'false negatives'. A list of compounds contributing to endpoint category evaluations is given in Appendix D. In addition, negative compounds identified in the studies were classified into 'true negatives' and 'false positives'. Appendix I2 summarizes the total number of citations for each endpoint category divided by species as well as the number of true and false positives and negatives for each endpoint category and species. In the following sections, especially the false negatives will be analysed in the individual studies' contexts for verification of results.

Human

These analyses show that for 131 human citations, 89 were true positives and 22 true negatives. 8 false positives were identified and 12 false negatives. Of the 12 false negatives, 10 citations were studies in stem-/progenitor cells and 2 citations studies in immortalized cells. Going one level deeper into the data, i.e. analyzing the cell types (Table 3) and test systems (Appendix I3) that the false positives were generated with, reveals that the 11 false negatives belonged to the endpoint categories **'mitochondrial dysfunction/oxidative stress/apoptosis'** (11 citations) and **'axonopathies'** (1 citation).

Table 3: Total number of citations for endpoint categories grouped for human cell types. Predictivity analyses were performed for each endpoint category and cell type by analysing true positives (t.p.), false positives (f.p.), true negatives (t.n.) and false negatives (f.n.).

	f.n.	f.p.	t.n.	t.p.	total #
Human	**12**	**8**	**22**	**89**	**131**
cell free				**3**	**3**
Stimulation of Cholinergic Neurotransmission				3	3
Immortalized cells	**2**			**47**	**49**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	2			35	37
Redox Cycling				11	11
Stimulation of dopaminergic Neurotransmission				1	1
primary cells				**11**	**11**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				10	10
Stimulation of Cholinergic Neurotransmission				1	1
primary tissue				**2**	**2**
Cytoskeletal Alterations				2	2
stem/progenitor cells	**10**	**8**	**22**	**25**	**65**
Axonopathies	1			1	2
Cytoskeletal Alterations				2	2
Inhibition of Cholinergic Neurotransmission				2	2
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	9			16	25
Negative		8	22		30
Neuroinflammation				1	1
Other				1	1
Redox Cycling				3	3

False negative data for the endpoint categories **'mitochondrial dysfunction/oxidative stress/apoptosis'** was generated with LUHMES cells (2 citations), which belong to the immortalized cells, with hiPSC-derived neurons (4 citations) or mixed cultures (2 citations) and to undifferentiated neurospheres (3 citations).

These false negatives in the **LUHMES cells** were the endpoint LDH leakage upon treatment with MPP+ within 1 study (Smirnova *et al*, 2016). However, LUHMES cells had 30 true positives for this endpoint. These 30 citations were derived from 5 publications using a total of 4 compounds (cyanide, methylmercury, MPP+ and rotenone). True positives for MPP+ came from the same study by Smirnova *et al*, (2016), yet measured viability via the resazurin assay. Thus, it is not the test system producing a false negative, but the test method. Interestingly, in the LUHMES cell system rotenone produced true positives for both, the LDH leakage and the resazurin assay (Smirnova *et al*, 2016). Thus, the LDH assay cannot be per se judged as non-applicable for assessing cell death in LUHMES cells. Possibly, it is a matter of detection time because apoptotic cells primarily do not release LDH because they are forming apoptotic bodies, which are then eliminated by immune cells. From a total of 32 compounds tested, this cell system has a false negative rate of 6.25% becoming 0% when considering the LDH aspect.
The 3 false negatives in the **hiPSC-derived cells** were diverse endpoint measures (electrophysiological recording, DAPI staining, caspase 3/7 assay) with rotenone, methylmercury, chlropyrifos-oxon or arsenite published in 3 articles (Zagoura *et al*, 2017; Druwe *et al*, 2015; Li *et al*, 2005). With such cells, also true positives (5 citations) were identified using hiPSC-derived mixed cultures and 5 citations using hiPSC-derived neurons. Compounds were also rotenone (Zagoura *et al*, 2017b), and here again, one study compiled positive and negative endpoints measures for the same test system and compound: while effects on DAPI staining and electrophysiological recordings were negative, immunocytochemistry and qPCR were positive. Hence, rotenone administered for 24 hours had no acute cytotoxic effect on the cells, yet induced Nrf2 translocation, reduced Keap cytoplasmic localization, and subsequently induced the Nrf2-dependent gene NQOR1. Moreover, rotenone induced GFAP and reduced tyrosin hydroxylase immunoreactivity (Zagoura *et al*, 2017b). These data suggest that the hiPSC-derived mixed cultures detect rotenone effects on neurons and glia and thus the false negative is not a 'real' false negative, just reflects the ability of the cells to compensate for rotenone effects on mitochondria, at least for 24 hours. Considering the Zagoura et al. study as not a 'real' false negative, the false negative rate of hiPSC-derived mixed cultures is 0%.
Also methylmercury and chlorpyrifos-oxon had true positive and false negative citations in one study (Druwe *et al*, 2015). In this study not mixed cultures, but pure hiPSC-derived neurons **(iCell neurons)** were used. Methylmercury and chlorpyrifos-oxon reduced a protease marker of cell viability, yet did not activate caspase-3/7. For arsenite, neither was affected (false negatives). It is possible that the false negative effects are due to lack of astrocytes in the cultures because astrocytes are crucial for neuronal protection, yet can also confer to neurotoxicity (Kubik & Philbert, 2015; Maurer & Philbert, 2015). Other studies measured effects of methylmercury on mitochondria function and LDH leakage (true positives; Wilson *et al*, 2014). No caspase activity was measured here. This work uses hES-derived neurons purchased from Aruna Biomedical, Inc. (Athens, GA, USA), also in absence of glia. Interestingly, 12 positive and negative NT compounds used in this study produced true positive data for mitochondria-related/cytotoxicity endpoints for 6-hydroxydopamine, acrylamide and methylmercury, true

negative data for the same endpoints for saccharin, nadolol, metformin, amoxicillin, while for celecoxib and ascorbic acid mitotracker was true negative, but MTT was false positive. Diphenhydramin was false positive for both tests, mitotracker and MTT. From a total of 9 compounds tested, this cell system has a false negative rate of 44.44%.
Also false negative data was collected from studies in **undifferentiated NPC**, a cell type representing the NPC niche in the adult hippocampus. However, one of the false negative citations was due to timing, i.e. 2-13 hours rotenone exposure did not activate caspase-3, while 24 hours exposure produced a true positive result (Li *et al*, 2005). The two additional false negatives belonged to non-activation of caspase-9. Caspase-9 is the first caspase interacting with cytochrome c released from the mitochondria and should be activated before caspase-3. Hence, timing of endpoint determination might be an issue here, and not that the cells do not undergo apoptosis as seen by the activated caspase-3. Because of the uncertainties and the overall low number in compounds studied, a predictivity analysis does not seem reasonable.
In the endpoint category **'axonopathies'**, one true and one false positive was identified. Again, these belong to the same study (Wilson *et al*, 2015) and show different effects of acrylamide on hES-derived neuronal cell viability and neuronal morphology indicating that neuronal morphology is a more sensitive endpoint than cell viability.

Table 4: List of negative compounds

True negative		
Compound	**CAS#**	**Species**
Acetaminophen	103-90-2	Mouse
Phthalate		Mouse
1,2 Propandiol	57-55-6	Rat
Quinmerac	90717-03-6	Rat
Paraquat	1910-42-5	Rat
Salicylic acid	69-72-7	Rat
Ibuprofene	15687-27-1	Rat
(2,4-Dichlorophenoxy)acetic acid	2702-72-9	Rat
Amoxicillin	26787-78-0	Human
Celecoxib	169590-42-5	Human
D-sorbitol	50-70-4	Human
L-ascorbic acid	50-81-7	Human
Metformin hydrochhloride	1115-70-4	Human
Nadolol	42200-33-9	Human
Saccharin	82385-42-0	Human
Acetaminophen	103-90-2	Human
Amoxicillin	26787-78-0	Human
Glyphosate	1071-83-6	Human
Rotenone	83-79-4	Rat
Cyanide	57-12-5	Rat

False positives		
Compound	**CAS#**	**Species**
Aniline	62-53-3	Rat
Celecoxib	169590-42-5	Human
Diphenhydramine	147-24-0	Human
L-ascorbic acid	50-81-7	Human
Nadolol	42200-33-9	Human
Saccharin	82385-42-0	Human

Of the **negative compounds** documented within this review with human models (Table 3), 22 citations were true negatives and 8 citations **false positives**. Looking at the compounds behind these data, saccharin or nadolol altered neuronal morphology (false positive), while it did not alter mitochondrial transmembrane potential (assessed by mitotracker), cytotoxicity (LDH release) or cell viability (MTT assay) and was thus a true negative on these endpoints in hES-derived neurons (Wilson *et al*, 2015). Similarly, celecoxib reduced cell viability (false positive), yet had no effect on mitochondrial transmembrane potential or neuronal morphology (true negative). L-ascorbic acid altered cell viability and neuronal morphology with no effect on mitochondrial transmembrane potential (Wilson *et al*, 2015). Acetaminophen and glyphosate behave as a true negative in the caspase and protease marker cell viability assays (Druwe *et al*, 2015) and also amoxicillin is negative in all tested endpoints (Druwe *et al*, 2015; Wilson *et al*, 2015). Sorbitol and metformin did not alter any tested endpoint either (true negative; Wilson *et al*, 2015). In contrast, there are compounds like diphenhydramine, which is stated as a non-toxic compound, but produces effects in all endpoints tested (cell viability, mitochondrial transmembrane potential, neuronal morphology) and accordingly is a false positive. However, antihistamines like diphenhydramine are known to have acute CNS effects in humans (Simons E, 1994). Therefore, the a priori classification of diphenhydramine by the authors might have been incorrect (Wilson *et al*, 2015).

Taken together, these analyses show that for the cell types, the number of false negatives and false positives (Table 4) seems over-estimated. Most of the false-negatives are due to multiple cytotoxicity/cell viability assays, where most of the time the LDH assay is the least sensitive. Effectively, when cells undergo apoptosis, LDH is not, or to a later time point released because apoptotic cells are primarily eliminated via apoptotic bodies (Figure 8). Of the false positives 4 might be due to a priori misclassification.

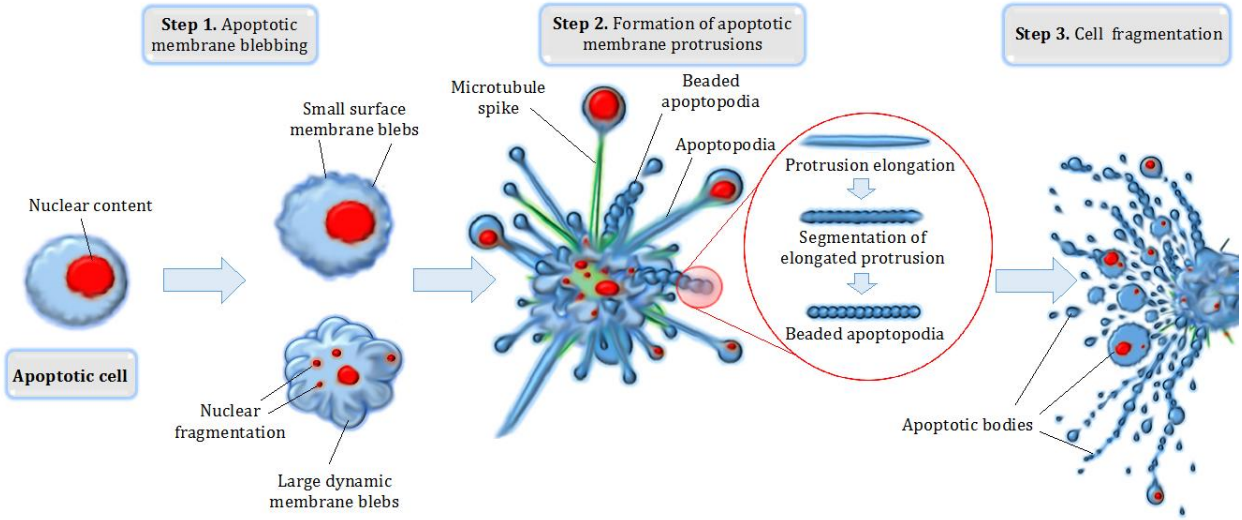


Figure 8 The process of apoptosis. By Aaron Smith, Michael AF Parkes, Georgia K Atkin-Smith, Rochelle Tixeira, Ivan KH Poon - Wikiversity:Draft:WikiJournal of Medicine/Cell disassembly during cell death, CC BY 4.0, https://commons.wikimedia.org/w/index.php?curid=59865845

Performance of human test systems is summarized in the following bar graphs. For each species, these are divided into one graph with true positives and negatives on the top and one graph with false positives and negatives on the bottom. The data has to be regarded with caution, especially for the false negatives as discussed above.

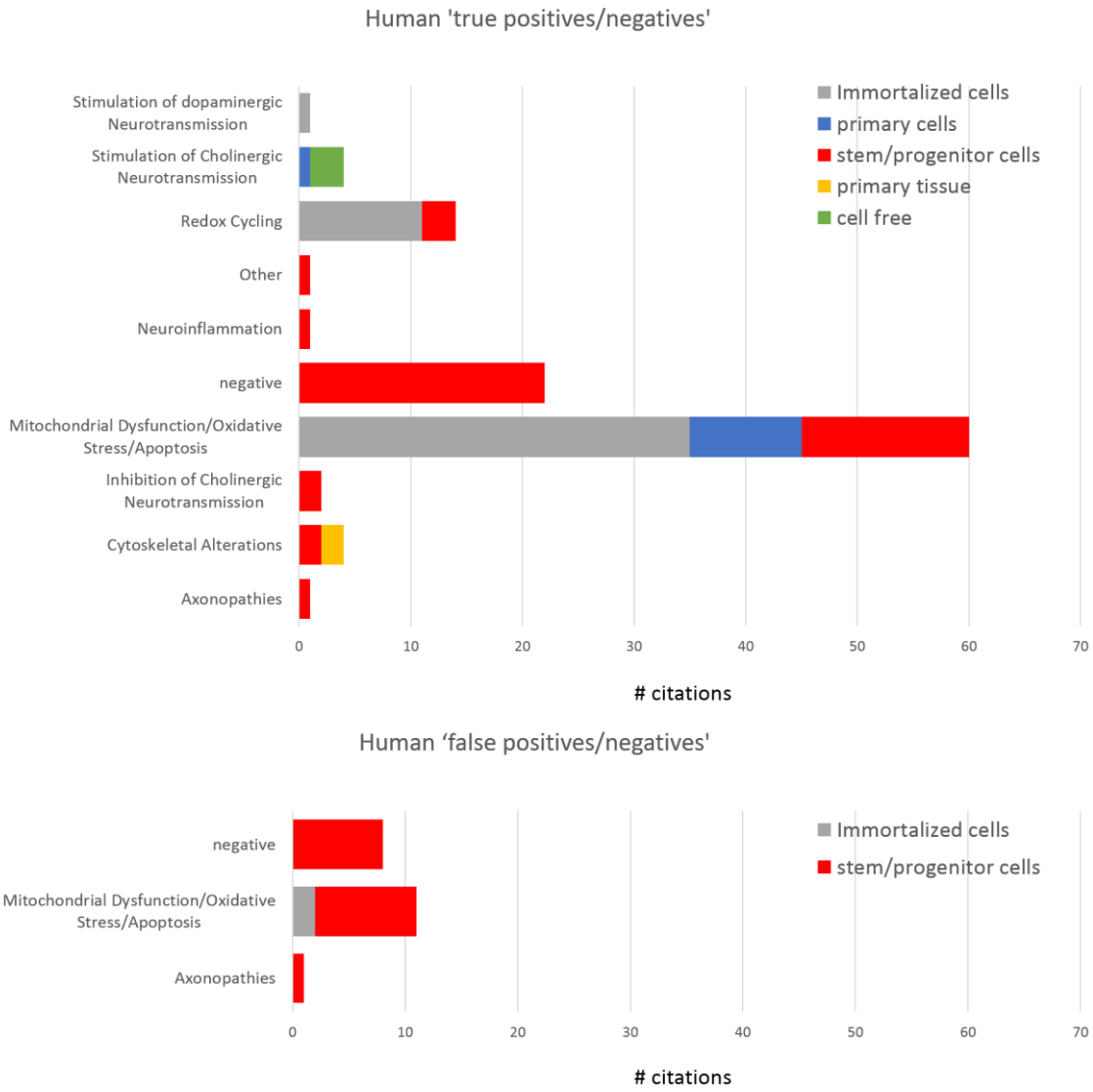


Figure 9 Performance analyses of human cell types with regards to the ability to identify compounds of the respective endpoint categories correctly. The false negatives of the endpoint category have to be regarded with caution (see discussion above).

This graph (Figure 9) demonstrates that immortalized cells, which are mainly LUHMES cells and also a few citations with immortalized mesencephalic cells (MESC2.10; Appendix I3) are valuable to assess few endpoint categories with a very good prediction for these endpoints. Primary cells are hardly ever used and are due to ethical and practical reasons not suited for testing purposes. Stem-/progenitor cell-based methods dominate the different endpoint categories, but compared to the sum of endpoint categories relevant for neurotoxicity evaluation, these are only a few. False compound classification is in reality less than shown in this graph as already stated above. Thus, stem-/progenitor cell-based test systems have the ability to assess multiple endpoint categories and seem suited for neurotoxicity evaluation. Clearly more research is

needed to enlarge the endpoint categories with these cell models as they are unlimitedly available and bear no ethical concerns.

Mouse

Over a total of 174 mouse citations, 163 resulted true positive, 2 true negative and 9 false negative. Endpoint categories reporting false negative are **'mitochondrial dysfunction/oxidative stress/apoptosis'** (6 citations), **'Redox-cycling'** (2 citations) and **'stimulation of cholinergic transmission'** (1 citations). Several endpoints have been considered in these categories.

Table 5: Total number of citations for endpoint categories grouped for mouse cell types. Predictivity analyses were performed for each endpoint category and cell type by analysing true positives (t.p.), false positives (f.p.), true negatives (t.n.) and false negatives (f.n.).

	f.n.	f.p.	t.n.	t.p.	total #
Mouse	**9**		**2**	**163**	**174**
cell free				**3**	**3**
Stimulation of Cholinergic Neurotransmission				3	3
Immortalized cells	**2**			**25**	**27**
Axonopathies				1	1
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				12	12
Redox Cycling	2			10	12
Stimulation of Glutamatergic Neurotransmission				2	2
primary cells	**7**		**2**	**107**	**116**
Activation of Sodium Channels				9	9
Altered Calcium Signaling				1	1
Axonopathies				1	1
Enzyme Inhibition				1	1
Inhibition of Chloride Channels				1	1
Inhibition of Cholinergic Neurotransmission				4	4
Inhibition of GABAergic Neurotransmission				7	7
Inhibition of Sodium Channels				2	2
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	6			58	64
Redox Cycling				1	1
Negative (compound used as negative)			2		2
Neuroinflammation				6	6
Neurotransmission in General				1	1
Other				2	2
Stimulation of Cholinergic Neurotransmission	1			8	9
Stimulation of dopaminergic Neurotransmission				1	1
Stimulation of GABAergic Neurotransmission				1	1
Stimulation of Glutamatergic Neurotransmission				4	4

primary tissue	**8**	**8**
Effects on Other Neuronal Receptors	2	2
Inhibition of dopaminergic Neurotransmission	1	1
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	3	3
Stimulation of Cholinergic Neurotransmission	2	2
stem/progenitor cells	**19**	**19**
Altered Calcium Signaling	1	1
Inhibition of Cholinergic Neurotransmission	12	12
Inhibition of Glycinergic Neurotransmission	3	3
Other	2	2
Stimulation of Cholinergic Neurotransmission	1	1

In the endpoint category **'mitochondrial dysfunction/oxidative stress/apoptosis',** dieldrin results as false negative for general cell death in mouse mesencephalic mixed glia-neuron cells when determined by counting cells visualized by DAPI-staining. When assessing specific cytotoxicity for DA- or GABAergic neurons by means of TH+ and GAD+ neurons count respectively, the compound is a true positive. This points to the relevance of the test system and again here, similar to the human studies, in light of the whole study, lack of general cytotoxicity cannot be considered as a false negative for the test system (Sanchez-Ramos *et al*, 1998). Different ·-carboline studies found 3 false negative citations in 2 publications (Hamann *et al*, 2006; 2008). All three were 9-methyl-beta-carboline xHCl, which belongs to the ·-carboline compound class, yet is not a neurotoxic, but a neuroprotective member of this class (Hamann *et al*, 2008). Hence, primary mesencephalon cells from mouse brain do not have any 'real' false negatives in the endpoint categories **'mitochondrial dysfunction/oxidative stress/apoptosis'.**

In the endpoint category '**Redox-cycling',** paraquat resulted as false negative for ROS production measured with DCF on BV2 microglial cells, when measurements were performed within 5h and for viability (assessed by MTT test) within 24h (Miller *et al*, 2007). Both these parameters are significantly affected at 24h and 48h respectively for Paraquat 50 ·M, suggesting a delayed cell death phenomenon (Miller *et al*, 2007). These results are consistent with what is observed in neuronal cultures from mice (mesencephalon and hippocampal immortalized neurons) and N9 immortalized microglia, where viability and oxidative stress are affected within 24 h but at higher concentrations (300-500 ·M) (Zhao *et al*, 2017; Lee *et al*, 2015a). Thus, is not the test system that produces the false negative but the experimental conditions chosen.

In the endpoint category **'stimulation of cholinergic transmission'**, glyphosate produced a false negative result in cortical mouse cultures grown on MEAs (Vassallo *et al*, 2017). This was in agreement with cortical rat cultures of different laboratories in the same study as well as in other studies (McConnell *et al*, 2012; Valdivia *et al*, 2014; Alloisio *et al*, 2015). In this study, chlorpyrifos oxon, deltamethrin and domoic acid were identified as true positives. Chlorpyrifos oxon is in the same endpoint category than glyphosate. In addition, glyphosate did not produce a single positive hit in this data collection. Hence, time and dose/concentration have to be questioned for the identified MoA. It can also be considered that the identified MoA might not be solely responsible for glyphosate toxicity.

Taken together, the false negatives identified in the data evaluation, were for the most part not 'real' false negatives, giving the mouse systems have a very good predictivity. Over all, performance of mouse test systems is summarized in the following bar graphs. For each species, these are divided into one graph with true positives and negatives on the top and one graph with false positives and negatives on the bottom. The data has to be regarded with caution, especially for the false negatives as discussed above.

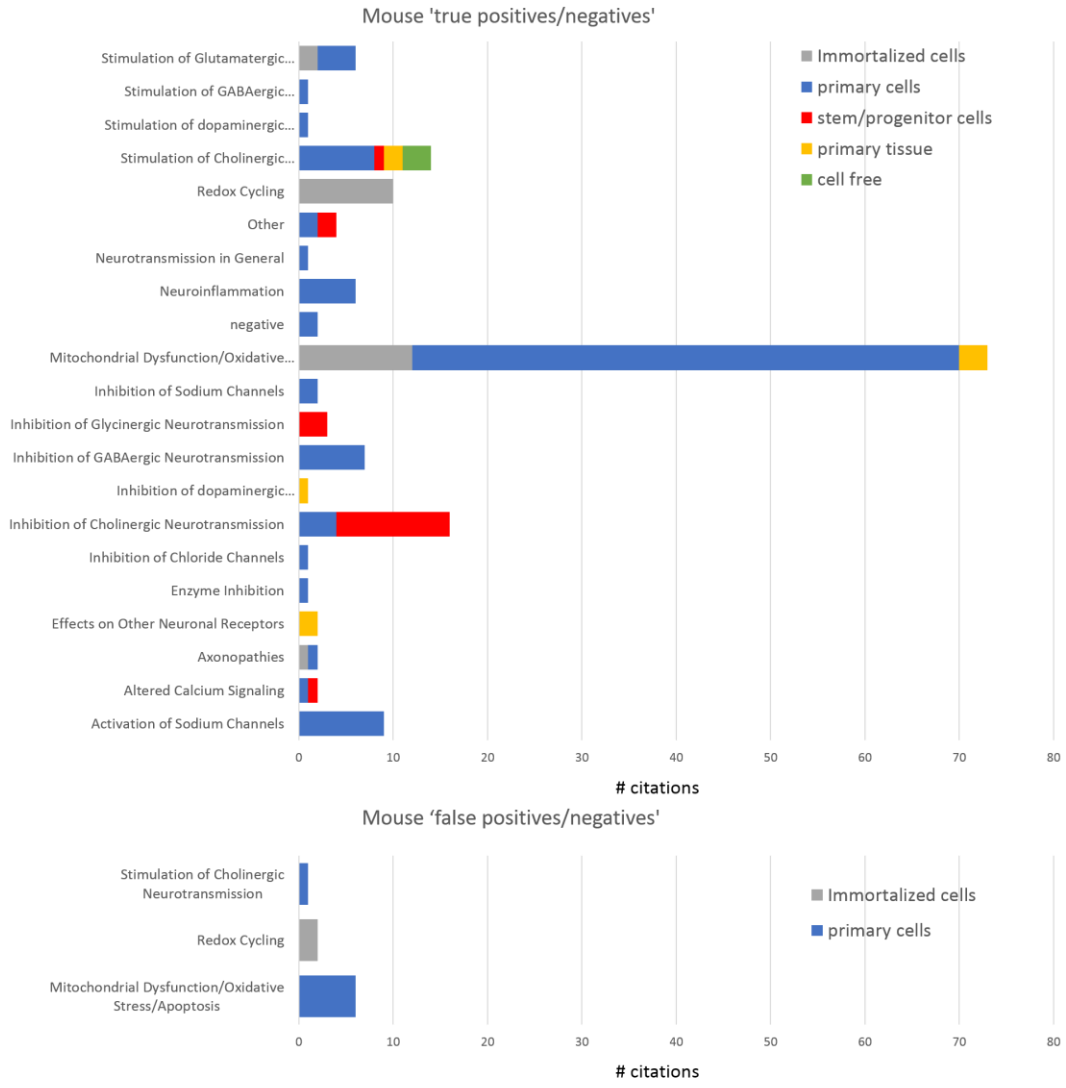


Figure 10 Performance analyses of mouse cell types with regards to the ability to identify compounds of the respective endpoint categories correctly. The false negatives of the endpoint category have to be regarded with caution (see discussion above).

Figure 10 illustrates that in the mouse over all there are very few false positive or false negative studies identified. Immortalized mouse cells, which are mainly BV-2 cells, similar to human immortalized cells, have the ability to study few endpoint categories with a very good prediction

for these endpoints. Primary cells dominate mouse studies. 18 endpoint categories are studied with this cell type, few false negatives that are discussed above and do not really classify as false when going back into the studies. Stem-/progenitor cell-based methods of mouse are few. These are used for studying cholinergic neurotransmission-related endpoints, as are also cell free methods. Data of primary mouse cells will serve as a good reference when working with and comparing data to human stem-/progenitor cell-based methods. Such species comparisons were previously performed with mouse and rat tumor cells (Schmuck & Ahr, 1997; Canete & Diogene, 2010, 2008, Hong *et al*, 2016, 2013; Huff & Abou-Donia, 1995; Case *et al*, 2016; Zhang *et al*, 2007; Campanha *et al*, 2014) and can be transferred to human stem-/progenitor cell-based methods. Such analyses can validate human *in vitro* methods for a variety of endpoints relevant for different neurotoxicity MoA.

Rat

Evaluating rat citations, of the 644 total citations, 529 were true positives and 22 true negatives. 85 false negatives were identified and 8 false positive. Of the 85 false negatives, 65 citations were studies in primary cells and 16 citations studies in primary tissue. Going one level deeper into the data, i.e. analyzing the cell types (Table 6) and test systems (Appendix I5) that the false negatives were generated with, reveals that the 85 false negatives mainly belonged to the endpoint categories **'mitochondrial dysfunction/oxidative stress/apoptosis'** (32 citations) and **'neuroinflammation'** (11 citation). These distribute between the cell types primary cells and primary tissue. The remaining are scattered across a large variety of endpoint categories.

Table 6: Total number of citations for endpoint categories grouped for rat cell types. Predictivity analyses were performed for each endpoint category and cell type by analysing true positives (t.p.), false positives (f.p.), true negatives (t.n.) and false negatives (f.n.).

	f.n.	f.p.	t.n.	t.p.	total #
Rat	**85**	**8**	**22**	**259**	**644**
Immortalized cells	**4**			**38**	**42**
Altered Calcium Signaling				1	1
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				17	17
Neuroinflammation				2	2
Redox Cycling	4			18	22
primary cells	**65**	**1**	**17**	**334**	**408**
Activation of Sodium Channels	5			18	23
Altered Calcium Signaling				5	5
Axonopathies				10	10
Cytoskeletal Alterations	4			1	5
Inhibition of Adrenergic Neurotransmission				2	2
Inhibition of Cholinergic Neurotransmission	3	2		4	9
Inhibition of dopaminergic Neurotransmission	1			2	2

	f.n.	f.p.	t.n.	t.p.	total #
Inhibition of GABAergic Neurotransmission	5			33	38
Inhibition of Glycinergic Neurotransmission				4	4
Inhibition of Sodium Channels				1	1
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	21		2	176	199
Myelin Toxicity				2	2
Negative (compound used as negative)		1	15		16
Neuroinflammation	11			20	31
Neurotransmission in General	1			5	6
Other				4	4
Redox Cycling	3		1	39	43
Stimulation of Cholinergic Neurotransmission	9	3	4	38	54
Stimulation of dopaminergic Neurotransmission				9	9
Stimulation of GABAergic Neurotransmission				1	1
Stimulation of Glutamatergic Neurotransmission	3			10	13
primary tissue	**16**			**98**	**114**
Activation of Sodium Channels				6	6
Altered Calcium Signaling	2			12	14
Axonopathies				5	5
Effects on Other Neuronal Receptors	3			9	12
Enzyme Inhibition				4	4
Inhibition of dopaminergic Neurotransmission				6	6
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	11			35	46
Redox Cycling				2	2
Neurotransmission in General				2	2
Stimulation of Cholinergic Neurotransmission				12	12
Stimulation of dopaminergic Neurotransmission				2	2
Stimulation of GABAergic Neurotransmission				1	1
Stimulation of Glutamatergic Neurotransmission				2	2
stem/progenitor cells				**7**	**7**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				5	5
Redox Cycling				2	2

False negative data for the endpoint categories **'mitochondrial dysfunction/oxidative stress/apoptosis'** was generated with primary glia (4 citations), primary microglia (2 citations), primary neurons (9 citations), re-aggregating brain cell cultures (3 citations), isolated mitochondria (9 citations) or synaptosomes (2 citations).

The false negatives in the **primary glia cells (astrocytes)** concern acrylonitrile cytotoxicity measured via MTT and LDH assays, lipid peroxidation measured by F_2-isoprostanes and Nrf2 protein expression. Acrylonitrile caused a strong induction of the cellular glutathione content (Caito *et al*, 2013). Therefore, these cells are able to compensate oxidative damage by increas-

ing defense mechanisms. As there is no cell damage involved, this test system is not suited as a stand-alone NT test system due to its limited sensitivity towards ROS.
The 2 false negatives in the **primary microglia cells** concern acrylonitrile cytotoxicity measured via MTT and LDH assays. Acrylonitrile causes lipid peroxidation measured by F_2-isoprostanes and Nrf2 protein expression. Acrylonitrile caused an induction of the cellular glutathione content, yet, GSH levels in microglia are much lower than in primary astrocytes suggesting the higher susceptibility of microglia towards ROS compared to astrocytes (Caito *et al*, 2013). Although there is no cell death involved, this test system shows signs of oxidative stress, yet it is not suited as a stand-alone NT test system due to its limited cell type composition.
The 9 false negatives in the **primary neuronal cultures** concern 5 different studies. D-amphetamine induces apoptotic cell death in primary fetal cortical cultures with decreasing bcl-X_L and increasing bcl-X_S. Yet, bax gene expression is unaltered (false-negative; (Stumm *et al*, 1999). This cannot be considered as a 'real' false negative because here bcl-$X_{L/S}$ are the driving factors for apoptosis that the cell system is able to detect. The low-concentration study by Nogueira *et al*, (2014) show that D-amphetamine promoted significant mitochondrial dysfunction and elicited neuronal death in primary mixed neuronal/glial cultures. Moreover, in some cultures caspase 3 activity was activated. The 4 false negatives in this study are based on experimenting with exposure time and concentration. In total, this test system is well suitable for studying D-amphetamine-induced ROS mitochondrial dysfunction and apoptotic cell death at exposure levels that are relevant for human exposure. Also Hondebrink *et al*, (2016) studied the effects of D-amphetamine on neuronal cell death. However, they measured neutral red uptake 30 min after exposure, which is far too short for detecting cell death. As these well-characterized, mixed neuronal/glial cultures are used for measuring electrical activity using multiwell microelectrode arrays (MEA), this cell model is well suited for neurotoxicity analyses. In the study by (Gao *et al*, 2002) one false negative was detected, i.e. no reduction of TH-positive neurons in neuron-enriched cultures derived from mesencephalon by rotenone treatment. Because the neuron/glia co-cultures are responsive to rotenone-induced reduction in TH+ cells, neuron-enriched cultures are not best-suited for this MoA. Glia, especially microglia presence enhances rotenone toxicity. Also lead produced false negative results in neuron-enriched primary cultures because it did not induce cytotoxicity (Fujimura & Usuki, 2012). However, it caused neuronal degeneration. Three different MoA were identified for lead (Appendix D), which do not always include cell death. Thus, it is questionable if this false negative result is a 'real' result or if lead is acting via a MoA not involving cell death. Studies with primary neuronal cultures also identified 163 true positives. Considering that not all of the 9 false negative citations are 'real', primary rat neuronal cultures have a false negative rate < 10%. In primary glia and microglia this is higher with 18% and 40%.
The 9 **false negative** data for the endpoint category **'mitochondrial dysfunction/oxidative stress/apoptosis'** in isolated organelles include rat brain as well as liver mitochondria. Here, lack of effect on oxygen consumption (Sayre *et al*, 1991), ROS formation (Fonck & Baudry, 2003), protein or lipid oxidation (Taskiran *et al*, 2007) were the reasons for categorization. Rat brain or liver mitochondria have 37.5% and 100% false negative rate.
The 11 **false negative** data for the endpoint category **'neuroinflammation'** were derived from one study (Cookson & Pentreath, 1994a) using toluene, isopropanol, hexane, acrylamide, MPTP, tin and lead for the endpoints astrocyte death and astrogliosis in primary astrocyte cultures. This paper studied altogether 20 positive compounds with a false negative rate of 55%.
Interestingly, out of all these studies, there was only one citation identified as **false positive**. This study by Defranchi *et al*, (2011) identified 6 compounds (1,2-propandiol, quinmerac,

paraquat, salicylic acid, ibuprofene, 2,4-dichlorophenoxy acetic acid) as true negatives, yet found aniline as a false positive and 5 compounds as true positives (tin, toluene, cicotine, fipronil, carbaryl) for disturbance of neuronal network activity of mixed neuronal/glia cultures growing on MEAs.
Performance of rat test systems is summarized in the following bar graphs. For each species, these are divided into one graph with true positives and negatives on the top and one graph with false positives and negatives on the bottom. The data has to be regarded with caution, especially for the false negatives as discussed above.

Over all, across the highest number of false negatives, where most of are in the end no false negatives with regards to the whole test system evaluation, concerns measures for cytotoxicity. In *in vitro* toxicology, choosing the right cytotoxicity assay, or a combination of different ones is not trivial and needs to be chosen wisely (Méry *et al*, 2017). It is recommended to use a larger variety of measures for the endpoint cell death because cells can die by different modes, which are sometimes difficult to distinguish, and it is thereby essential that the measurement of cell death involves complementary methodologies. Another important point concerns the timing of experimentation as apoptosis often occurs very quickly, whereas necrosis is much slower; thus, the use of cytotoxicity and viability assays is necessary so as to correctly assess different cell death pathways (Kepp *et al*, 2011).

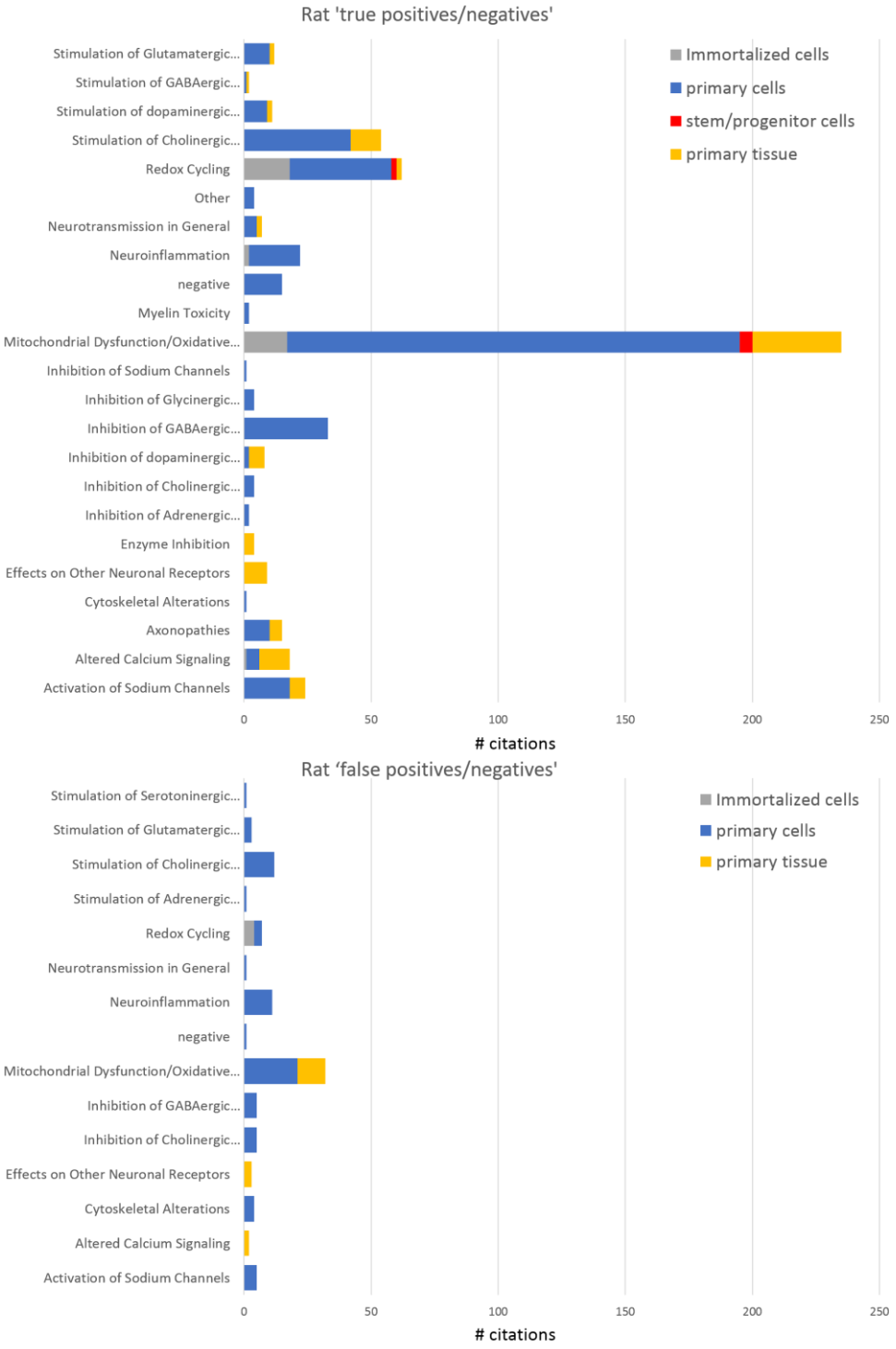


Figure 11 Performance analyses of rat cell types with regards to the ability to identify compounds of the respective endpoint categories correctly. The false negatives of the endpoint category have to be regarded with caution (see discussion above).

Figure 11 illustrates that most endpoint categories (23) are studied in rat and that the rat is the best studied species of the five species evaluated in this report. There are hardly any immortalized rat cells used, the few studies on the same endpoint categories than in human and mouse immortalized cells relating to apoptosis and cell death. Rat primary cells dominate these studies and all 23 endpoint categories can be assessed with this cell type. Most false classifications fall again in the apoptosis-cell death endpoint category. Of these, primary neuronal cultures (which often are neuronal-glia co-cultures) perform better than pure astrocyte or microglia cultures. Stem-/progenitor cell-based methods of rat are even fewer than from mouse and negligible. Such data of primary rat cells will – similar or in addition to the primary mouse cells - serve as a good reference when working with and comparing data to human stem-/progenitor cell-based methods. As already stated above, such species comparisons were previously performed with mouse and rat tumor cells (Schmuck & Ahr, 1997; Canete & Diogene, 2010, 2008, Hong *et al*, 2016, 2013; Huff & Abou-Donia, 1995; Case *et al*, 2016; Zhang *et al*, 2007; Campanha *et al*, 2014) and can be transferred to human stem-/progenitor cell-based methods. Such analyses can validate human *in vitro* methods for a large variety of endpoints relevant for different neurotoxicity MoA identified in this analysis.

Chicken and Xenopus

The analyses were also performed for methods based on chicken and Xenopus. The few Xenopus studies had no false negative, only true positive data. Thus, negative compounds are missing in the experimental set-up. Yet, the measurements of direct effects of compounds on overexpressed receptors is straight-forward and confirms MoA of DDT, terfluthrin, deltamethrin, ·-conotoxin, chorotoxin and tetradotoxin. Although not in a physiological context, these methods seem useful as a screen for molecular initiating events (MIEs). Unfortunately, the throughput of this test system is low. Table 7 summarizes the individual test systems of the xenopus assays and Figure 13 shows its performance.

Table 7: Total number of citations for endpoint categories grouped for Xenopus cell types. Predictivity analyses were performed for each endpoint category and cell type by analysing true positives (t.p.), false positives (f.p.), true negatives (t.n.) and false negatives (f.n.).

	true positive	total #
primary cells	**10**	**10**
neurolemma in Xenopus Oocytes	**6**	**6**
Activation of Sodium Channels	2	2
Inhibition of Calcium Channels	1	1
Inhibition of Chloride Channels	1	1
Inhibition of Sodium Channels	2	2
sodium channels in Xenopus Oocytes	**4**	**4**
Activation of Sodium Channels	4	4
total	**10**	**10**

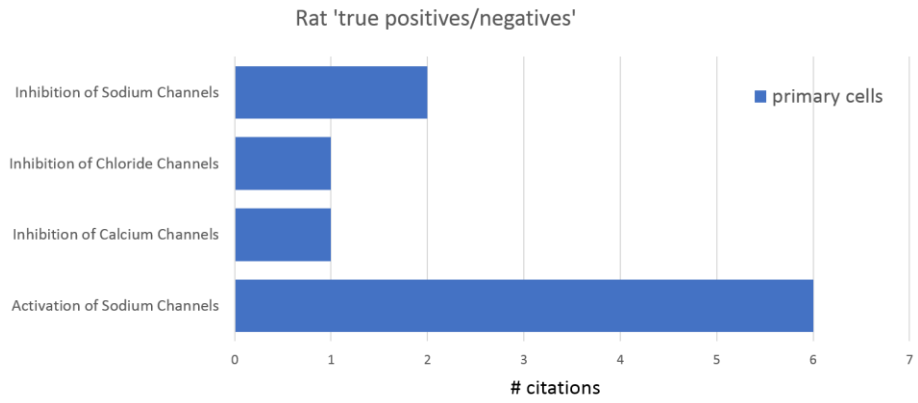


Figure 12 Performance analyses of Xenopus cell types with regards to the ability to identify compounds of the respective endpoint categories correctly. The false negatives of the endpoint category have to be regarded with caution (see discussion above).

Data analyses for 18 chicken citations revealed 15 true positives and 3 false negatives. Most true positive citations (10) were acetylcholinesterase inhibition using parathion and paraoxon in presence of S9 mix (Sawyer *et al*, 1992), in the endpoint category **'enzyme inhibition'**. As in the other species, false negatives are due to lack of ROS-related endpoints in the endpoint category **'mitochondrial dysfunction/oxidative stress/apoptosis'** upon chicken cerebellar granule cells arsenic exposure. However, reduced GSH was measured indicating the compensatory abilities of the test system (Castro-Coronel *et al*, 2011). Table 8 summarizes the individual test systems of the chicken assays and Figure 14 shows its performance.

Table 8: Total number of citations for endpoint categories grouped for chicken cell types. Predictivity analyses were performed for each endpoint category and cell type by analysing true positives (t.p.), false positives (f.p.), true negatices (t.n.) and false negatives (f.n.).

	false negative	true positive	total#
primary cells	**3**	**13**	**16**
cerebellar bergmann glia	**3**	**2**	**5**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	2	2	4
Other	1		1
primary neurons		**10**	**10**
Enzyme Inhibition		10	10
Spheroids		**1**	**1**
Activation of Chloride Channels		1	1
primary tissue		**2**	**2**
brain homogenate		**2**	**2**
Delayed Neuropathy		1	1
Stimulation of Cholinergic Neurotransmission		1	1
Total	**3**	**15**	**18**

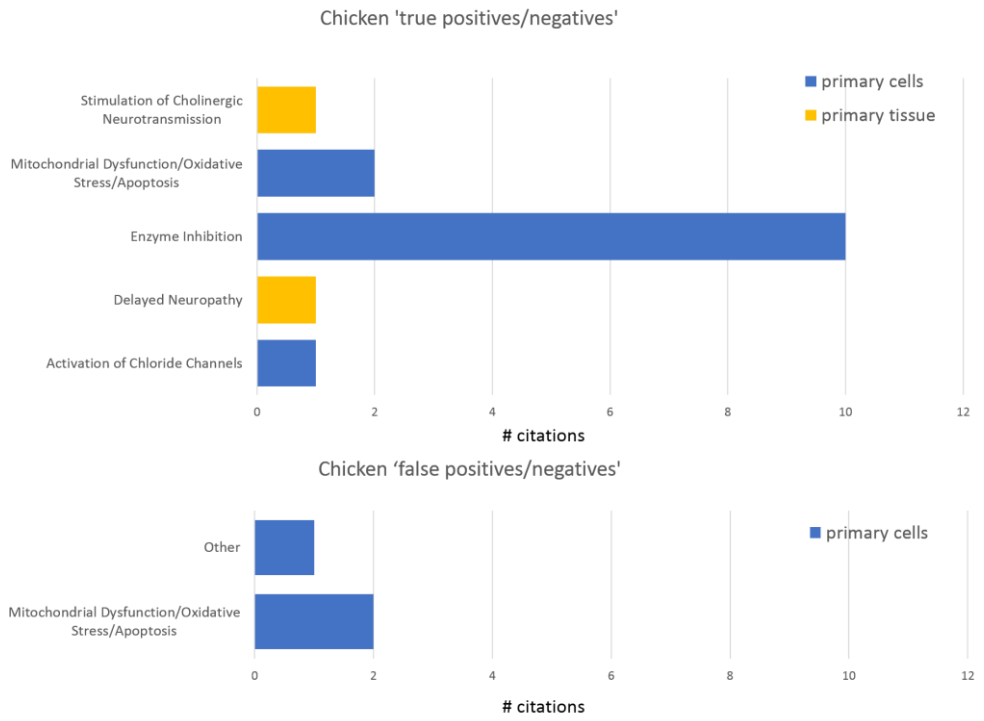


Figure 13 Performance analyses of chicken cell types with regards to the ability to identify compounds of the respective endpoint categories correctly. The false negatives of the endpoint category have to be regarded with caution (see discussion above).

Summary evaluation of the false negatives across all species and cell types/test systems reveal that measures for cytotoxicity are the main issue identifying false negatives incorrectly. In *in vitro* toxicity evaluation, choosing the right cytotoxicity assay, or a combination of different ones is not trivial and needs to be chosen wisely (Méry *et al*, 2017). It is recommended to use a larger variety of measures for the endpoint cell death because cells can die by different modes, which are sometimes difficult to distinguish, and it is thereby essential that the measurement of cell death involves complementary methodologies. Another important point concerns the timing of experimentation as apoptosis often occurs very quickly, whereas necrosis is much slower; thus, the use of cytotoxicity and viability assays is necessary so as to correctly assess different cell death pathways (Kepp *et al*, 2011). Hence, if multiple viability/cytotoxicity assays display contradictive results, the negative one should not be taken as a false negative.

Cell Types and Endpoint categories

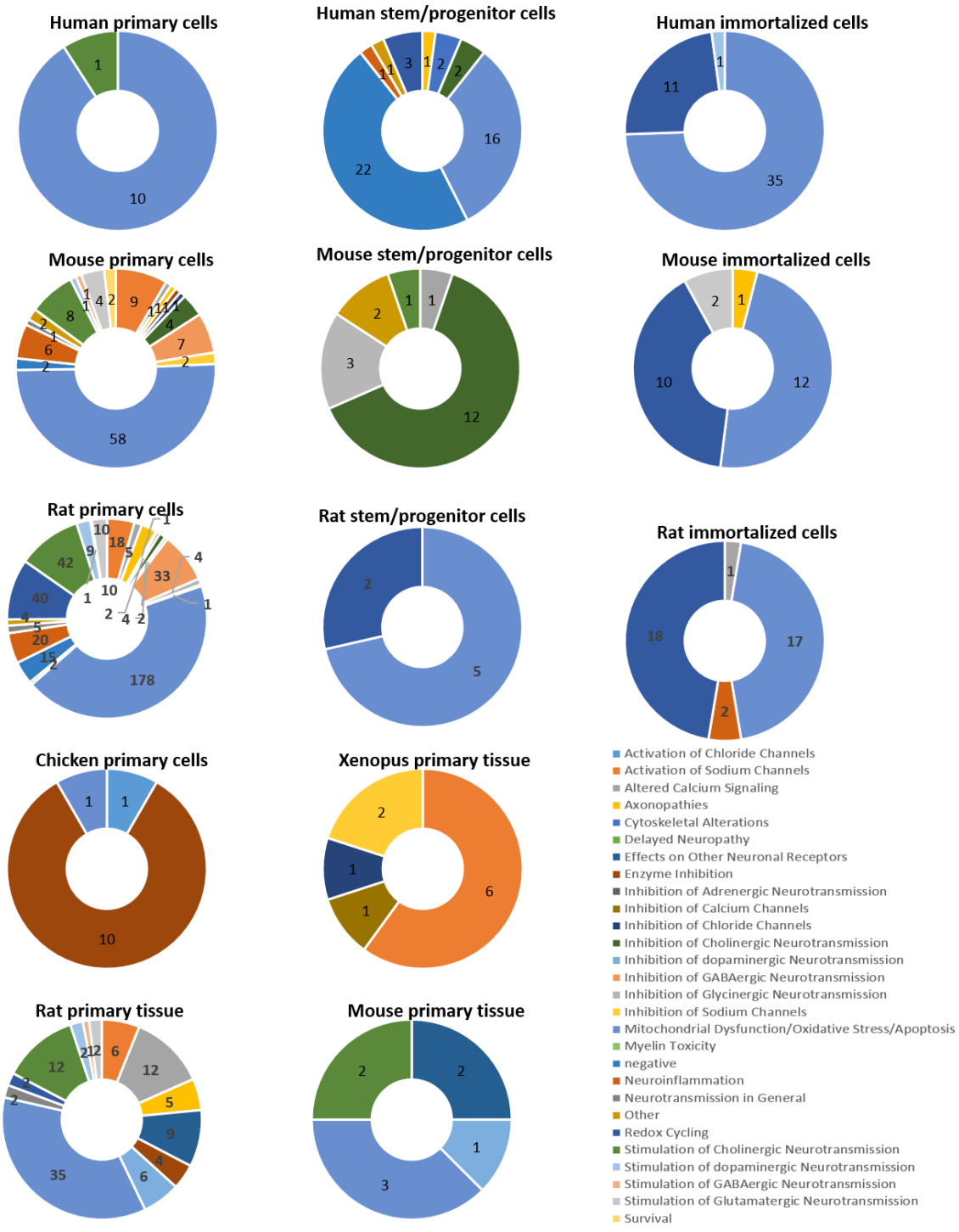


Figure 14 Endpoint categories covered by the different cell types and tissues across the species human, mouse, rat, chicken and xenopus from true positive and true negative data. Information on test systems behind the cell types is given in Appendix I1.

After performance analyses of the different cell types, now the analyses determine the endpoint categories that each cell type from each species can successfully evaluate. Here, the primary MoA analyses that we performed before data collection is extremely helpful, because every true positive and negative result has physiological relevance. This means, that every cell type/test system positive in this evaluation is suitable for neurotoxicity testing of this endpoint category summarizing certain MoA (Appendix D). The donut graphs in Figure 14 summarize these endpoint categories for each cell type. The specific Test systems behind can then be found in Appendix I1.

This way of presenting the data shows on the first glimpse that some cell types are capable of studying a large variety of endpoint categories (primary mouse and rat cells) and that others are more limited in their abilities. Human stem-/progenitor cells, the cell type with the largest promise for human cell-based neurotoxicity tests, are on the way to become useful for neurotoxicity testing applications. However, clearly more work is needed to establish methods that can detect as many endpoint categories in human stem-/progenitor cells as in the rodent primary cells.

Published *in vitro* methods beyond chemical testing

The literature was also mined for studies that are promising methods, yet are not targeted to chemical testing. We identified 123 studies and included 34 of them (Appendix H). Because we already found a high number of rodent studies evaluating NT endpoints, we concentrated on promising studies using human systems. 8 publications were found; of those, 3 studies employ hiPSC-derived neural cells, 2 hESC-based neuronal cells, 3 NSC-derived neurons and 2 primary astrocytes. Because working with ESC is of ethical concern and thus not applicable for regulatory purposes, these studies were not included into the evaluation.

Malik *et al*, (2014) evaluated effects of a compound library with pharmaceuticals on iPSC, iPSC-derived neurons and fetal astrocytes. The set-up in a screening format is very useful (Figure 15). hiPSC-derived neurons were seeded at a density of 30,000 cells/well and assayed on day 10 when >90% of cells were ·(III)tubulin-positive. In the same study, also hiPSC, NSC and fetal astrocytes were used as test systems. There are two drawbacks of the study. For one, there does not seem to be glia cells in the cultures and secondly, the only endpoint assessed was cytotoxicity. Co-culture systems of multiple cell types have a higher physiological relevance and increase application domains of assays. With regards to cytotoxicity evaluation, our data set analysed above clearly show that there are a high number of misclassified compounds due to cytotoxicity measures. Therefore, it is not advisable to use only cytotoxicity screens as stand alone assays for NT evaluation.

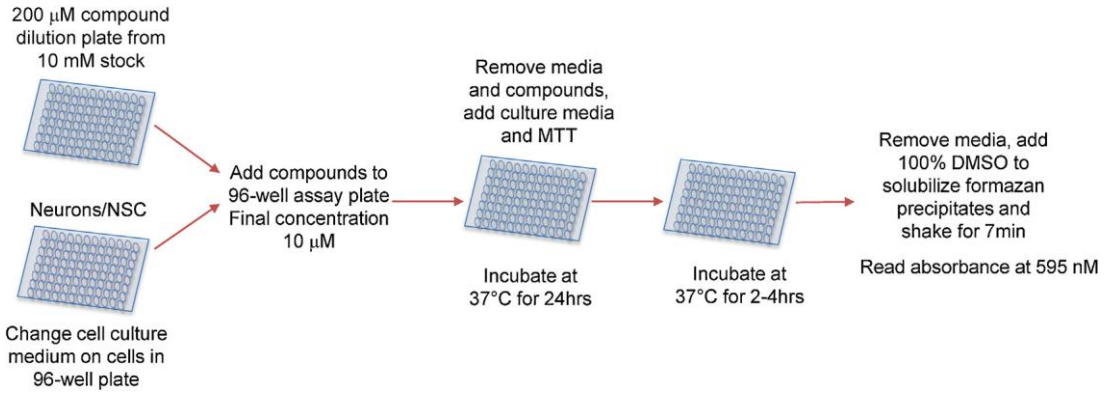


Figure 15 Flow chart depicting how hiPSC-derived neurons were used for compound testing in a 96-well format. Cytotoxicity of pharmaceuticals was assessed. (from: Malik et al. 2014)

Seidel *et al*, (2017) use three different hiPSC lines (2 IMR-90 lines and the 4603c27 line) for differentiation into NPC. After timely defined cultivation of NPC they are differentiated into the neural lineage producing neurons and also astrocytes in a controlled and defined manner. Readout in this study is MEA activity, an endpoint warranted for NT testing because a broad number of endpoint categories can be assessed when different neuronal receptors are present in the cells. This study shows reactivity towards dopamine, GABA, serotonin, acetylcholine and glutamic acid, not to norepinephrine. One drawback of the study is the quantification of MEA measurements. Instead of using different chips or experiments as individual 'n' numbers, single electrodes were used as the statistical unit. Therefore, reproducibility and standardization between experiments cannot be assessed.

A recent study by Tukker *et al*, (2016) used different commercially available hiPSC-derived, mature neurons with and without astrocyte co-culture in comparison to rat primary cortical neurons. Readouts are MEA activity and calcium signalling. Calcium transients of individual iCell neurons are generated upon treatment with glutamate, GABA and acetylcholine. Glutamate and GABA strongly reduced mean spike rate. Limitations of the pure neuronal cultures are absence of bursting and absence of astrocytes. Moreover, more thorough characterization of the cell system is necessary. Similar to the study by Seidel *et al*, (2017), statistical analyses are an issue. Here, one well of a 48-well plate is the statistical unit, not electrodes. However, reproducibility and standardization between experiments can also not be assessed with this statistical procedure.

hiPSC-derived neuronal differentiation was also performed by (Yan *et al*, 2016), who differentiated the hiPSC line iPSK3 to embryoid bodies that further developed to 3D NPC aggregates. These aggregates were plated for neuronal differentiation and used for changing biomolecule-dependent neuronal patterning of brain organoids. This is an interesting approach, especially with regards to brain region-specific differentiation. Others have used the concept of producing NPC as an intermediate cell population with a distinct proliferation capacity to generate reliable terminal differentiated neural tissue (neuron-glia and neuron subtype proportions) in manageable time intervals (Li *et al*, 2011; Hofrichter *et al*, 2017). More work is needed here to understand if such methods can be used for reproducible generation of brain region-specific neural cultures with human cells.

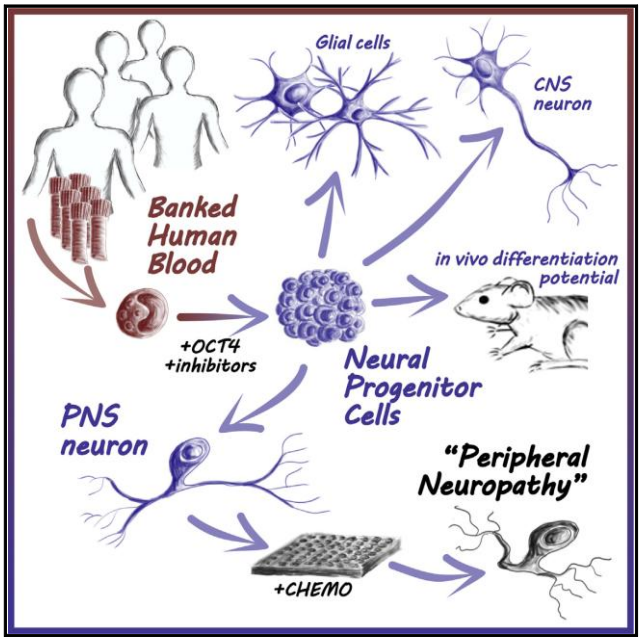


Figure 16 Generation of neural cell types from banked human blood. (from Lee *et al*, 2015)

An interesting report from Lee *et al*, (2015) employed banked human blood as a direct cell source for reprogramming into NPC (Figure 16). Neuronal cultures generated by this method are well-characterized, express a variety of neuronal and glia markers. The NPCs differentiate into central (dopaminergic) and peripheral (nociceptive) neurons, the latter were used for chemotherapy-induced neurotoxicity evaluation. This is an interesting concept also showing the plethora of differentiation protocols available. Most hiPSC-derived neural cells use fibroblasts. However, also other body cells can be used for reprogramming into hiPSC or direct iNPC.

Another cell type that is clearly underrepresented for toxicity studies, yet is of high importance, are oligodendrocytes. Oligodendrocytes can be generated from primary human NPC (Baumann *et al*, 2015; Dach *et al*, 2017), however, to differentiate standardized, large amounts of oligodendrocytes, hiPSC-based methods are preferred. While most oligodendrocyte differentiation protocols take a long period of time (60-150 days; e.g.) and show limited efficiency (Wang *et al*, 2013; Douvaras *et al*, 2014; Djelloul *et al*, 2015), a recent protocol published in PNAS allows rapid and efficient generation of oligodendrocytes from hiPSC (Ehrlich *et al*, 2017).

In summary, most studies using differentiating hiPSC do characterize neurons. This is a highly important topic and overall hiPSC have the ability to differentiate into a variety of neuronal subtypes. In contrast, information on glia, i.e. astro- and oligodendroglia is sparse. For regulatory applications, defined protocols need to be established, that in a reproducible way generate neural cells from hiPSC, cover a large variety of compounds' MoA for NT assessment in a reasonable time frame that are applicable to multiwell plate testing. The EU-Joint Research Center goes into this direction by publishing standard operating procedures for their hiPSC cultivation and differentiation procedures in the DB-ALM database online (https://ecvam-dbalm.jrc.ec.europa.eu/methods-and-protocols/method-summary/differentiation-of-induced-

pluripotent-stem-cells-into-post-mitotic-neurons-and-glial-cells-(mixed-culture)/key/m_1961; protocol number 165 and 166) as well as the accompanying research paper (Zagoura *et al*, 2017a). More studies in this way are needed to produce hiPSC-based cultures tailored for neurotoxicity evaluation *in vitro*.

3.1.1. Grey literature search

The grey literature search for *in vitro* and ao methods did not retrieve any additional information that was not already published. Results of the grey literature search are summarized in Appendix F.

Next to the grey literature search a promising approach for defined neurotoxicity testing was published as poster at the SOT meeting in March 2018 (Saavedra et al. 2018; NeuCyte, CA, USA), yet this work has not been published as a primary research paper. The specialty of this approach lies in the direct reprogramming of hiPSCs into highly functional neurons of defined subtypes and into glia cells. These cells are then seeded in a defined ratio of excitatory neurons, inhibitory neurons and astrocytes onto MEAs, thereby providing a defined human neuronal/glial co-culture platform for comprehensive electrophysiological measurements using MEAs (SynFire iN Cell Technology). When combining defined ratios of glutamatergic to GABAergic neurons together with astroglial cells, robust neuronal activity, including synaptically driven spontaneous synchronized network bursting can be recorded at 3-4 weeks post seeding (Figure 17).

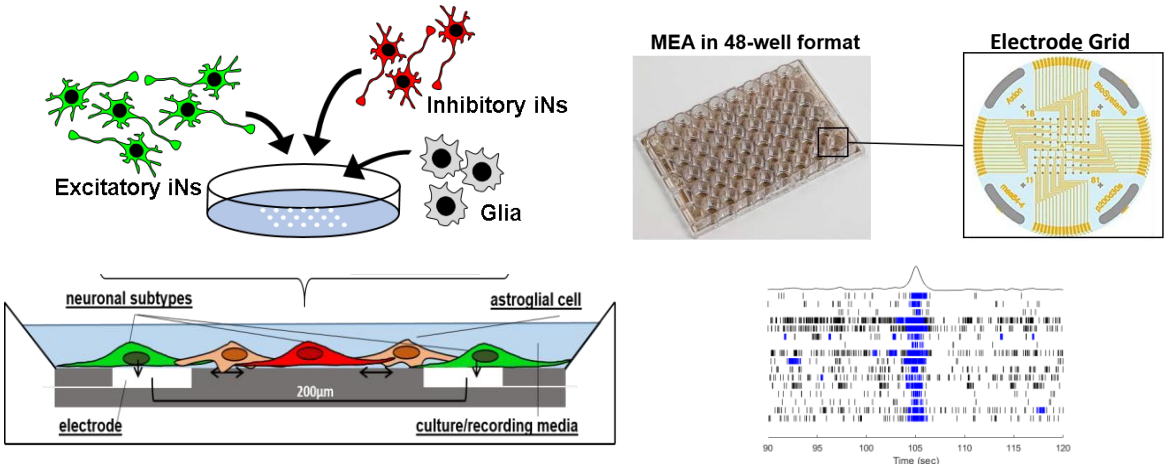


Figure 17 Combination of iN technology with human primary glial supporter cells on 48-well MEA plates to develop a pure human neural co-culture system consisting of glutamatergic excitatory neurons (140K cells/well), GABAergic inhibitory neurons (60K cells/well), and astrocytes (70K cells/well). Reproducible formation of spontaneous synchronized neuronal network activity can be detected 3-4 weeks after plating.

Due to parallel acquisition of multiple parameters, the platform allows detailed characterization of neurotoxicity effects of test compounds. Specifically in this work the platform was used to specifically and quantitatively assess chemically-induced seizure-like activity in a semi high-throughput setting by studying the following defined MoA: Inhibition of $GABA_A$ receptors, stimulation of muscarinic AChR, potassium channel blockage, D2 receptor antagonism, norepinephrine and serotonine re-uptake inhibition and 5-HT_{2A} receptor antagonism. Therefore, 11 compounds with clinically reported seizurogenic effects in patients were studied, which partially remained undetected in rodent-based testing. The human system identified specific alterations in neuronal activity of all test compounds in a concentration-dependent manner, and determined seizure-like firing patterns in the most potent subset. Hence, this is a proof-of-concept study of a novel iPSC-based neuronal/glial *in vitro* approach for the assessment of seizure liabilities of chemical compounds in a human-relevant cell context. The defined cellular context of this work is so far unique, eliminates variability due to individual experiment hiPSC differentiation and thus deserves more MoA analyses in the future.

3.2. Cell based blood brain barrier models

The blood brain barrier (BBB) plays a critical role for any chemical-induced neurotoxic action to the CNS as it determines the availability of a potential neurotoxicant at the target organ. BBB permeability therefore presents a key characteristic to be considered in the evaluation of the neurotoxic potential of a chemical in humans. This is especially true when the NT potential of a compound is determined based on an alternative model (*in vitro*, alternative organism) that does lack or is very limited in the representation of pharmacokinetics of the CNS.
Anatomically the BBB consists of brain microvascular endothelial cells (BMECs), the non-cellular basement membrane (BM) and cells from the neurovascular unit, astrocytes, pericytes, microglia and neurons (Aday et al., 2016; Banerjee et al., 2016; Figure 18B). The barrier function of the BBB consists of several key characteristics. Tight and adherent junctions expressed by BMEC seal the gap between two adjacent cells and prevents the paracellular diffusion of hydrophilic compounds (Figure 18D). Efflux transport systems like the ABC-type transporters transport hydrophobic molecules out of the cell back into the lumen and thereby prevent them from reaching the CNS. The high metabolizing capacity of astrocytes metabolizes hydrophobic compounds that permeated into the neurovascular unit (Banerjee *et al*, 2016b; Helms *et al*, 2016b). In addition to several mechanisms that prevent the penetration of chemicals into the CNS, there are several influx transport systems that provide the CNS with necessary nutrients but might also cause the influx of undesired chemicals. Influx mechanisms are thereby the carrier-mediated transport, receptor-mediated endocytosis and transcytosis and adsorptive-mediated endocytosis and transcytosis (Palmer and Alavijeh, 2013; Figure 18D).
For an alternative model to correctly predict the BBB permeability its key structure and mechanism need to be represented *in vitro*. Therefore, an ideal *in vitro* model should form tight junctions, express transporters, enzymes, macromolecular and immune cell trafficking and signalling and be suitable for rapid screening of BBB permeability of potential CNS toxicants.

There are a number of well-characterized cell-based models that are able to assess the BBB permeability of a chemical as summarized in two reviews from Helms et al. (2016) and Wolff et al. (2015). These models are mainly developed from rat, mouse, bovine, porcine as well as hu-

man material and are based on primary-, tumor/immortalized- or stem/progenitor cells. The most common set up is thereby a transwell cell culture system with either mono-culture, non-contact co-culture, contact co-culture or a triple culture. In addition, several microchip-based BBB models have been developed in the last 10 years.

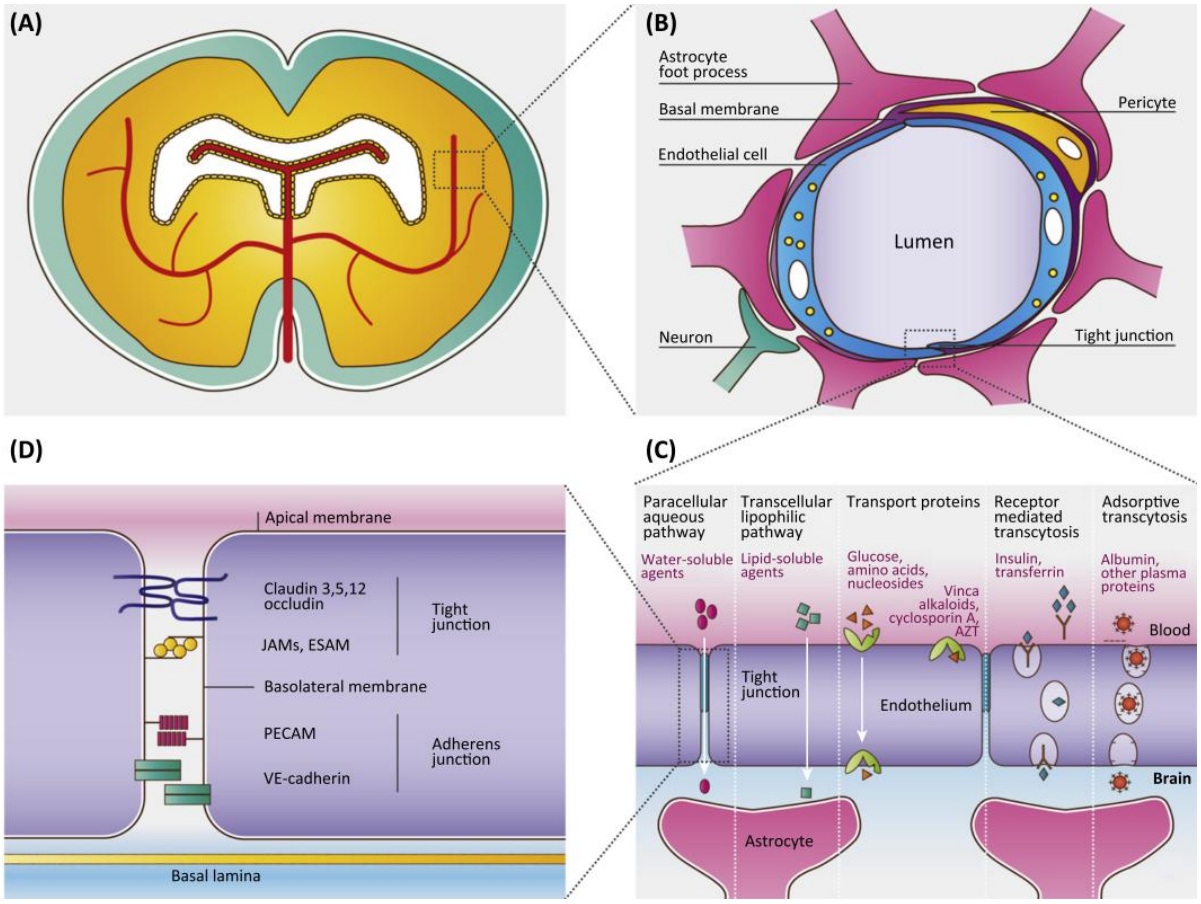


Figure 18: Blood–Brain Barrier (BBB): structure and function. (A) Schematic picture of the blood brain barriers. (B) Schematic picture of the BBB's structure in magnification. The lumen surrounded by endothelial cells (ECs) that are in close association with pericytes, astrocytic end feet and neurons all separated by the basal membrane. (C) different routes of transport across the BBB into the CNS. (D) Structure of Tight junctions (Aday et al., 2016).

The most simplistic cellular BBB model is the mono-culture, which consists of a uniform layer of brain endothelial cells (BECs) that is seeded on a semipermeable membrane filter (Figure 19A). This model has been generated from different cell types and various species. Key advantages of a mono-culture model are its simplicity and cost effectiveness as well as the suitability for high-throughput permeability screening (Banerjee *et al*, 2016b). Major drawbacks, however, are the absence of other influencing cellular components like pericytes, astrocytes or microglia as well as inadequate barrier properties as demonstrated by low transendothelial electrical resistance (TEER) compared to co-cultures (Banerjee *et al*, 2016b). The use of astrocyte conditioned media or dynamic models that introduce shear stress to the system both increase barrier properties (Siddharthan *et al*, 2007). Kim (2009) demonstrated that shear stress, together with astrocyte-conditioned medium can increase the TEER of primary human brain microvascular endo-

thelial cells (HBMECs) by a factor of 3, from 500 to 1500 ohm/cm^2. The use of astrocytes or pericytes co-cultured with brain endothelial cells (BECs) in general increases tight junction characteristics and improves the paracellular barrier properties (Nakagawa *et al*, 2007). Because of the lack of astrocyte or pericytes induction, results from mono-cultures should generally be regarded with caution whereas co-culture systems that consider the interactions between endothelial cells ECs and the surrounding brain microenvironment are more promising models.

Co-culture models mostly combine endothelial cells with astrocytes or pericytes. Thereby two set ups are commonly used. In both set ups BEC are grown on the luminal side of a transwell membrane, whereas the second cell type (astrocytes or pericytes) are grown either on the bottom of the well (non-contact co-culture; Figure 19B) or directly on the abluminal side of the membrane (contact co-culture; Figure 19C). Similar to mono-cultures, co-culture models have also been generated from multiple cell types and species including different combinations of both. The key advantage of co-culture models is their ability to represent an astrocyte or pericyte induction with increased TEER values and improved permeability. They are, however, more complex in their overall handling causing high variations in research results and a limited throughput of permeability screening. Another disadvantage is that the thickness of semipermeable membrane inserts is much thicker than the basal membrane *in vivo* (Banerjee *et al*, 2016b). Similar to the mono-cultures the introduction of shear stress improves barrier function compared to static co-cultures models (Booth & Kim, 2012).

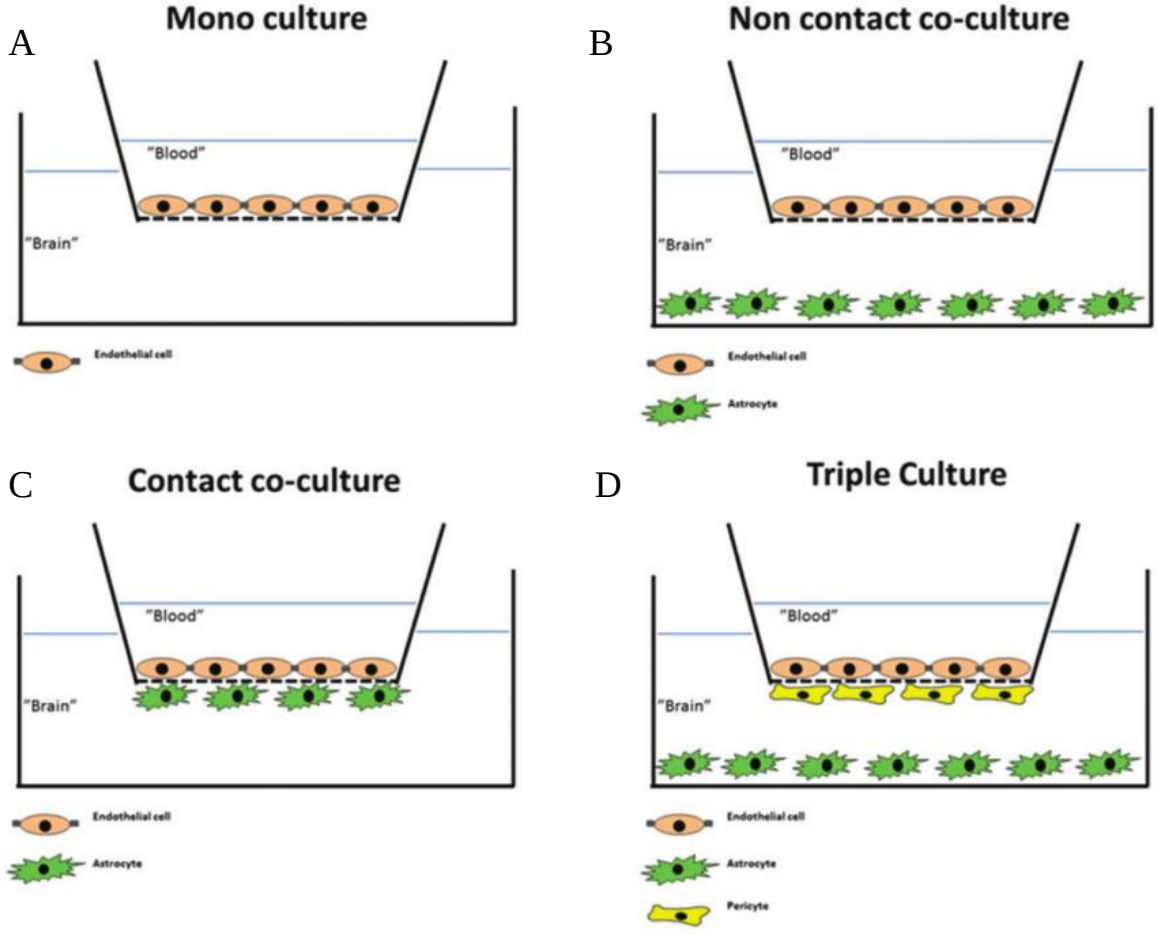


Figure 19 Common configuration of transwell culture system. (A) mono-culture with cells BECs seeded on the luminal site of a semipermeable membrane filter. (B) non-contact co culture with astrocytes seeded on the abluminal site on the bottom of the well. (C) contact co-culture with astrocytes seeded on the abluminal site of the semipermeable membrane. (D) triple-culture with pericytes seeded on the abluminal site of the semipermeable membrane and astrocytes seeded on the abluminal site on the bottom of the well (from Helms et al., 2016*).*

An extension of the co-culture models are triple-cultures that most commonly contain endothelial cells and astrocytes as first and second cell type. As a third cell type neurons have been used traditionally but have been substituted by pericytes in recent years (Wolff *et al*, 2015a). The classical set up of triple-culture models in a transwell format is demonstrated in Figure 19D. A key advantage of the triple culture model is the higher level of resemblance with *in vivo* brain physiology. It was for example demonstrated that triple-cultures of endothelial cells, astrocytes and neurons show a higher expression of zonula occludens-1 (ZO-1) and faster synthesis of occludin, both key proteins of tight junctions, with an overall higher TEER (Schiera *et al*, 2003; Xue *et al*, 2013). Overall triple-cultures show considerable improvement in BBB phenotype and paracellular permeability compared to co-culture models (Banerjee *et al*, 2016b). They are, however, very complex, more expensive and often show low reproducibility with a high variation of data, which makes them less suited for high throughput permeability screening.

In a study by Nakagawa et al. (2007) seven different models and set ups of either rat brain capillary endothelial cells in mono-culture, in co-culture with astrocytes or pericytes as contact or non-contact culture, or as triple-culture with pericytes and astrocytes (contact and non-contact) are compared based on their barrier tightness. The authors show that a triple-culture with pericytes as contact and astrocytes as non-contact culture gives the best TEER and that all co-culture models show increased TEER and decreased permeability of the marker sodium fluorescein compared to the mono-culture model. In a follow up study Nakagawa et al. (2009) characterize the triple-culture with pericytes as contact- and astrocytes as non-contact-culture. They identify the expression of several crucial transport systems and demonstrate a good correlation (R^2=0.89) between the *in vitro* and *in vivo* permeability of 19 compounds. The observation by Nakagawa et al. (2007) could not be reproduced in a similar study that compared a similar set ups of immortalized human cerebral microvascular endothelial cells (hCMEC/D3) co-cultured with human cerebral astrocytes and/or human brain vesicular pericytes (HBVP). Here the overall TEER was noticeably lower than observed in Nakagawa et al. (2007; 40-60 compared to 50-400 ohm/cm^2) and the triple-cultures did not show improved TEER compared to mono-culture (Hatherell *et al*, 2011).

The comparison of these two studies demonstrates some of the major limitations in comparing different BBB models and identification of a suitable model. For one, not all studies provide TEER values and permeability of tracer substances. However, a number of studies demonstrate that TEER values alone are not sufficient to determine the quality of the BBB (summarized in Wolff et al., 2015) which is why both values should be assessed and ideally accompanied with a correlation of permeability data from the BBB model and *in vivo* data from compounds with different properties related to their ability to enter the CNS. Another important aspect is species differences in BBB permeability. There are several studies that demonstrate different expression of transporters, tight junctions, and receptors between the BBB of rodents and primates

(Uchida *et al*, 2013, 2011; Shawahna *et al*, 2011; Ito *et al*, 2011; Hoshi *et al*, 2013). Positron emission tomography studies for example indicate differences in pharmacokinetics of P-gp substrates. Thereby the substrates [18F]-altanserin and [11C] -R205171 have a 4.5- and 8.6-fold greater penetration efficiency in human BBB compared to rodents (Syvänen *et al*, 2008). Another example is the tight junction protein claudin-5 which in human brain microvessels is two-fold lower expressed than in those of rats and other primates (Shawahna *et al*, 2011). These species differences ultimately cause differences in the paracellular and transporter/receptor based transport. It is therefore desirable to find suitable *in vitro* BBB models that are based on human cells.

A general limitation of human models is to acquire primary human material on a regular basis, which makes the establishment of BBB models based on primary human BEC difficult. For human primary material that is offered commercially, there is often only sparse documentation on the source of the material. Nevertheless some BBB models based on primary human brain endothelial cells are presented in the literature (reviewed in Wolff et al., 2015). A primary model of human BMECs co-cultured in a dynamic *in vitro* set up with human astrocytes demonstrated good TEER (200-700 ohm/cm^2) and good correlation with *in vivo* permeability data (R^2 = 0.93) of 7 compounds (Cucullo *et al*, 2011). This study additionally demonstrates that a dynamic system reaches a higher TEER, better correlation of compound permeability compared to *in vivo* data and increased tight junctions formation as well as the expression of multidrug resistance transporters.

Human based models that circumvent aforementioned limitation can be based on immortalized brain endothelial cells or pluripotent stem cell (PSC), (including induced PSC), - derived BECs. The most common, easy to use and best-characterized immortalized line are human cerebral microvascular endothelial cells (hCMEC/D3). In its basic state TEER values from these cells are not comparable with the *in vivo* situation or animal based models. This model therefore offers barrier function for only large molecules, whereas small molecules relatively easily permeate the barrier (Helms *et al*, 2016b). Although there are several articles that show limited barrier function of this cell model (Eigenmann *et al*, 2013; Hatherell *et al*, 2011) different culture condition for example dynamic culture system and/or co-culture with pericytes and/or astrocytes might improve the barrier function and make this a useful model for permeability screening. A study that compared four immortalized capillary endothelial cell lines (hCMEC/D3, hBMEC, TY10 and BB19) concluded that hBMEC proved to be the most suitable and promising immortalized cell for a human *in vitro* BBB model in terms of barrier tightness and paracellular permeability. The TEER values however were low compared to previously described models and no correlation of permeability data with *in vivo* data was demonstrated.

In contrast to immortalized cells, PSC- derived brain endothelial cells have been shown to form a good barrier function with expression of tight junction proteins, a TEER between 250 and up to 5000 ohm/cm^2 depending on the culture condition and sucrose permeability's comparable to porcine and bovine models and below values for primary and immortalized human models (reviewed in Helms et al., 2016). In a study by Lippmann et al. (2012) the authors present a BBB model based on human PSC- derived endothelial cells in non-contact co-culture with primary rat astrocytes. This model demonstrated well -organized tight junctions, appropriate expression of nutrient transporters and polarized efflux transporter activity. The TEER was measured to be around 1,450 ohm/cm^2 and permeability was in good correlation with *in vivo* data

(R^2 = 0.98) of a group of 7 compounds. The same authors demonstrated in a study from 2016 that retinoic acid treated iPSC derived BMEC that were co-cultured with primary human brain pericytes as well as astrocytes and neurons derived from human neural progenitor cells (NPCs) were able to form a BBB model with a TEER of up to 5,000 ohm/cm^2 and low sucrose permeability (Lippmann *et al*, 2015). The TEER however was very dependent on the type of iPSC that was used an could drop to below 500 ohm/cm^2. Other recent studies generated models based on adult human endothelial progenitor cells (Ponio *et al*, 2014) and hematopoietic stem/progenitor cells (Cecchelli *et al*, 2014). Especially the latter study generated a stable and reproducible human BBB model with good correlation with *in vivo* permeability data of 9 compounds (R^2=0.84). These studies demonstrate that the generation of BBB models based on PSC-derived BECs with superior barrier function compared to other human BBB models is in principal possible, making this an promising method for future permeability screening.

Until now, there is no model that can be recognized as the gold standard for predictable high-throughput screening application for BBB permeability. Nevertheless, several models have been proven useful. Primary BMECs isolated from animal brain tissue have shown to form tight barriers with low permeability and several models show good correlation with *in vivo* data (reviews in Helms et al., 2016; Wolff et al., 2015); however these comparisons are limited and the number of drugs tested very small (<10). The increasing knowledge on species differences lead to the general view, that human based systems are better suited for the prediction of BBB permeability *in vivo* (Aday *et al*, 2016). Primary human BMECs and immortalized human cells possess only moderate barrier properties (Helms *et al*, 2016b). These barrier properties can be improved when cells are co-cultured with astrocytes or pericytes or by the introduction of shear stress in a dynamic system. Recent research demonstrates that PSC- derived BMEC in co-culture with astrocytes, pericytes and/or neurons and under specific culture conditions form a very tight barrier that are able to produce TEER values comparable to *in vivo* data (Lippmann *et al*, 2015). Because of the short time period these cells are available they still lack some validation regarding functional expression of transporters, efflux pumps and receptors. There is also still some uncertainty in the reproducibility regarding culture condition, stability of the culture, culture set up and the type of PSC that is used. Nevertheless, PSC- derived BMEC are a promising model for chemical or drug permeability screening and need further characterization and development. In addition promising and established human based models should undergo validation of the transport and permeability properties using a large set of compounds and drugs with very different properties regarding *in vivo* BBB permeability and sufficient *in vivo* data.

3.3. In silico

3.3.1. Introduction:

Chemical substances can only exert neurotoxic effects if they are able to reach the brain. Before they can reach the brain, they have to cross the blood-brain barrier (BBB or BB). As previously described, the BBB is a very complex physical barrier that determines which substances are allowed to enter the brain. In this section computational (*in silico*) models for estimating BBB passage and for predicting neurotoxicity as such are discussed.
The selection procedure described in the methodology section resulted in 1365 publications. The title-abstract screening resulted in 169 publications for the full-text screening. The full-text

screening resulted in 54 publications. Details of the publications are presented in the Data Collection Sheet (DCS), rows 3 to 56. The publications are arranged in alphabetical order by first author's last name. The columns of the DCS sheet are structured according to the 5 OECD principles for Quantitative-Structure-Activity relationship (QSAR) models for regulatory use (OECD guidance No 69 on QSAR validation).
Additionally, websites offering QSAR models for hazard assessment were explored. Fourteen websites were found to contain *in silico* tools useful for NT evaluation.
A QSAR is a quantitative relationship between a biological activity (e.g., toxicity), which may be qualitative (yes/no) or quantitative, and one or more molecular descriptors that are used to predict the activity. A molecular descriptor is a structural or physicochemical property of a molecule, or part of a molecule, which specifies a particular characteristic of the molecule and is used as an independent variable in a QSAR (Worth et al., 2011)[3].

3.3.2. Models on blood-brain-barrier (BBB) permeation and neurotoxicity:

A total of 39 publications with in some cases more than one model per publication is the result of the full paper screening on BBB passage (Table 10). All models consider the passive diffusion; four of the papers include additional descriptors to also cover active transport. Most of the models are quantitative (calculate log BB = log ((C_{brain}/C_{blood})); some are qualitative, indicating whether the substance is able to reach the central nervous system (CNS) or not (CNS+/-, or BB+/-) with a defined cut-off value). The algorithms are developed by means of a wide range of machine learning methods, going from univariate and multivariate linear regression generation with a small number of parameters (e.g. Bujak *et al*, 2015) to the development of complex (non)-linear regressions with artificial neural networks (ANN) (e.g. Yan *et al*, 2013) and other machine learning methods such as support vector machine (SVM) (e.g. Golmohammadi *et al*, 2017), partial least square (PLS; (Cuadrado *et al*, 2007) and random forest analysis (RFA) (e.g. (Polishchuk *et al*, 2016). In the field of machine learning, the computer has the ability to "learn" from data (i.e. progressively improve performance on a specific task), without being explicitly programmed. The model developer chooses the number and kind of parameters (descriptors), the computer calculates their weights (coefficients) and the intercept. At the end an algorithm or final model (e.g. ANN), which is in fact a very complex algorithm with linear and non-linear correlations, is derived. These machine learning techniques can also be used to generate a qualitative outcome (BBB permeation: yes/no).
An overview of the compliance of the models with the five OECD criteria for QSAR-model evaluation is presented in Table 10. In general the models comply very well with the exception of the description of the applicability domain (OECD principle 3). For models without a published algorithm, the compliance with criterion 2, i.e. an unambiguous algorithm, is unclear. For two models algorithms are not present but they may be in the supplementary information. For complex models, it is not possible to publish the algorithm in a paper. Transparency on the exact training data and architecture of the method (e.g. neural network, number of layers and nodes in the ANN) (to 'reproduce' the model) or availability of the final model (represents the algorithm), and the assurance that the model is not overfitted (too many variables present, leading to a low validation performance) are important factors in view of the 2nd OECD principle

[3] Applicability of QSAR analysis in the evaluation of developmental and neurotoxicity effects for the assessment of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment. Scientific report submitted to EFSA. Final version 31/05/11

on QSAR model validation. A linear model is generally considered as unambiguous algorithm (OECD report 69 on QSAR validation).

Table 9: Compliance of the QSAR models with the five OECD criteria:

OECD Principle	Result (details available in the DCS sheet
1. Defining the end-point	BBB permeation: 39 models CNS membrane toxicity ($IC_{12.5}$ for membrane fluidity): 1 model Acute neurotoxicity: 6 models (EC_{30} effective concentration at 30% of the maximum possible effect) Acetylcholinesterase (AChE) and neurotoxic esterase inhibition: 2 models Neurotoxicity effect towards ryanodine receptors (neurons): 2 models Formation of Reactive Oxygen Species (ROS): 1 model Three biochemical endpoints for cytotoxicity: 1 observed SAR (not a model) Decrease in cell dopamine content: 1 observed SAR (not a model) CNS/hyperactivity: 1 observed SAR (not a model)
2. Defining the algorithm	Algorithm available: 27 publications Final models (complex algorithms): 18 models (the models or their architecture and training set should be available via the supplementary information of by other means, to fulfil the 2nd OECD criterion 'unambiguous algorithm')) Not applicable: SAR observations: 3 models (line 14, 48 and 52) 3D models, no algorithm: 3 models An algorithm could not be built: 1 publication (Stenberg et al. 2011) Not available: 2 models (maybe in the supplementary information but that is not indicated) (line 31, and 38) and 50)
	Descriptors: described from in detail to general, mostly calculated; Number of descriptors: from unknown to 72
3. Defining the applicability domain	Domain explicitly mentioned: 4/37 Chemical indication/description of the training set chemicals: all models
	Limits of applicability: 4 publications provide physical chemical property boundaries of the training set chemicals; the other publications define no limits of applicability. Predictions for substances outside the training set property boundaries may be less accurate.
4. Internal/external validation	Yes for all models Coefficient of determination (R^2 or r^2) range: 0,54-0,99 Cross validation correlation coefficient(Q^2): 0,50-0,98 Accuracy range: 73-97%(where available)
5. Mechanistic basis	BBB permeation: passive diffusion; 4 models additionally include active transport for which the mechanistic basis is protein binding activity

	Ryanodine receptor (RyR) mediated mechanisms and associated neuronal signalling: influenced by altered calcium regulation

The fitness of an algorithm or model is measured as the coefficient of determination (R^2 or r^2) and internal cross-validation coefficient (Q^2). Commonly applied criteria for reliability and predictability are $R^2 > 0.6$ and $Q^2 > 0.5$ which were set by Tropsha *et al*, (2003). Other authors considered a QSAR model to have an acceptable predictive power if $q^2 > 0.65$ and $R^2 > 0.65$. The R^2 ranges from 0,54 to 0,99 (where available) and Q^2 ranges from 0,50 to 0,98 (where available). Most of the models are above the 0.65 limit. Moderate q^2 values may indicate that the model might only perform marginally in predicting the logBB of compounds outside the training set.
The accuracy, which is the ratio of correctly predicted positives and negatives to the total number of predicted compounds, gives an idea about the predictivity of the model. The accuracy of the studied models varies from 73 to 97% (where available).

A significant disadvantage of an ANN is the amount of training data needed. Experimental data on BBB or neurotoxicity are scarce. This may contribute to the explanation as to why the performance of an ANN model is not always better than the performance of a simple linear regression.

Two properties often returning as important factors linked to passive BBB diffusion are lipophilicity (LogP) and charge (total polar surface area). It has been appreciated in drug design for many years that drugs with logP values of 1.5-3.0 seem to have optimal abilities to diffuse through biological membranes (Burns & Weaver, 2004). The total polar surface area of a molecule however, is negatively correlated with passive diffusion. Also heteroatoms seem to have a negative impact on diffusion.

Note: Two authors developed a BBB-model from dynamic membrane computer simulations. Four models consider, besides passive passage, also active transport such as by the efflux pump P-glycoprotein (P-gp) (Chen *et al*, 2009; Garg & Verma, 2006; Iyer *et al*, 2002 and Suenderhauf *et al*, 2012) Whether a substance is actively transported through the BBB depends on the properties of that substance, but may also be influenced by the presence of another substance that acts as transport inhibitors. One model, trained with alcohols, estimates the impact on Na+/K+-ATPase and AchE activity (indicators of CNS membrane fluidity). A linear relationship between membrane fluidity and the octanol-water partitioning coefficient (log P) is found, indicating a significant effect of the alcohols on membrane fluidity. Based on these results, the authors suggest that the alcohols inhibit the Na+/K+-ATPase and AchE activity through a direct toxic action on the enzymes and/or through changing the membrane fluidity (El Yazal *et al*, 2001). This research touches the issue of mixture toxicity: active transport of chemicals into the CNS may be influenced by the presence of chemicals inhibiting membrane fluidity directly or indirectly.

3.3.3. Grey literature search:

Based on the gathered information, all websites of the first screening were retained in the final selection, with the exception of ToxMatch and Accelrys; ToxMatch does not cover the neurotoxicity endpoint or BBB permeation, and Accelrys offers a modelling environment, where the user can store his own developed models.
The finally collected models and tools are presented in the DCS sheet rows 56 to 69. The modelled endpoints are BBB permeation and neurotoxicity. Furthermore tools for read a cross were identified and one database that is searchable for *in silico*- and experimental results on neurotoxicity was found: the e-Chem portal developed by the OECD (row 70 in the DCS). The result of the *in silico* grey literature search is as follows:

- BBB permeation: 7 models
- Neurotoxicity: 3 models
- Read across tools: 3
- Database: 1.

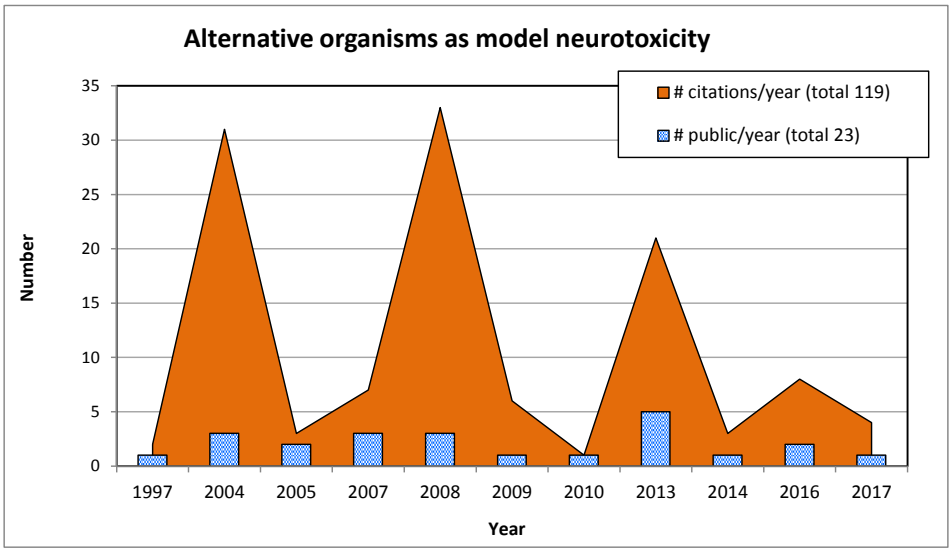
More details and the links to the models/tools are presented in the DCS sheet, rows 56 to 70 (first column: grey lit 1 to 15).
Some models have been developed to serve new-drug designers, although the websites do not state that the models are restricted to predictions for drugs. Other models explicitly state that they cover 'any' chemical.
It would be interesting to make a comparative study on the prediction of neurotoxicity or BBB permeation of pesticides and other chemical substances by these website-models and the best performing algorithms.

3.4. Alternative organisms

For 140 publications on methods with alternative organisms for neurotoxicity of chemicals with known MoA during the years 1990-2017, full texts were screened and relevant information in line with inclusion criteria was collected in a data collection sheet (Appendix H). For those retained for further analyses, a quality sheet was also completed to assess the reliability. However, a majority of publications was excluded based on main criteria such as 1) a study on DNT, 2) a study on neuroprotection (often in combination with the 3rd exclusion criterion), 3) only 1 single concentration of compound tested or 4) life stage was not an alternative stage. The latter exclusion criterion was only applicable for the zebrafish model, as often neurotoxicity studies were performed with juvenile or adult life stages which according to EU Directive are considered experimental animals.

Only 23 different publications representing alternative whole organism models with the fruit fly *Drosophila melanogaster* (n=7), the nematode worm *Caenorhabditis elegans* (n=12) or the zebrafish *Danio rerio* (n=4) were retained for further analysis in multiple ways. Each of the publications can cover multiple biological endpoints for neurotoxicity and studying one or more different test compounds. The latter resulted in a total of 119 citations (or hits) corresponding to these 23 papers. The time course for these publications and citations is presented in Figure 20.

a)



b)

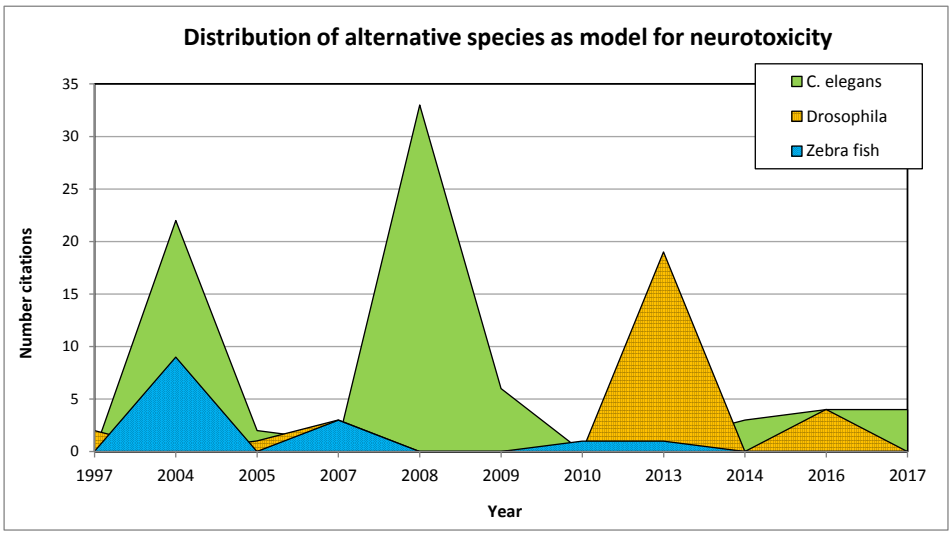


Figure 20 a) number of publications, covering multiple citations by endpoints and compounds for methods with alternative organisms as a function of year and b) number of citations with distribution for 3 alternative species across the past 20 years.

The analysis of total number of publications and of citations over time (Figure 20) did not show a real trend, but overall the numbers were rather low to find patterns. With respect to the number of citations there is high variability from year to year, but the high number of citations e.g. for 2004, 2008 and 2013 is rather due to few publications where multiple endpoints or compounds were simultaneously studied (Cole *et al*, 2004b; Rajini *et al*, 2008; Hosamani, 2013).

The graph on contribution of the different alternative model species (Figure 20b) clearly demonstrates that the nematode worm, *C. elegans* is used most based on the number of citations throughout the past 20 years.

The review clearly showed that *C. elegans* is most important alternative organisms for neurotoxicity studies, as can be derived from the number of publications which represent 52% (n= 12 publications), and the number of citations which goes up to 64% (n=76). This is followed by *Drosophila* as 2nd most important organism (7 publications, 24 % of citations) and least use of the zebrafish model (4 publications, 12% of citations) as alternative for neurotoxicity (Figure 21 a & b).

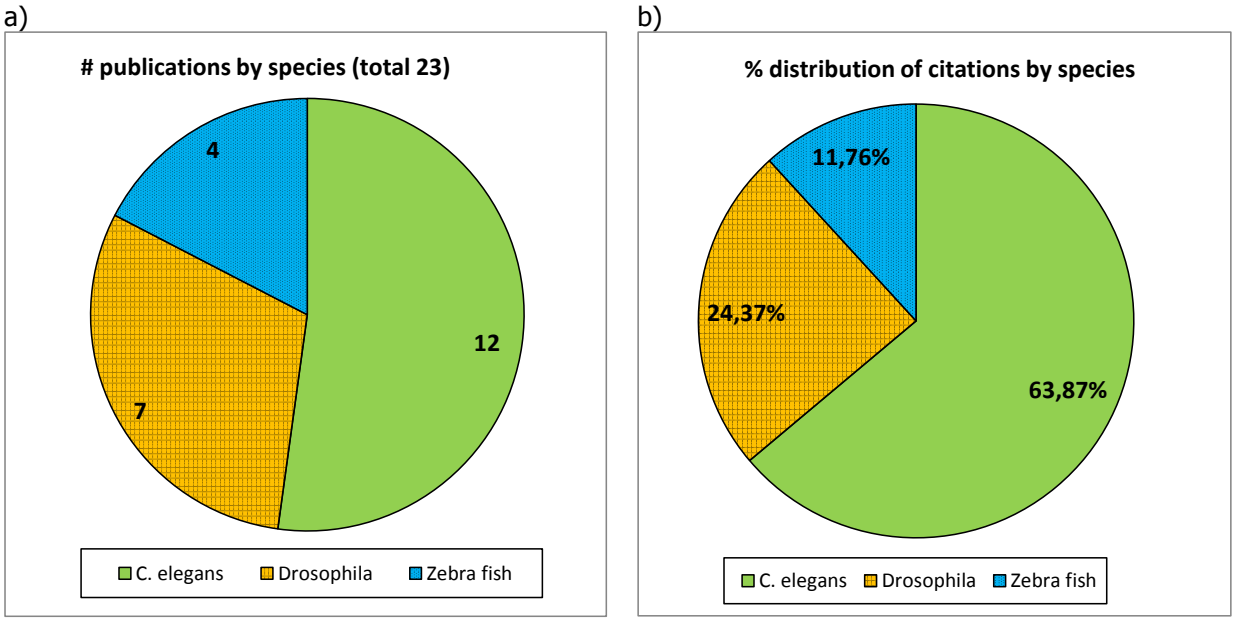


Figure 21 a) number of publications for alternative whole organism approaches grouped by species and b) distribution of citations by species as a % of total.

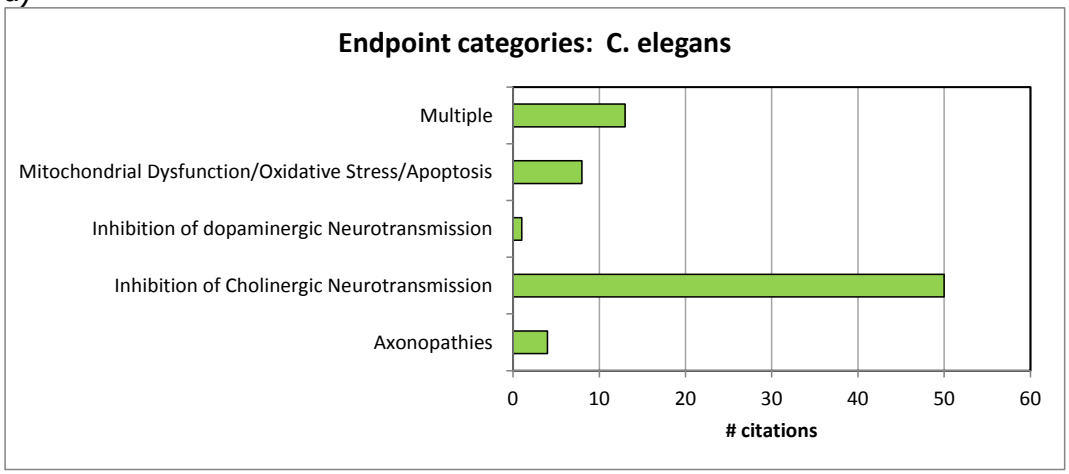
These data are then evaluated with regard to their performance within species group for certain MoA represented by endpoint categories as described in the methods section. A comprehensive list of the identified MoA for the list of neurotoxic compounds in relation to endpoint categories is available in Appendix D.

This further evaluation will only make use of individual citations (single rows in DCS sheet), grouped by species addressing endpoint categories. For endpoint categories, the list of specific endpoints studied will be listed as can be derived from the DCS for each of the species studied for selected neurotoxic compounds. Questions to be addressed within the data evaluation are: 1) which endpoint categories and specific endpoints are represented by each of the alternative model species, and 2) what is the predictive capacity of each alternative organism for each endpoint category considering the identified MOA and for each of the selected neurotoxic chemicals in the compound list?

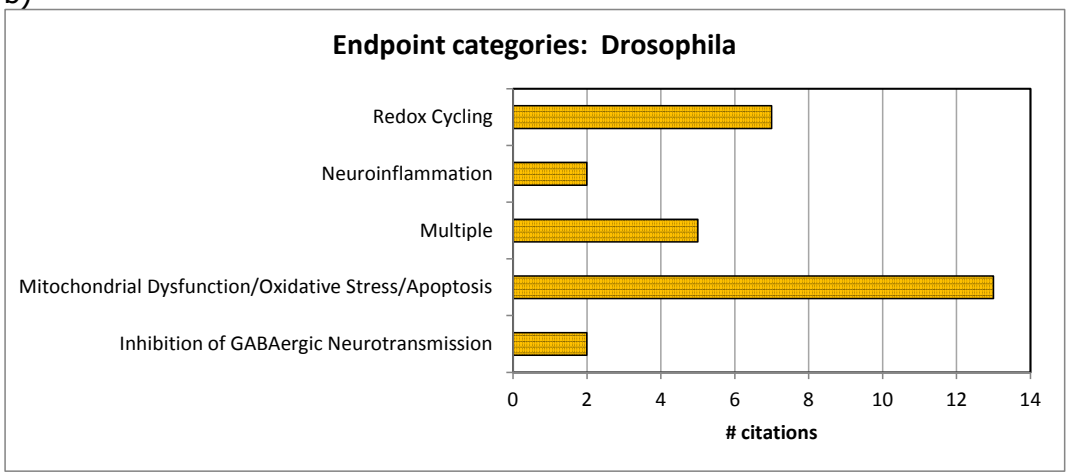
When evaluating the publications, only those citations that were in agreement with the identified MoA (1 or more) of the selected neurotoxic compounds were included in the DCS leading to the 119 citations. This procedure guarantees that an effect observed in model organism is a 'true' effect and allows the categorization into 'true positives' or 'false negatives'. In addition, negative compounds (though exceptional for alternative organism) which were identified in the studies have been classified into 'true negatives' or 'false positives'.

In the following graphs (Figure 22 a,b,c), the major endpoint categories studied for respectively *C. elegans*, *Drosophila* and zebrafish are shown. For each of the species, a difference is seen with respect to the main endpoint categories studied. It is not clear from the current review with limited number of publications whether these differences are rather occasional and biased due to the low number or whether there is a biological or technological background with respect to e.g. availability of methods to measure the endpoints.

a)



b)



c)

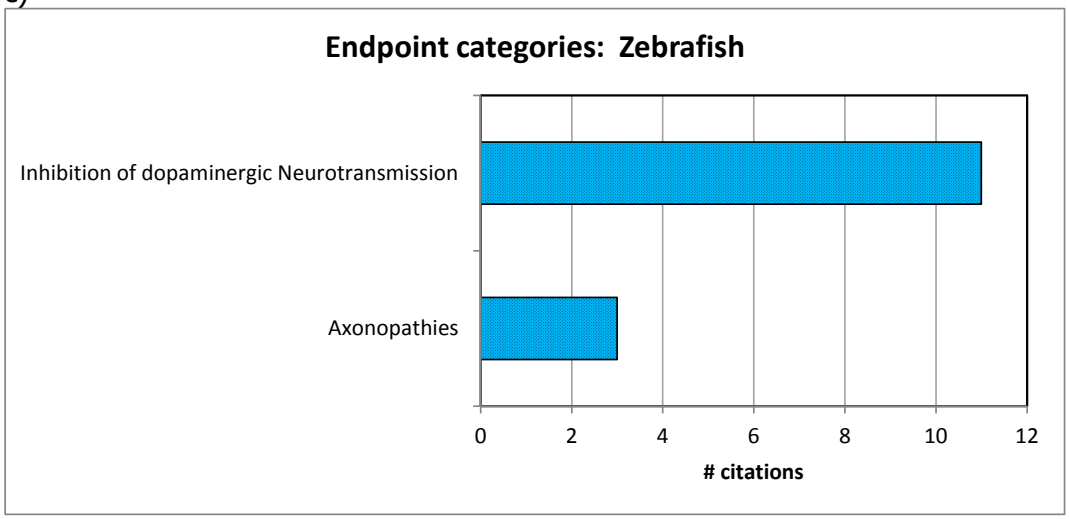


Figure 22 Distribution of # citations across endpoint categories within the different species, respectively a) *C. elegans*, b) Drosophila, and c) zebrafish. Endpoint categories were derived from the compounds' MoAs.

More detail on endpoints within categories and their predictive capacity using *C. elegans* is summarized in the next table (Table 10) and an overview for the compounds tested is presented in a bar chart (Figure 23).

The major endpoint category studied with *C. elegans* is inhibition of cholinergic transmission (48 citations), which is represented by specific endpoints as measurements of AChE activity (15 citations) and motor activity (22 citations). Especially the latter endpoint confirms the added value of a whole organism approach which is lacking in studies using *in vitro* assays (cell culture models). The predictive capacity for this endpoint category is high with 96% true positive results in *C. elegans*. Only for 1 compound (glyphosate), for both motor activity and AChE inhibition a false negative score was obtained. Effects on motor activity were dedicated to low pH of test solutions at high glyphosate levels, which did not occur after pH correction, and no effect on AChE activity was observed (Cole *et al*, 2004b). The 2nd most important endpoint category for *C. elegans* was 'multiple'. This category was identified throughout the analysis of methods for those cases where an endpoint (e.g. motor activity disturbance in alternative organisms), might be the result of different MOAs, which relate to more than 1, thus multiple, endpoint categories for some compounds. For this category, a high true positive score of 83.3 % was obtained for the main endpoint 'motor activity' (10/12 citations). Copper chloride and mebendazole were both considered as negative compounds by the authors for this assay. The results showed no effect for mebendazole for *C. elegans* motor activity and thus confirmed as true negative. On the other hand, disturbance of motor activity (locomotion and feeding behavior) was observed after exposure to copper chloride which scored false positive. The effects were however less than other known neurotoxic metals (Pb and Al).

Within the 3 other endpoint categories, respectively mitochondrial dysfunction & oxidative stress (8 citations), axonpathies (4 citations) and inhibition of dopaminergic neurotransmission all assays resulted into 100% true positive results.

The prediction of *C. elegans* assays for a diversity of compounds is presented as a bar chart for individual compounds tested (Figure 23). For a total of 76 citations, representing 27 compounds of different chemical classes, 96% true prediction was obtained for 72 positive neurotoxic compounds and 1 negative compound.

Table 10: Total number of citations for endpoint categories and specific endpoints, grouped for *C. elegans*. Predictivity analyses were performed for each endpoint category and specific endpoint by analyzing true positives, false positives, true negatives and false negatives.

C. elegans **Endpoint category (and specific endpoints)**	false negative	false positive	true negative	true positive
Axonopathies				4
Aggregation of α-synuclein protein				1
Learning ability				1
Motor activity				1
Sensory function				1
Inhibition of Cholinergic Neurotransmission	2			48
AChE activity	1			15
Activity feeding behaviour				1
Motor activity	1			22
Viability				10
Inhibition of dopaminergic Neurotransmission				1
Localisation of dopaminergic neurons				1
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				8
Mitochondrial Membrane Potential				1
Mitochondrial structure				1
Motor activity				3
Neurodegeneration				1
Oxygen consumption rate				1
RNAi of respiratory chain genes				1
Multiple		1	1	11
Metabolomics				1
Motor activity		1	1	10
Grand Total	**2**	**1**	**1**	**72**

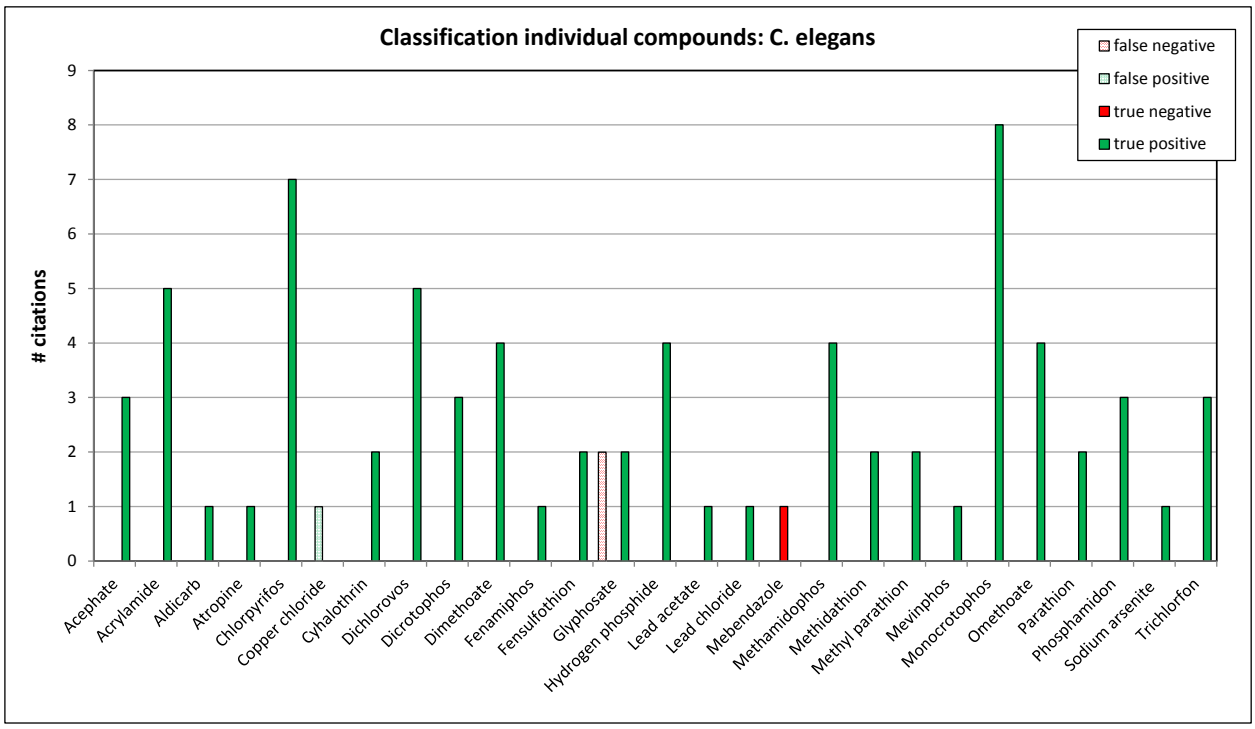


Figure 23: Total number of citations for each of the test compounds evaluated for *C. elegans* assays. Predictivity analyses were performed for each compound, considering different endpoints resulting into true positives, false positives, true negatives and false negatives.

More detail on endpoints within categories and their predictive capacity using *Drosophila* from 8 different publications with 29 citations is summarized in the next table (Table 11) and an overview for the compounds tested (n=4) is presented in a bar chart (Figure 24).
The major endpoint category studied with *Drosophila* is mitochondrial dysfunction/oxidative stress/apoptosis (13 citations), which is represented by 13 different specific endpoints (Figure 22; Table 11) with 100% true positive results. Also the endpoint category redox cycling, with 7 citations, resulted in 86% true positive results. Only 1 citation for endpoint oxidative stress (catalase assay) in *Drosophila* exposed to paraquat gave a false negative result among true positive results for other endpoints as motor activity (geotaxis) and morphological damage to dopaminergic neurons (Chaudhuri *et al*, 2007).

Except for 26 citations with paraquat, only 3 other neurotoxic compounds were tested with *Drosophila* assays, but all these resulted into true positives (Figure 24). No negative compounds were evaluated. Overall, the *Drosophila* assays across all endpoints and compounds resulted into 96% true prediction.

Table 11: Total number of citations for endpoint categories and specific endpoints, grouped for *Drosophila*. Predictivity analyses were performed for each endpoint category and specific endpoint by analyzing true positives, and false negatives.

Drosophila **Endpoint category (and specific endpoints)**	false negative	true positive
Inhibition of GABAergic Neurotransmission		2
Nerve firing		2
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis		13
AChE activity		1
Activity of anti-oxydant enzymes		1
Assay for Citric Acid Cycle Enzymes		1
BuChE activity		1
Electron Transport Enzymes: Complex I-III and Complex II-III Activity		1
Hydroperoxide generation		1
Lipid peroxidation		1
Mg2+ ATPase Activity		1
Mitochondrial Membrane Potential		1
Reduced GSH content		1
ROS generation		1
Superoxide generation		1
Tissue Iron Levels		1
Multiple		5
DNA damage in brain		1
Histopathological brain damage		1
Motor activity		2
Oxidative stress		1
Neuroinflammation		2
Oxidative stress		2
Redox Cycling	1	6
Metabolomics		1
Morphology dopaminergic neurons		1
Motor activity		2
Number of DA neurons		1
Oxidative stress	1	1
Grand Total	**1**	**28**

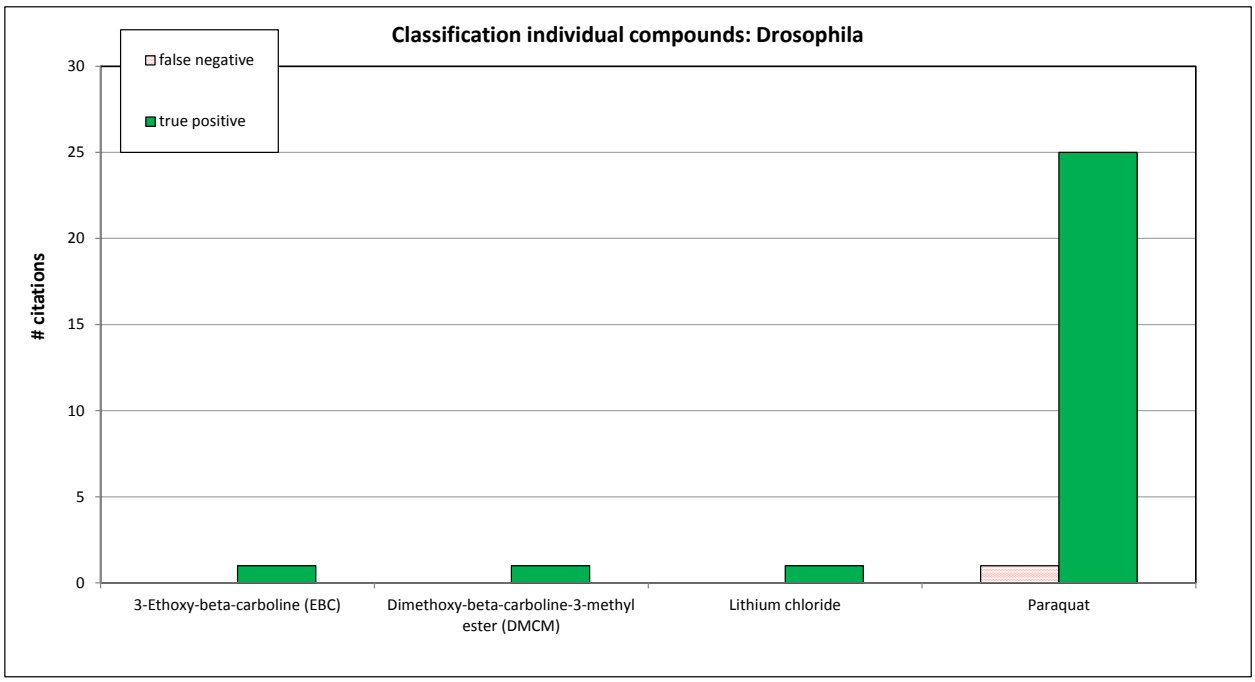


Figure 24 Total number of citations for each of the test compounds evaluated for *Drosophila* assays. Predictivity analyses were performed for each compound, considering different endpoints resulting into true positives and false negatives.

For the limited number of different publications (4), for 6 positive compounds tested resulting into 14 citations for the zebrafish model, the data analysis on endpoints within categories and their predictive capacity is presented in the next table (Table 12) and an overview for the compounds tested is presented in a bar chart (Figure 25).

The major endpoint category (Figure 22; Table 12) studied with zebrafish was inhibition of dopaminergic transmission (11 citations), with 3 specific endpoints (motor activity, morphology and tyrosine hydroxylase labeling of dopaminergic neurons). The majority of citations were part of only 1 publication with the study of 3 compounds, respectively paraquat, rotenone and MPTP (Bretaud *et al*, 2004). Except for the endpoint motor activity and tyrosine hydroxylase labeling for MPTP, all other endpoints and compounds resulted into false negative scores. So far there is no good explanation for this discrepancy.
On the other hand, for 5 citations in 3 other papers, the endpoint motor activity and 3 different endpoints covering the category axonopathies resulted in 100% true positive results. No negative compounds were evaluated for the zebrafish assays. Overall, the zebrafish assays across all endpoints and compounds resulted into 50% true prediction, and 50% false prediction.

Table 12: Total number of citations for endpoint categories and specific endpoints, grouped for zebrafish. Predictivity analyses were performed for each endpoint category and specific endpoint by analyzing true positives, and false negatives.

Zebrafish **Endpoint category (and specific endpoints)**	false negative	true positive
Axonopathies		3
Apoptosis		1
Myelin basic protein (MBP)		1
Structure of myelin fibers		1
Inhibition of dopaminergic Neurotransmission	7	4
Morphology dopaminergic neurons	3	
Motor activity	2	3
Tyrosine hydroxylase labeling for Dopaminergic neurons	2	1
Grand Total	**7**	**7**

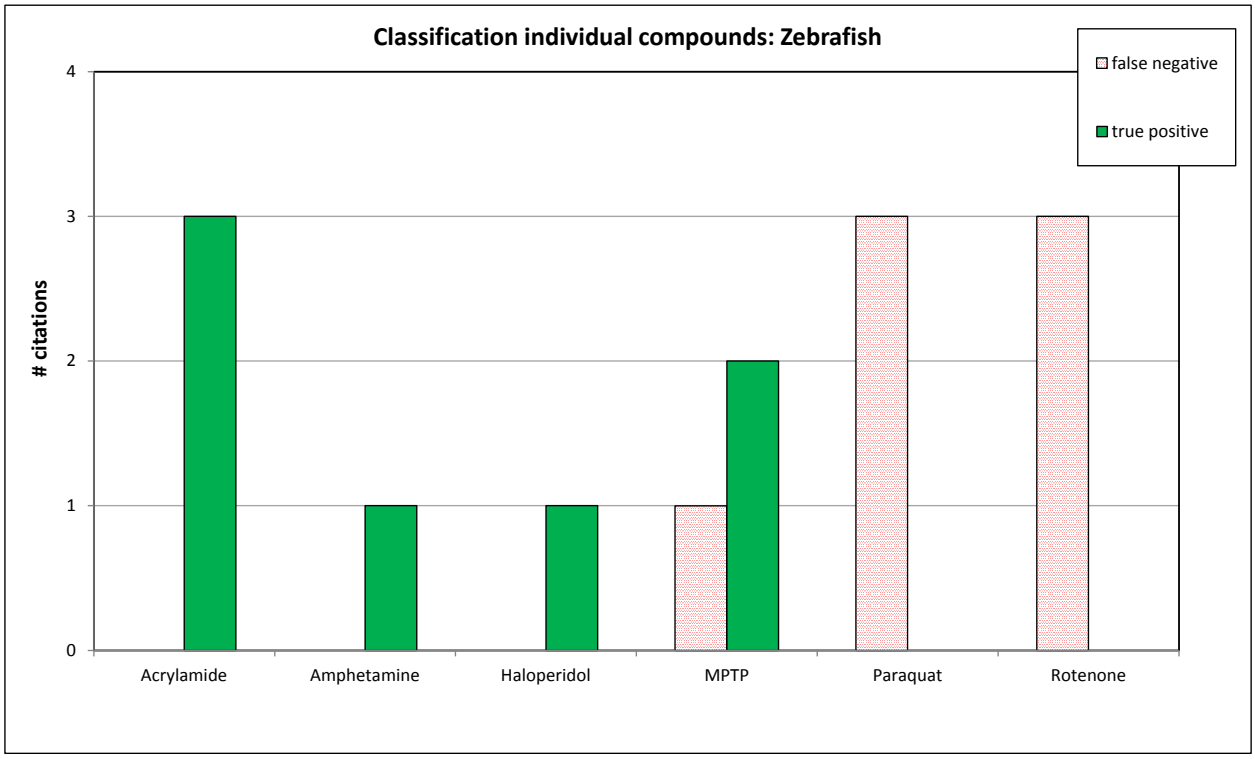


Figure 25 Total number of citations for each of the test compounds evaluated for zebrafish assays. Predictivity analyses were performed for each compound, considering different endpoints resulting into true positives and false negatives.

Summary alternative organisms related to MOA of neurotoxic compounds

The review resulted in a rather low number of publications, in total 23, representing 3 non-mammalian whole organism models which were retained for the analysis of methods covering a diversity of endpoint categories related to identified MOA of neurotoxic compounds. The latter allowed evaluating methods using *C. elegans*, *Drosophila* or zebrafish covering 119 citations.

We can summarize the results of the review through comparison of performance for each of the model systems to predict mode of action of compounds, based on endpoint categories tested (Figure 26).

The majority of data were available for *C. elegans* with 76 citations, representing 27 compounds (25 positive, 2 negative). The main endpoint category was inhibition of cholinergic transmission, with specific endpoints for AChE activity and motor activity. The latter endpoint confirms the need for the use of a whole organism approach. Often automated video tracking methods are used which allow medium to high-throughput assessment. The overall true prediction using the *C. elegans* was very high (96%) for a large group of compounds of different chemical classes (Figure 26a).

For the *Drosophila* model, 7 publications with 29 citations were evaluated with majority dedicated to paraquat (26 citations). The major endpoint category studied in the fly was mitochondrial dysfunction/oxidative stress/apoptosis. These and other endpoints studied, resulted in 97 % true prediction, but only 4 compounds, positive for neurotoxicity were part of the studies (Figure 26b)

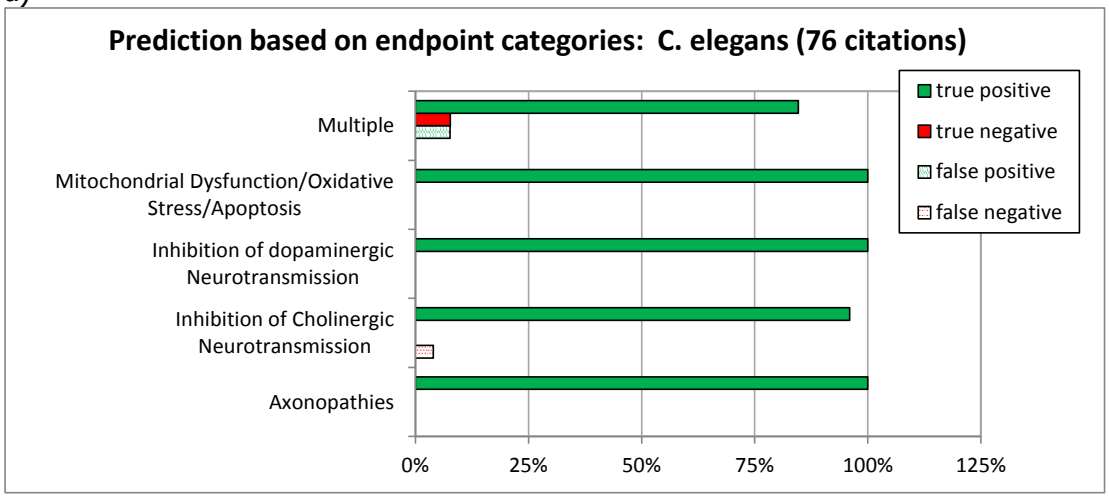
The zebrafish model represented the lowest number of publications (4) and citations (14) for 6 chemical compounds, classified as positive for neurotoxicity. Predictive capacity was least of the 3 alternative species with only 50% true prediction and 50% false prediction (Figure 26c). The latter bad result is based on tests performed within only 1 publication.

Overall, we can conclude that the whole organism approaches using *C. elegans* and *Drosophila* both provide high true prediction of MOA (96-97%). Most data resulted from studies with the nematode worm which seems widely used, applied for a diversity of numerous compounds and a lot of protocols are available both for mechanistic studies (molecular and cellular events) as well as medium to high-throughput methods for effect assessment, or adverse outcome such as motor activity. New developments related to lab-on-chip approaches and standardization of protocols might be needed to consider this assay as a low cost, and fast screening tool within an IATA for neurotoxicity assessment. As the *C. elegans* nervous system is a simplified system, further needs for research might go to the zebrafish model, representing a vertebrate organism with known high gene conservation compared to mammals. It appeared in this review that this model has not yet been sufficiently studied as an alternative model organism for neurotoxicity, though promising assays for developmental neurotoxicity and assays with juvenile and adult fish exist. Beyond the period of embryo development (2-3 dpf), when major brain structure and cellular features are present (Nishimura *et al*, 2015), exposure studies to neurotoxic compounds for many more endpoints are required to demonstrate the validity of this model. However, exposure and endpoint analysis should be finished within 5-6dpf when the model is not yet considered an experimental model. The whole animal approaches, which offer advantages to study complex processes such as molecular, cellular and signaling events in the CNS and PNS need

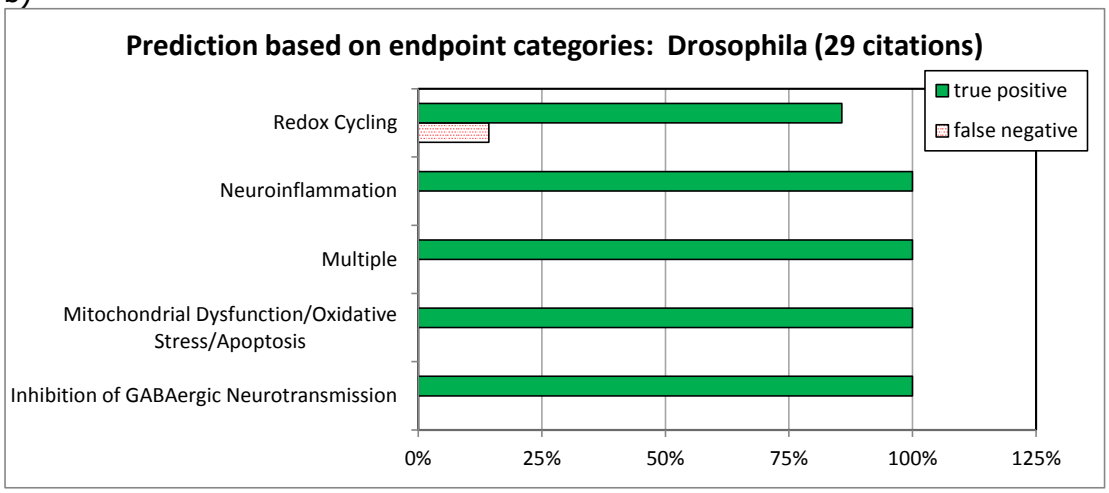
further consideration to fill gaps identified using *in silico* & vitro approaches, and enhance *in vivo* predictions.

Another recommendation, applicable to any of test methods (ao, *in vitro*,..) to be evaluated refers to the need to include also negative compounds for performance testing and assessment of prediction.

a)



b)



c)

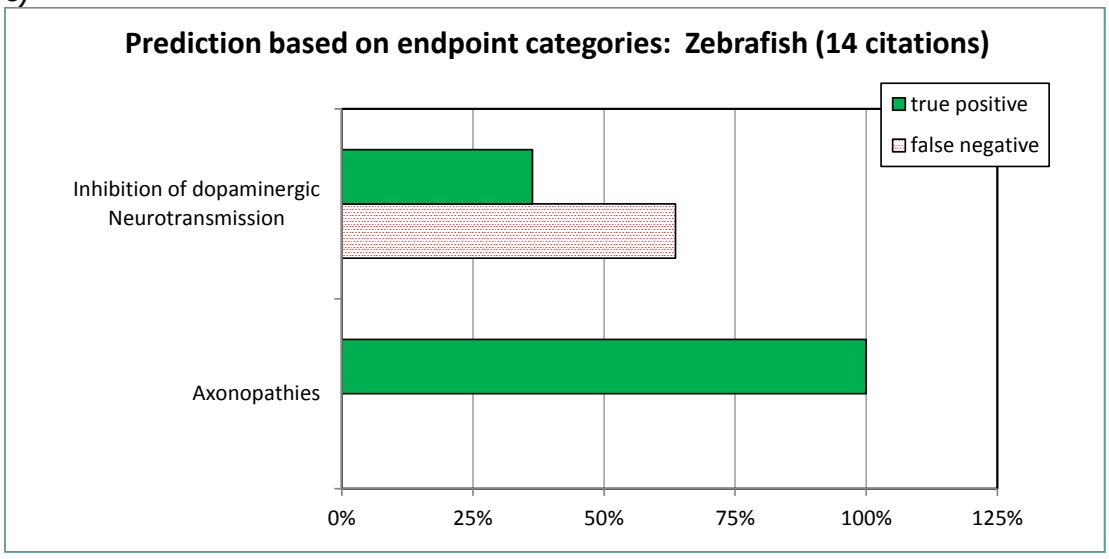


Figure 26 summary of results, based on endpoint categories and % prediction by analyzing true positives, true negatives, false positives and false negatives across compounds tested for each of the alternative model species, respectively a) *C. elegans*, b) *Drosophila*, and c) zebrafish.

4. Readiness analyses

A readiness check of testing methods was performed according to the procedure described in Bal-Price *et al*, (2018). Here, 13 criteria for evaluation are sorted into three phases. Each area has various sub items and the number of points that can be obtained is indicated in Appendix J. Phase I (green) includes the basic features of the test method as they would be provided by academic researchers. They include biological plausibility of the test method, features of the test system, and the availability of controls. A high number of points can be obtained for test system description (10 out of 35), as this is very important at early stages of test development. However, still two thirds of the points come from other areas not to be neglected. The second phase (blue) relates to the implementation of a test for practical applications in industry or for regulatory purposes. Here, the relation to a testing strategy, good robustness, and the availability of a prediction model are important. The third phase (yellow) is optional as not each test method is used for a screening approach. Notably, not all points apply to all tests. In the preliminary rating scheme suggested here, these items are then scored positive automatically. Each phase is evaluated independently, and then categorized into one of four readiness classes (A-D). In the figure, an example is given for the rating of the cMINC (UKN2) test method. It would score as 'A' (largely ready) in phase I, and as 'B' in phase II. For phase III, it would score as 'A' (Figure 27).

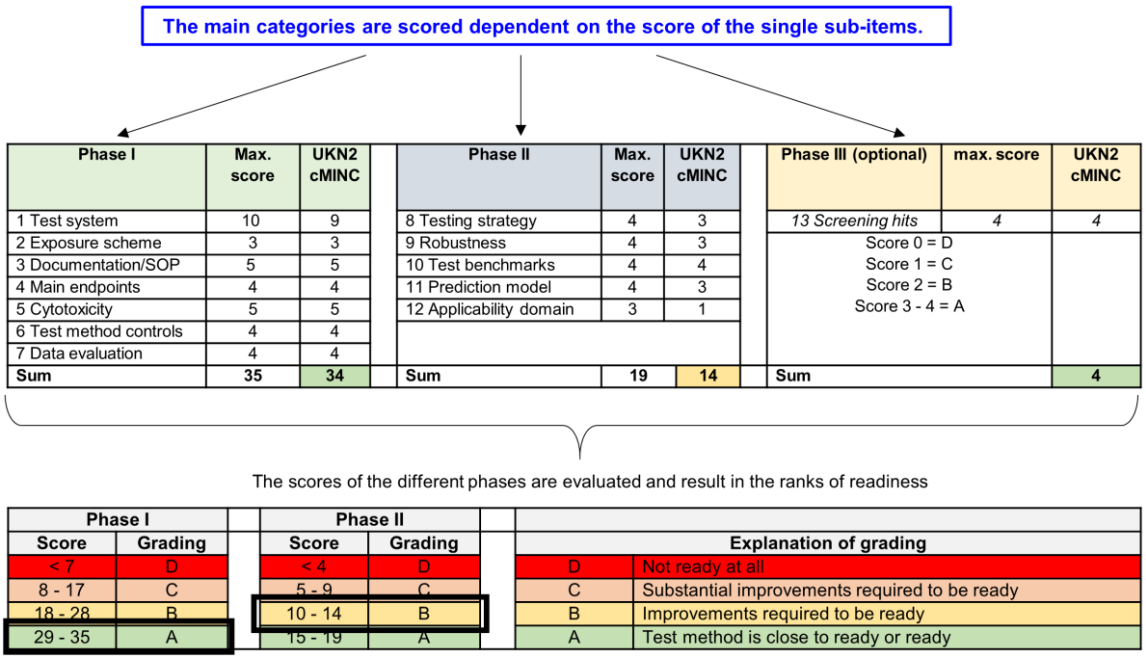


Figure 27 Set-up and example of a test system readiness check as published in Bal-Price et al. 2018.

During the review of the studies, 8 publications were marked as applicable to a readiness check (Smirnova *et al*, 2016; Zagoura *et al*, 2017a; Cookson & Pentreath, 1994b; Krug *et al*, 2013; Zurich *et al*, 2013; van Vliet *et al*, 2007; Nicolas *et al*, 2014; Sirenko *et al*, 2014). Scoring of these studies was performed across phase I, II and III as indicated above. The results of this scoring are shown in Table 13 and the detailed score results in Appendix K. These analyses show that there is not a high readiness level within the test systems evaluated. The highest score for phase I is a 'B', indicating that there is room for improvement even on the basic scientific description level of the methods. Also scores for phase II across all evaluations do not exceed 'B', most studies scoring 'C' or 'D'. The best category was phase III concerning screening. Here, most studies reached an 'A'. Overall, the study with the highest scores (Zurich et al. 2013) are 3D primary rat cultures, followed by studies using 2D (Krug et al. 2013) or 3D (Smirnova et al. 2016) immortalized human LUHMES cells, and a hiPSC model (Zagoura et al. 2016).

Table 13: Scores for readiness of in vitro test systems from 8 publications.

1st author	Smirnova	Zagoura	Cookson	Krug	Zurich	van Vliet	Nicolas	Sirenko
Year	2016	2016	1994	2013	2013	2007	2014	2014
Phase I -	21	20	17	23	20	11	10,5	12,5

score								
Phase I - grade	**B**	**B**	**C**	**B**	**B**	**C**	**C**	**C**
Phase II - score	6	8	9	9	11	4	5	2
Phase II - grade	**C**	**C**	**C**	**C**	**B**	**D**	**D**	**D**
Phase III - score	4	3	3	4	3	0	0	0
Phase III - grade	**A**	**A**	**A**	**A**	**A**	**D**	**D**	**D**

For the review of studies using alternative organisms, 3 publications were retained to be suitable for a readiness check (Shukla *et al*, 2014; Cole *et al*, 2004a; Irons *et al*, 2010). Scoring of these studies was performed across phase I, II and III as indicated above. The results of this scoring are shown in Table 14 and the detailed score results in Appendix K. Some (sub)criteria, as defined based on in vitro methods, were not considered relevant for the organism approach (NR), and got the maximum score applicable (italic in appendix).

The highest score for each of the papers in phase I is a 'B', indicating that similar to the in vitro methods, there is need to improve the methods at different levels (e.g. use of positive and negative controls, technical documentation). Also scores for phase II are low, with 'B' for the nematode model, and even 'C' for the 2 other publications. There is general lack of information on the prediction model, test benchmarks, robustness testing and application domain. The phase III scoring was little bit better with 2 studies reaching the 'A' score. The overall score for the *C. elegans* models (Cole *et al*, 2004a) was the highest, followed by the fruit fly model (Shukla *et al*, 2014) and least for the zebrafish model (Irons *et al*, 2010). In the latter study, the linkage of the relevant neurotoxic adverse outcome measurements (disturbed behaviour) to changes of early key events as part of AOP should have been of added value and needs further investigation (see also low score for criteria 4, 8 & 12).

Table 14: Scores for readiness of alternative organisms based test systems from 3 publications.

1st author	Shukla et al.	Cole et al.	Irons et al.
Year	2014	2004	2010
Species ao	Fruit fly	Nematode worm	Zebrafish
Phase I - score	22	23	19
Phase I - grade	**B**	**B**	**B**
Phase II -	9	11	6

score			
Phase II - grade	**C**	**B**	**C**
Phase III - score	3	3	2
Phase III - grade	**A**	**A**	**B**

5. Summary

Testing for neurotoxicity of compounds including plant protection products is performed in rats according to the OECD Guideline 424 (Neurotoxicity studies in rodents) and 426 (Developmental Neurotoxicity Studies). However, both these methods use complex *in vivo* tests, which are often too laborious and expensive and might also not well reflect the human situation because of inter-species variation. It is now recognised that the future of chemical safety assessment must move away from animal tests towards a combination of complementary approaches that address functional mechanistic endpoints tied to adverse outcomes of regulatory concern. This does not only concern the EU, but also the US has just released a draft strategy to reduce the use of vertebrate animals in chemical testing. This Toxic Substances Control Act (TSCA) requires EPA to develop a Strategic Plan by mid 2018 to promote the development and implementation of alternative test methods and strategies to reduce, refine or replace vertebrate animal testing (https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/alternative-test-methods-and-strategies-reduce). On this background, this systematic review was performed under a contract with EFSA to evaluate information on assessment methods in the field of neurotoxicity (NT). Therefore, a systematic and comprehensive literature search and collection of scientific literature and all other relevant grey literature and website information (in English) from 1990 until 2017 on the state of the art NT testing methods including *in vitro* test methods, novel and alternative non-mammalian models and *in silico* methods was performed. In addition, mode-of-action (MoA) analyses for 248 individual neurotoxic compounds, 23 compound classes and 212 natural neurotoxins were carried out and collected. While analysing the publications, only these studies were included that related to the identified MoA(s) of the respective compounds. In this way, only studies with known physiological relevance were included in the final evaluation strongly increasing the meaningfulness of the data. However, there still were endpoints detected that did not show an anticipated effect or negative compounds having an effect. This produced false negative and false positive data. These were analysed in more detail and the true positives were used for defining the assays' application domains.

The analyses of the *in vitro* studies revealed:

- From 9088 studies after the first search, 1803 were included for full text screening. From those we included 209 for *in vitro*, 39 for *in silico* and 23 for alternative organisms.
- The main studies of the publication selection for *in vitro* studies were performed with rat cells, followed by mouse, human and only few studies were included working with chicken and xenopus cells.

- These contain primary cells, primary tissue, stem-/progenitor cells, immortalized cells or cell free methods.
- Primary cells were generated from different brain regions with rat primary cells prepared from the largest number of brain regions followed by mouse and chicken.
- Stem-/progenitor cell-based models were mainly from human the human species, here predominantly hiPSC.
- According to the MoA analyses, 27 endpoint categories were defined that contain grouped endpoints and reflect key events (KE) of neurotoxicity.
- In all *in vitro* models, 'mitochondrial dysfunction/oxidative stress/apoptosis' was the most studied endpoint
- No test system has so far been shown to study all endpoint categories. However, multiplexing of endpoints within one model is advisable.
- Multiple especially primary cells in combination are able to cover a wide spectrum of endpoint categories.
- With regards to human stem-/progenitor cell methods, promising cell systems are on the way, yet there is a lot of data missing regarding their ability to detect the endpoint categories that can be studied in rodent primary cells.
- Special attention has to be given to glia cells. These are understudied cells when it comes to hiPSC. Here, oligodendrocyte and microglia data are even more sparse than data on astrocytes.
- More glia models and neuron-glia co-culture models are needed to multiplex endpoint evaluations in one system that contains interactions between the different cell types, neurons and glia.
- BBB models need further development.
- Protocol harmonization and definition of culture and quality standards are necessary to reproducibly produce defined test systems based on hiPSC.
- This systematic review report now contains usable data on cell methods and application domains that can be used for assembly of a NT testing strategy covering multiple NT MoA. Moreover, it can be used for choosing the right test system when a certain MoA of a compound is suspected.
- Readiness analyses indicate that there is more specific test method development needed.

The analyses of the *in silico* studies revealed:

- *The in silico* part in this systematic review was dedicated to models for assessment of chemicals passing the blood-brain barrier or of neurotoxicity.
- Selected 54 QSARs publications, relevant for NT, of which 39 on BBB permeation The QSARs available in the publications were developed from data on drugs and chemicals.
- Two properties often returning as important factors linked to passive BBB diffusion are lipophilicity (LogP) and charge (total polar surface area). Lipophilicity is positively correlated with passive diffusion. The total polar surface area of a molecule and heteroatoms, however, seem to have a negative impact on diffusion. These properties should be considered in the assessment of NT.
- Most of the mhe models do not consider active transport, none considers metabolites of compounds.
- There is limited experimental data for chemicals and pesticides on blood-brain barrier passage. Publications from different authors often refer to the same data sets.

The analyses of the alternative organism studies revealed:

- Rather low number of publications, in total 22, representing 3 non-mammalian whole organism models.
- A diversity of endpoint categories (in total 8 for ao) and specific endpoints related to identified MOA of neurotoxic compounds using *C. elegans*, *Drosophila* or zebrafish as ao models covering 119 citations for evaluation.
- The majority of data were available for *C. elegans* with 76 citations, representing 27 compounds (25 positive, 2 negative). The main endpoint category was inhibition of cholinergic transmission, with specific endpoints for AChE activity and motor activity, the latter confirming the added value of a whole organism approach among alternative models
- The major endpoint category studied in the fly was mitochondrial dysfunction/oxidative stress/apoptosis.
- Despite its role as a vertebrate model, the zebrafish assays were poorly available for neurotoxicity studies. Restrictions are likely due to the limited time period to be used as a non-animal methods, as early life stages are considered DNT, while after 5dpf the zebrafish is considered an animal. Nevertheless, the zebrafish model should be further explored for NT assays, especially for the period 3-5dpf when the metabolic system is developed.
- The true prediction using *C. elegans* (96%) and *Drosophila* (97%) was very high, while only 50% for zebrafish, but the latter was based on only 4 publications.

6. Conclusion

This systematic review identified a variety of cell types and test systems covering a broad variety of endpoint categories. These endpoint categories are representative for the MoA of neurotoxic compounds identified through a MoA analysis. Most predictive and covering a large variety of endpoints are primary rodent cells or tissues, next to the *C. elegans* model. Working with tissues was historically a prominent model, however, nowadays primary and stem cell models have superseded these tissue models.

With regards to species-specific effects, working with human methods is warranted. Therefore, although primary rodent cells are capable of assessing a large variety of endpoints, they do not solve the species issue. This is why lately working with hiPSC-based neural methods has evolved. So far, we do not have enough data on these test systems to understand if they could possibly substitute primary rodent cells for NT testing purposes and thus momentarily the primary rodent cell is the best method we have for NT evaluation. Nevertheless, species comparisons with primary rodent cells to validate hiPSC methods are probably very useful. Especially functional endpoints like electrical activity measures on MEAs are well-studied with rat cultures, yet sparse with hiPSC. Another shortcoming of stem cell-based methods is the lack of brain region-specificity. Rodent cell preparations from different brain regions reflect brain region-specific toxicities. To achieve similar results with hiPSC *in vitro* using human cells one needs defined differentiation protocols that reflect respective brain regions. Such are sporadically available, but are not standard procedures yet. Besides neurons, glia represent targets for neurotoxicants. Astrocytes and microglia can protect against neurotoxicity or exacerbate chemical ef-

fects on neurons, e.g. by releasing pro-inflammatory cytokines. Oligodendrocytes are targets for myelin toxins. Cell methods for glia toxicity are rare and glia endpoint categories are underrepresented. There is a strong need to put some method development focus on glia cells, astrocytes, oligodendrocytes and microglia. Preferably, one would want all these cell types in neuronal-glia co-cultures. This is also true for BBB models.

Until we possess fit-for-regulatory-purposes-hiPSC test systems for NT evaluation, i.e. covering endpoint categories and with this known NT MoA comprehensively, the results from this study can be used to select test systems according to suspected MoA for NT testing. The complementarity of assays to screen for AOP events in a reproducible way, and correctly predicted MOA should be considered including the best performing model systems (*in vitro*, ao, *in silico*) and endpoints methods identified in this review. Harmonisation of procedures and development of standard protocols with data interpretation will be necessary to enhance regulatory confidence and implementation of non-animal alternatives for neurotoxicity. Specific test method development needs to be accelerated for using test systems in a regulatory context.

7. Summary of recommendations

7.1. General

- There is consensus that there is a need for alternative methodologies that can more rapidly and cost effectively screen large numbers of chemicals for their potential to cause NT or investigate MoA to provide information on human relevance.
- As part of an IATA, the different alternative approaches *in silico*, *in vitro* and alternative organisms should be evaluated for their performance (predictivity) for regulatory needs, while considering time and cost-efficiency.
- To demonstrate performance, especially for the ao methods, but relevant for all assays, there is a need to identify negative compounds for neurotoxicity.
- Test method development for regulatory purposes is needed.

7.2. In vitro

- Cells taken out of the *in vivo* context maintain certain cellular and molecular functions they hold *in vivo.* Cellular composition and dimension of *in vitro* models are crucial. This aspect needs special attention when it comes to human stem cell-derived systems.
- Human embryonic stem cells (hESC) were employed, yet posing an ethical issue on their use and differences in national laws for working with such material.
- Human hiPSC are ethically without concern and are therefore a useful alternative to embryonic stem cells. When using hiPSC-derived neural cultures, mixed neuron and glia cultures are preferred, i.e. containing astrocytes, oligodendrocytes and microglia.

- Easy to use, highly reproducible protocols for establishing such mixed cell, hiPSC-based *in vitro* cultures need to be produced. Reliable protocols need establishment in a systematic way.
- Brain region-specific cell methods derived from hiPSC need establishment.
- Glia role in neurotoxicity need to be implemented with a particular attention to co-culture systems.
- BBB models need improvement.
- When working with hiPSC protocols, quality assurance needs to be implemented into the procedures.
- For endpoint determination, a guidance on how to use and interpret viability/cytotoxicity assays in NT *in vitro* studies is highly recommended.
- A NT *in vitro* test battery covering identified and relevant MoA for NT is recommended. Therefore, assays as test methods with relevant controls and standard operation procedures have to be set up for covering most important MoA. To link the human *in vitro* testing to rodent *in vivo* studies and validate the stem cell-derived systems, it is advised to include rodent primary cultures into the studies.
- Chemicals representing compound classes with defined MoA need to challenge the human and rat *in vitro* testing battery thereby producing reliable reference data.

7.3. Alternative organisms

- Both whole organism approaches using *C. elegans* and *Drosophila* showed high true prediction of MOA (96-97%). Especially for the nematode worm, which seems widely used and applied for a diversity of compounds with a lot of protocols for mechanistic studies (molecular and cellular events) and motor activity are available. In order to strengthen the suitability of this model, standardization of protocols might be necessary and validation studies to demonstrate its wide applicability.
- Current assays with alternative organisms show automation towards increased throughput methods by using lab-on-chip approaches, video tracking for motor activity, or high-throughput detection of fluorescent signals in transgenic models (detection of MIE, reporter systems). These developments offer future potential for low cost, fast screening of large battery of compounds to be prioritized for NT assessment.
- The zebrafish model, recognized as a highly relevant DNT model (Fritsche *et al*, 2015) and representing a vertebrate organism with known high gene conservation compared to mammals, needs further exploration as an alternative model organism for neurotoxicity. Beyond the period of embryo development (2-3 dpf), when major brain structure and cellular features are present (Nishimura et al, 2015), and the liver metabolic system is active (Goldstone & Stegeman, 2012) a multitude of endpoints, as listed for MOA can be measured in an automated way, in multi-well set-up.
- The whole animal approaches, which offer advantages to study complex processes such as molecular, cellular and signaling events in the CNS and PNS need further consideration to fill gaps identified using *in silico* & vitro approaches, and enhance *in vivo* predictions aiming to predict human neurotoxicity.

7.4. In silico

- There is limited experimental data for chemicals and pesticides on blood-brain barrier passage. Publications from different authors often refer to the same data sets. There is a need to extend the data sets for neurotoxic chemicals.
- To fully profit from the advantages of complex machine learning techniques such as Artificial Neural Networks and Support Vector Machines, the generation of new experimental data (log BB) for use as training set should be increased drastically. Mixture toxicity is considered in one model, indicating that a substance may influence the BBB-passage of another substance. Further research on the mechanism of mixture toxicity and additional experimental data on the mixture effect are needed
- It would be interesting to make a comparative study on the prediction of neurotoxicity or BBB permeation of pesticides and other chemical substances by the models provided on websites (grey literature result) and the best-performing published algorithms.

7.5. Possible EFSA follow-up activities

- Experimental project sponsorship to systematically set up a NT hiPSC-based testing strategy with primary rodent models as references. Test methods need further development. Challenging this testing battery with a test set of chemicals from different compound classes covering the proposed MoAs.
- Experimental project sponsoring evaluation of NT MoA in the zebrafish by using different compound classes.
- *In silico* analyses of the outcomes with comparisons across the different models.

BIBLIOGRAPHY

Aday S, Cecchelli R, Hallier-Vanuxeem D, Dehouck MP & Ferreira L (2016) Stem Cell-Based Human Blood-Brain Barrier Models for Drug Discovery and Delivery. *Trends Biotechnol.* 34: 382–393 Available at: http://www.ncbi.nlm.nih.gov/pubmed/26838094 [Accessed March 2, 2018]

Alloisio S, Nobile M & Novellino A (2015) Multiparametric characterisation of neuronal network activity for in vitro agrochemical neurotoxicity assessment. *Neurotoxicology* 48: 152–165

Bal-Price A, Crofton KM, Leist M, Allen S, Arand M, Buetler T, Delrue N, FitzGerald RE, Hartung T, Heinonen T, Hogberg H, Bennekou SH, Lichtensteiger W, Oggier D, Paparella M, Axelstad M, Piersma A, Rached E, Schilter B, Schmuck G, et al (2015a) International STakeholder NETwork (ISTNET): creating a developmental neurotoxicity (DNT) testing road map for regulatory purposes. *Arch. Toxicol.* 89: 269–287

Bal-Price A, Crofton KM, Sachana M, Shafer TJ, Behl M, Forsby A, Hargreaves A, Landesmann B, Lein PJ, Louisse J, Monnet-Tschudi F, Paini A, Rolaki A, Schrattenholz A, Suñol C, van Thriel C, Whelan M & Fritsche E (2015b) Putative adverse outcome pathways relevant to neurotoxicity. *Crit. Rev. Toxicol.* 45: 83–91

Bal-Price A, Hogberg H, Crofton K & Daneshian M (2018) t4 Workshop Report Recommendation on Test Readiness Criteria for New Approach Methods in Toxicology: Exemplified for Developmental Neurotoxicity1. Available at: http://www.altex.ch/resources/BalPrice_of_180223_v2.pdf [Accessed March 12, 2018]

Banerjee J, Shi Y & Azevedo HS (2016a) In vitro blood–brain barrier models for drug research: state-of-the-art and new perspectives on reconstituting these models on artificial basement membrane platforms. *Drug Discov. Today* 21: 1367–1386

Banerjee J, Shi YJ & Azevedo HS (2016b) In vitro blood-brain barrier models for drug research: state-of-the-art and new perspectives on reconstituting these models on artificial basement membrane platforms. *Drug Discov. Today* 21: 1367–1386

Baumann J, Dach K, Barenys M, Giersiefer S, Goniwiecha J, Lein PJ & Fritsche E (2015) Application of the Neurosphere Assay for DNT Hazard Assessment: Challenges and Limitations. *Methods Pharmacol. Toxicol.*: 1–29 Available at: http://link.springer.com/10.1007/7653_2015_49 [Accessed November 28, 2016]

Booth R & Kim H (2012) Characterization of a microfluidic in vitro model of the blood-brain barrier (μBBB). *Lab Chip* 12: 1784 Available at: http://xlink.rsc.org/?DOI=c2lc40094d [Accessed March 7, 2018]

Bretaud S, Lee S & Guo S (2004) Sensitivity of zebrafish to environmental toxins implicated in Parkinson's disease. *Neurotoxicol Teratol* 26: 857–864

Bujak R, Struck-Lewicka W, Kaliszan M, Kaliszan R & Markuszewski MJ (2015) Blood-brain barrier permeability mechanisms in view of quantitative structure-activity relationships (QSAR). *J Pharm Biomed Anal* 108: 29–37

Burns J & Weaver DF (2004) A mathematical model for prediction of drug molecule diffusion across the blood-brain barrier. *Can. J. Neurol. Sci.* 31: 520–527

Caito S, Yu YC & Aschner M (2013) Differential response to acrylonitrile toxicity in rat primary astrocytes and microglia. *Neurotoxicology* 37: 93–99

Campanha HM, Carvalho F & Schlosser PM (2014) Active and peripheral anionic sites of acetylcholinesterase have differential modulation effects on cell proliferation, adhesion and neuritogenesis in the NG108-15 cell line. *Toxicol Lett* 230: 122–131

Canete E & Diogene J (2008) Comparative study of the use of neuroblastoma cells (Neuro-2a) and neuroblastorna x glioma hybrid cells (NG108-15) for the toxic effect quantification of marine toxins. *Toxicon* 52: 541–550

Canete E & Diogene J (2010) Improvements in the use of neuroblastoma x glioma hybrid cells (NG108-15) for the toxic effect quantification of marine toxins. *Toxicon* 55: 381–389

Case AJ, Agraz D, Ahmad IM & Zimmerman MC (2016) Low-Dose Aronia melanocarpa Concentrate Attenuates Paraquat-Induced Neurotoxicity. *Oxid Med Cell Longev* 2016: 5296271

Castro-Coronel Y, Del Razo LM, Huerta M, Hernandez-Lopez A, Ortega A & Lopez-Bayghen E (2011) Arsenite exposure downregulates EAAT1/GLAST transporter expression in glial cells. *Toxicol Sci* 122: 539–550

Cecchelli R, Aday S, Sevin E, Almeida C, Culot M, Dehouck L, Coisne C, Engelhardt B, Dehouck M-P & Ferreira L (2014) A Stable and Reproducible Human Blood-Brain Barrier Model Derived from Hematopoietic Stem Cells. *PLoS One* 9: e99733 Available at: http://dx.plos.org/10.1371/journal.pone.0099733 [Accessed March 9, 2018]

Chang LW (1995) Handbook of Neurotoxicology. Infroma Healthcare.

Chaudhuri A, Bowling K, Funderburk C, Lawal H, Inamdar A, Wang Z & O'Donnell JM (2007) Interaction of genetic and environmental factors in a Drosophila parkinsonism model. *J Neurosci* 27: 2457–2467

Chen G, Bower KA, Xu M, Ding M, Shi X, Ke Z-J & Luo J (2009) Cyanidin-3-Glucoside Reverses Ethanol-Induced Inhibition of Neurite Outgrowth: Role of Glycogen Synthase Kinase 3 Beta. *Neurotox. Res.* 15: 321–331

Choi J, Polcher A & Joas A (2016) Systematic literature review on Parkinson's disease and Childhood Leukaemia and mode of actions for pesticides. *EFSA Support. Publ.* 13: n/a-n/a Available at: http://dx.doi.org/10.2903/sp.efsa.2016.EN-955

Cole RD, Anderson GL & Williams PL (2004a) The nematode Caenorhabditis elegans as a model of organophosphate-induced mammalian neurotoxicity. *Toxicol Appl Pharmacol* 194: 248–256

Cole RD, Anderson GL & Williams PL (2004b) The nematode Caenorhabditis elegans as a model of organophosphate-induced mammalian neurotoxicity. *Toxicol. Appl. Pharmacol.* 194: 248–256

Cookson MR & Pentreath VW (1994a) Alterations in the glial fibrillary acidic protein content of primary astrocyte cultures for evaluation of glial cell toxicity. *Toxicol Vitr.* 8: 351–359

Cookson MR & Pentreath VW (1994b) ALTERATIONS IN THE GLIAL FIBRILLARY ACIDIC PROTEIN-CONTENT OF PRIMARY ASTROCYTE CULTURES FOR EVALUATION OF GLIAL-CELL TOXICITY. *Toxicol. Vitr.* 8: 351-

Costa LG, Giordano G, Guizzetti M & Vitalone A (2008) Neurotoxicity of pesticides: a brief review. *Front. Biosci.* 13: 1240–1249

Cuadrado MU, Ruiz IL & Gomez-Nieto MA (2007) QSAR models based on isomorphic and nonisomorphic data fusion for predicting the blood brain barrier permeability. *J Comput Chem* 28: 1252–1260

Cucullo L, Hossain M, Puvenna V, Marchi N & Janigro D (2011) The role of shear stress in Blood-Brain Barrier endothelial physiology. *BMC Neurosci.* 12: 40 Available at: http://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-12-40 [Accessed March 8, 2018]

Dach K, Bendt F, Huebenthal U, Giersiefer S, Lein PJ, 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 Available at: http://www.ncbi.nlm.nih.gov/pubmed/28317842 [Accessed April 6, 2017]

Defranchi E, Novellino A, Whelan M, Vogel S, Ramirez T, van Ravenzwaay B & Landsiedel R (2011) Feasibility Assessment of Micro-Electrode Chip Assay as a Method of Detecting Neurotoxicity in vitro. *Front Neuroeng* 4: 6

Djelloul M, Holmqvist S, Boza-Serrano A, Azevedo C, Yeung MS, Goldwurm S, Frisén J, Deierborg T & Roybon L (2015) Alpha-Synuclein Expression in the Oligodendrocyte Lineage: an In Vitro and In Vivo Study Using Rodent and Human Models. *Stem Cell Reports* 5: 174–184 Available at: http://www.ncbi.nlm.nih.gov/pubmed/26235891 [Accessed March 28, 2018]

Douvaras P, Wang J, Zimmer M, Hanchuk S, O'Bara MA, Sadiq S, Sim FJ, Goldman J & Fossati V (2014) Efficient Generation of Myelinating Oligodendrocytes from Primary Progressive Multiple Sclerosis Patients by Induced Pluripotent Stem Cells. *Stem Cell Reports* 3: 250–259 Available at: http://www.ncbi.nlm.nih.gov/pubmed/25254339 [Accessed March 28, 2018]

Druwe I, Freudenrich TM, Wallace K, Shafer TJ & Mundy WR (2015) Sensitivity of neuroprogenitor cells to chemical-induced apoptosis using a multiplexed assay suitable for high-throughput screening. *Toxicology* 333: 14–24

Ehrlich M, Mozafari S, Glatza M, Starost L, Velychko S, Hallmann A-L, Cui Q-L, Schambach A, Kim K-P, Bachelin C, Marteyn A, Hargus G, Johnson RM, Antel J, Sterneckert J, Zaehres H, Schöler HR, Baron-Van Evercooren A & Kuhlmann T (2017) Rapid and efficient generation of oligodendrocytes from human induced pluripotent stem cells using transcription factors. *Proc. Natl. Acad. Sci.* 114: E2243–E2252 Available at: http://www.pnas.org/content/114/11/E2243.short [Accessed February 18, 2018]

Eigenmann DE, Xue G, Kim KS, Moses A V, Hamburger M & Oufir M (2013) Comparative study of four immortalized human brain capillary endothelial cell lines, hCMEC/D3, hBMEC, TY10, and BB19, and optimization of culture conditions, for an in vitro blood–brain barrier model for drug permeability studies. *Fluids Barriers CNS* 10: 33 Available at: http://fluidsbarrierscns.biomedcentral.com/articles/10.1186/2045-8118-10-33 [Accessed March 8, 2018]

Faqi AS (2013) A comprehensive guide to toxicology in preclinical drug development Academic Press

Fonck C & Baudry M (2003) Rapid reduction of ATP synthesis and lack of free radical formation by MPP+ in rat brain synaptosomes and mitochondria. *Brain Res* 975: 214–221

Fritsche E, Alm H, Baumann J, Geerts L, Hakansson H, Masjosthusmann S & Witters H (2015) Literature review on in vitro and alternative Developmental Neurotoxicity (DNT) testing methods. *EFSA Support. Publ.* 778: 1–186

Fujimura M & Usuki F (2012) Differing effects of toxicants (methylmercury, inorganic mercury, lead, amyloid beta, and rotenone) on cultured rat cerebrocortical neurons: differential expression of rho proteins associated with neurotoxicity. *Toxicol Sci* 126: 506–514

Gao DW, Wang P, Yang L, Peng YZ & Liang H (2002) Study on the screening of molecular structure parameter in QSAR model. *J Env. Sci Heal. A Tox Hazard Subst Env. Eng* 37: 601–609

Garg P & Verma J (2006) In silico prediction of blood brain barrier permeability: an Artificial Neural Network model. *J Chem Inf Model* 46: 289–297

Goldstone J V. & Stegeman JJ (2012) Methodological Approaches to Cytochrome P450 Profiling in Embryos. In *Methods in molecular biology (Clifton, N.J.)* pp 265–275. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22669670 [Accessed March 12, 2018]

Golmohammadi H, Dashtbozorgi Z & Khooshechin S (2017) Prediction of Blood-to-Brain Barrier Partitioning of Drugs and Organic Compounds Using a QSPR Approach. *Acta Physico-Chimica Sin.* 33: 1160–+

Grandjean P & Landrigan P (2006) Developmental neurotoxicity of industrial chemicals. *Lancet*

368: 2167–2178

Grandjean P & Landrigan PJ (2014) Neurobehavioural effects of developmental toxicity. *Lancet Neurol.* 13: 330–338 Available at: http://www.ncbi.nlm.nih.gov/pubmed/24556010 [Accessed August 18, 2016]

Hamann J, Rommelspacher H, Storch A, Reichmann H & Gille G (2006) Neurotoxic mechanisms of 2,9-dimethyl-beta-carbolinium ion in primary dopaminergic culture. *J Neurochem* 98: 1185–1199

Hamann J, Wernicke C, Lehmann J, Reichmann H, Rommelspacher H & Gille G (2008) 9-Methyl-beta-carboline up-regulates the appearance of differentiated dopaminergic neurones in primary mesencephalic culture. *Neurochem Int* 52: 688–700

Hatherell K, Couraud PO, Romero IA, Weksler B & Pilkington GJ (2011) Development of a three-dimensional, all-human in vitro model of the blood-brain barrier using mono-, co-, and tri-cultivation Transwell models. *J. Neurosci. Methods* 199: 223–229 Available at: https://www.sciencedirect.com/science/article/pii/S0165027011002706 [Accessed March 7, 2018]

Helms HC, Abbott NJ, Burek M, Cecchelli R, Couraud P-O, Deli MA, Förster C, Galla HJ, Romero IA, Shusta E V, Stebbins MJ, Vandenhaute E, Weksler B & Brodin B (2016a) In vitro models of the blood–brain barrier: An overview of commonly used brain endothelial cell culture models and guidelines for their use. *J. Cereb. Blood Flow Metab.* 36: 862–890

Helms HC, Abbott NJ, Burek M, Cecchelli R, Couraud PO, Deli MA, Forster C, Galla HJ, Romero IA, Shusta E V, Stebbins MJ, Vandenhaute E, Weksler B & Brodin B (2016b) In vitro models of the blood-brain barrier: An overview of commonly used brain endothelial cell culture models and guidelines for their use. *J. Cereb. Blood Flow Metab.* 36: 862–890

Hofrichter M, Nimtz L, Tigges J, Kabiri Y, Schröter F, Royer-Pokorac B, Hildebrandt B, Schmuck M, Epanchintsev A, Theiss S, Adjaye J, Egly J-M, Krutmann J & Fritsche E (2017) Comparative performance analysis of human iPSC-derived and primary neural progenitor cells (NPC) grown as neurospheres in vitro. *Stem Cell Res.*: under revision

Hondebrink L, Verboven AHA, Drega WS, Schmeink S, de Groot MWGDM, van Kleef RGDM, Wijnolts FMJ, de Groot A, Meulenbelt J & Westerink RHS (2016) Neurotoxicity screening of (illicit) drugs using novel methods for analysis of microelectrode array (MEA) recordings. *Neurotoxicology* 55: 1–9 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0161813X16300651

Hong S, Kim JY, Hwang J, Shin KS & Kang SJ (2013) Heptachlor induced mitochondria-mediated cell death via impairing electron transport chain complex III. *Biochem Biophys Res Commun* 437: 632–636

Hong WS, Pezzi HM, Schuster AR, Berry SM, Sung KE & Beebe DJ (2016) Development of a Highly Sensitive Cell-Based Assay for Detecting Botulinum Neurotoxin Type A through Neural Culture Media Optimization. *J Biomol Screen* 21: 65–73

Hosamani R (2013) ACUTE EXPOSURE OF *Drosophila melanogaster* TO PARAQUAT CAUSES OXIDATIVE STRESS AND MITOCHONDRIAL DYSFUNCTION. *Arch. Insect Biochem. Physiol.* 83: 25–40

Hoshi Y, Uchida Y, Tachikawa M, Inoue T, Ohtsuki S & Terasaki T (2013) Quantitative atlas of blood-brain barrier transporters, receptors, and tight junction proteins in rats and common marmoset. *J. Pharm. Sci.* 102: 3343–3355 Available at: http://www.japha.org/article/S0022-3549(15)30936-9/abstract [Accessed March 7, 2018]

Huff RA & Abou-Donia MB (1995) In vitro effect of chlorpyrifos oxon on muscarinic receptors and adenylate cyclase. *Neurotoxicology* 16: 281–290

Irons TD, MacPhail RC, Hunter DL & Padilla S (2010) Acute neuroactive drug exposures alter locomotor activity in larval zebrafish. *Neurotoxicol Teratol* 32: 84–90

Ito K, Uchida Y, Ohtsuki S, Aizawa S, Kawakami H, Katsukura Y, Kamiie J & Terasaki T (2011) Quantitative Membrane Protein Expression at the Blood – Brain Barrier of Adult and Younger Cynomolgus Monkeys. *Library (Lond).* 100: 3939–3950 Available at: http://www.japha.org/article/S0022-3549(15)31938-9/abstract [Accessed March 7, 2018]

Iyer M, Mishra R, Han Y & Hopfinger AJ (2002) Predicting Blood–Brain Barrier Partitioning of Organic Molecules Using Membrane–Interaction QSAR Analysis. *Pharm. Res.* 19: 1611–1621

Jönnson & Wim (2009) The cost of dementia in Europe. *Pharmacoeconomics* 27: 391–403

Kepp O, Galluzzi L, Lipinski M, Yuan J & Kroemer G (2011) Cell death assays for drug discovery. *Nat. Rev. Drug Discov.* 10: 221–237

Kim Y V (2009) Human astrocytes/astrocyte conditioned medium and shear stress enhance the barrier properties of human brain microvascular endothelial cells. *Brain*: 39–50 Available at: https://www.sciencedirect.com/science/article/pii/S0006899307003307 [Accessed March 6, 2018]

Klimisch H-J, Andreae M & Tillmann U (1997) A Systematic Approach for Evaluating the Quality of Experimental Toxicological and Ecotoxicological Data. *Regul. Toxicol. Pharmacol.* 25: 1–5

Krug AK, Kolde R, Gaspar JA, Rempel E, Balmer N V, Meganathan K, Vojnits K, Baquie M, Waldmann T, Ensenat-Waser R, Jagtap S, Evans RM, Julien S, Peterson H, Zagoura D, Kadereit S, Gerhard D, Sotiriadou I, Heke M, Natarajan K, et al (2013) Human embryonic stem cell-derived test systems for developmental neurotoxicity: a transcriptomics approach. *Arch Toxicol* 87: 123–143

Kubik LL & Philbert MA (2015) The Role of Astrocyte Mitochondria in Differential Regional Susceptibility to Environmental Neurotoxicants: Tools for Understanding Neurodegeneration. *Toxicol. Sci.* 144: 7–16

Landrigan PJ, Sonawane B, Butler RN, Trasande L, Callan R & Droller D (2005) Early environmental origins of neurodegenerative disease in later life. *Environ. Health Perspect.* 113: 1230–3 Available at: http://www.ncbi.nlm.nih.gov/pubmed/16140633 [Accessed May 25, 2017]

Lee HJ, Han J, Jang Y, Kim SJ, Park JH, Seo KS, Jeong S, Shin S, Lim K, Heo JY & Kweon GR (2015a) Docosahexaenoic acid prevents paraquat-induced reactive oxygen species production in dopaminergic neurons via enhancement of glutathione homeostasis. *Biochem Biophys Res Commun* 457: 95–100

Lee JH, Mitchell RR, McNicol JD, Shapovalova Z, Laronde S, Tanasijevic B, Milsom C, Casado F, Fiebig-Comyn A, Collins TJ, Singh KK & Bhatia M (2015b) Single Transcription Factor Conversion of Human Blood Fate to NPCs with CNS and PNS Developmental Capacity. *Cell Rep* 11: 1367–1376

Leist M & Hartung T (2013) Inflammatory findings on species extrapolations: humans are definitely no 70-kg mice. *Arch. Toxicol.* 87: 563–567

Leist M, Ringwald A, Kolde R, Bremer S, van Thriel C, Krause K-H, Rahnenführer J, Sachinidis A, Hescheler J & Hengstler JG (2013) Test systems of developmental toxicity: state-of-the art and future perspectives. *Arch. Toxicol.* 87: 2037–2042

Li J, Spletter ML, Johnson DA, Wright LS, Svendsen CN & Johnson JA (2005) Rotenone-induced caspase 9/3-independent and -dependent cell death in undifferentiated and differentiated human neural stem cells. *J Neurochem* 92: 462–476

Li W, Sun W, Zhang Y, Wei W, Ambasudhan R, Xia P, Talantova M, Lin T, Kim J, Wang X, Kim WR, Lipton SA, Zhang K & Ding S (2011) Rapid induction and long-term self-renewal of primitive neural precursors from human embryonic stem cells by small molecule inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* 108: 8299–304 Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1014041108 [Accessed March 9, 2018]

Lippmann ES, Al-Ahmad A, Azarin SM, Palecek SP & Shusta E V. (2015) A retinoic acid-enhanced, multicellular human blood-brain barrier model derived from stem cell sources. *Sci. Rep.* 4: 4160 Available at: https://www.nature.com/articles/srep04160 [Accessed March 8, 2018]

Lippmann ES, Azarin SM, Kay JE, Nessler RA, Wilson HK, Al-Ahmad A, Palecek SP & Shusta E V (2012) Derivation of Blood-Brain Barrier Endothelial Cells from Human Pluripotent Stem Cells. *Nat. Biotechnol.* 30: 783–791 Available at: https://www.nature.com/articles/nbt.2247 [Accessed March 8, 2018]

Malik N, Efthymiou AG, Mather K, Chester N, Wang X, Nath A, Rao MS & Steiner JP (2014) Compounds with species and cell type specific toxicity identified in a 2000 compound drug screen of neural stem cells and rat mixed cortical neurons. *Neurotoxicology* 45: 192–200

Matthews RA (2008) Medical progress depends on animal models - doesn't it? *J. R. Soc. Med.* 101: 95–98

Maurer LL & Philbert MA (2015) The mechanisms of neurotoxicity and the selective vulnerability of nervous system sites. *Handb. Clin. Neurol.* 131: 61–70

McConnell ER, McClain MA, Ross J, LeFew WR & Shafer TJ (2012) Evaluation of multi-well microelectrode arrays for neurotoxicity screening using a chemical training set. *Neurotoxicology* 33: 1048–1057

Méry B, Guy J-B, Vallard A, Espenel S, Ardail D, Rodriguez-Lafrasse C, Rancoule C & Magné N (2017) In Vitro Cell Death Determination for Drug Discovery: A Landscape Review of Real Issues. *J. Cell Death* 10: 117967071769125

Miller RL, Sun GY & Sun AY (2007) Cytotoxicity of paraquat in microglial cells: Involvement of PKCdelta- and ERK1/2-dependent NADPH oxidase. *Brain Res* 1167: 129–139

Nakagawa S, Deli MA, Kawaguchi H, Shimizudani T, Shimono T, Kittel Á, Tanaka K & Niwa M (2009) A new blood–brain barrier model using primary rat brain endothelial cells, pericytes and astrocytes. *Neurochem. Int.* 54: 253–263 Available at: https://www.sciencedirect.com/science/article/pii/S0197018608001976 [Accessed March 7, 2018]

Nakagawa S, Deli MA, Nakao S, Honda M, Hayashi K, Nakaoke R, Kataoka Y & Niwa M (2007) Pericytes from Brain Microvessels Strengthen the Barrier Integrity in Primary Cultures of Rat Brain Endothelial Cells. *Cell. Mol. Neurobiol.* 27: 687–694 Available at: http://link.springer.com/10.1007/s10571-007-9195-4 [Accessed March 7, 2018]

Nicolas J, Hendriksen PJ, van Kleef RG, de Groot A, Bovee TF, Rietjens IM & Westerink RH (2014) Detection of marine neurotoxins in food safety testing using a multielectrode array. *Mol Nutr Food Res* 58: 2369–2378

Nishimura Y, Murakami S, Ashikawa Y, Sasagawa S, Umemoto N, Shimada Y & Tanaka T (2015) Zebrafish as a systems toxicology model for developmental neurotoxicity testing. *Congenit. Anom. (Kyoto).* 55: 1–16 Available at: http://www.ncbi.nlm.nih.gov/pubmed/25109898 [Accessed March 9, 2018]

Nogueira TB, da Costa Araujo S, Carvalho F, Pereira FC, Fernandes E, Bastos ML, Costa VM & Capela JP (2014) Modeling chronic brain exposure to amphetamines using primary rat neuronal cortical cultures. *Neuroscience* 277: 417–434

Ntzani EE, Ntritsos G CM, Evangelou E & Tzoulaki I (2013) Literature review on epidemiological studies linking exposure to pesticides and health effects. *EFSA Support. Publ.* 10:

Ockleford C, Adriaanse P, Berny P, Brock T, Duquesne S, Grilli S, Hernandez-Jerez AF, Bennekou SH, Klein M, Kuhl T, Laskowski R, Machera K, Pelkonen O, Pieper S, Smith R, Stemmer M, Sundh I, Teodorovic I, Tiktak A, Topping CJ, et al (2017) Investigation into experimental toxicological properties of plant protection products having a potential link to Parkinson's disease and childhood leukaemia. *EFSA J.* 15: Available at: http:https://doi.org/10.2903/j.efsa.2017.4691

OECD (2013) Revised Guidance Document on Developing and Assessing Adverse Outcome Pathways. *Ser. Test. Assess.* 184:

Palmer AM & Alavijeh MS (2013) Overview of experimental models of the blood-brain barrier in CNS drug discovery. *Curr Protoc Pharmacol* 62: Unit 7.15.

Palmer AM, Alavijeh MS, Palmer AM & Alavijeh MS (2013) Overview of Experimental Models of the Blood-Brain Barrier in CNS Drug Discovery. In *Current Protocols in Pharmacology* p 7.15.1-7.15.30. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Patlewicz G, Kuseva C, Kesova A, Popova I, Zhechev T, Pavlov T, Roberts DW & Mekenyan O (2014) Towards AOP application – Implementation of an integrated approach to testing and assessment (IATA) into a pipeline tool for skin sensitization. *Regul. Toxicol. Pharmacol.* 69: 529–545

Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, Macleod M, Mignini LE, Jayaram P & Khan KS (2007) Comparison of treatment effects between animal experiments and clinical trials: systematic review. *BMJ* 334: 197

Polishchuk P, Tinkov O, Khristova T, Ognichenko L, Kosinskaya A, Varnek A & Kuz'min V (2016) Structural and Physico-Chemical Interpretation (SPCI) of QSAR Models and Its Comparison with Matched Molecular Pair Analysis. *J Chem Inf Model* 56: 1455–1469

Ponio JB-D, El-Ayoubi F, Glacial F, Ganeshamoorthy K, Driancourt C, Godet M, Perrière N, Guillevic O, Couraud PO & Uzan G (2014) Instruction of Circulating Endothelial Progenitors In Vitro towards Specialized Blood-Brain Barrier and Arterial Phenotypes. *PLoS One* 9: e84179 Available at: http://dx.plos.org/10.1371/journal.pone.0084179 [Accessed March 9, 2018]

Rajini PS, Melstrom P & Williams PL (2008) A Comparative Study on the Relationship Between Various Toxicological Endpoints in *Caenorhabditis elegans* Exposed to Organophosphorus Insecticides. *J. Toxicol. Environ. Heal. Part A* 71: 1043–1050

Sanchez-Ramos J, Facca A, Basit A & Song S (1998) Toxicity of dieldrin for dopaminergic neurons in mesencephalic cultures. *Exp Neurol* 150: 263–271

Sawyer TW, Weiss MT & Dickinson T (1992) EFFECT OF METABOLISM ON THE ANTICHOLINESTERASE ACTIVITY OF PARATHION AND PARAOXON IN PRIMARY NEURON CULTURES. *Toxicol. Vitr.* 6: 569–574

Sayre LM, Wang FJ, Arora PK, Riachi NJ, Harik SI & Hoppel CL (1991) DOPAMINERGIC NEUROTOXICITY INVIVO AND INHIBITION OF MITOCHONDRIAL RESPIRATION INVITRO BY POSSIBLE ENDOGENOUS PYRIDINIUM-LIKE SUBSTANCES. *J Neurochem* 57: 2106–2115

Schiera G, Bono E, Raffa MP, Gallo A, Pitarresi GL, Liegro I & Savettieri G (2003) Synergistic effects of neurons and astrocytes on the differentiation of brain capillary endothelial cells in culture. *J. Cell. Mol. Med.* 7: 165–170 Available at: http://doi.wiley.com/10.1111/j.1582-

4934.2003.tb00215.x [Accessed March 6, 2018]

Schmuck G & Ahr HJ (1997) Improved in vitro method for screening organophosphate-induced delayed polyneuropathy. *Toxicol. Vitr.* 11: 263–270

Seidel D, Jahnke HG, Englich B, Girard M & Robitzki AA (2017) In vitro field potential monitoring on a multi-microelectrode array for the electrophysiological long-term screening of neural stem cell maturation. *Analyst* 142: 1929–1937

Shawahna R, Uchida Y, Declèves X, Ohtsuki S, Yousif S, Dauchy S, Jacob A, Chassoux F, Daumas-Duport C, Couraud PO, Terasaki T & Scherrmann JM (2011) Transcriptomic and quantitative proteomic analysis of transporters and drug metabolizing enzymes in freshly isolated human brain microvessels. *Mol. Pharm.* 8: 1332–1341 Available at: http://pubs.acs.org/doi/abs/10.1021/mp200129p [Accessed March 7, 2018]

Shukla AK, Pragya P, Chaouhan HS, Tiwari AK, Patel DK, Abdin MZ & Chowdhuri DK (2014) Heat Shock Protein-70 (Hsp-70) Suppresses Paraquat-Induced Neurodegeneration by Inhibiting JNK and Caspase-3 Activation in Drosophila Model of Parkinson's Disease. *PLoS One* 9:

Siddharthan V, Kim Y V., Liu S & Kim KS (2007) Human astrocytes/astrocyte-conditioned medium and shear stress enhance the barrier properties of human brain microvascular endothelial cells. *Brain Res.* 1147: 39–50 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0006899307003307 [Accessed March 7, 2018]

Simons E FR (1994) H1-Receptor Antagonists. *Drug Saf.* 10: 350–380

Sirenko O, Hesley J, Rusyn I & Cromwell EF (2014) High-Content High-Throughput Assays for Characterizing the Viability and Morphology of Human iPSC-Derived Neuronal Cultures. *Assay Drug Dev Technol* 12: 536–547

Smirnova L, Harris G, Delp J, Valadares M, Pamies D, Hogberg HT, Waldmann T, Leist M & Hartung T (2016) A LUHMES 3D dopaminergic neuronal model for neurotoxicity testing allowing long-term exposure and cellular resilience analysis. *Arch Toxicol* 90: 2725–2743

Stumm G, Schlegel J, Schafer T, Wurz C, Mennel HD, Krieg JC & Vedder H (1999) Amphetamines induce apoptosis and regulation of bcl-x splice variants in neocortical neurons. *Faseb j* 13: 1065–1072

Suenderhauf C, Hammann F & Huwyler J (2012) Computational Prediction of Blood-Brain Barrier Permeability Using Decision Tree Induction. *Molecules* 17: 10429–10445

Syvänen S, Lindhe Ö, Palner M, Lindhe O, Kornum BR, Rahman O, Långström B, Knudsen GM & Hammarlund-Udenaes M (2008) Species differences in blood-brain barrier transport of three PET radioligands with emphasis on P-glycoprotein transport. *Drug Metab. Dispos.* 37: 635–643 Available at: http://dmd.aspetjournals.org/content/37/3/635.short [Accessed March 7, 2018]

Tanner CM, Goldman SM, Ross GW & Grate SJ (2014) The disease intersection of susceptibility and exposure: chemical exposures and neurodegenerative disease risk. *Alzheimers. Dement.* 10: S213-25

Taskiran D, Nesil T & Alkan K (2007) Mitochondrial oxidative stress in female and male rat brain after ex vivo carbon monoxide treatment. *Hum Exp Toxicol* 26: 645–651

Tollefsen KE, Scholz S, Cronin MT, Edwards SW, de Knecht J, Crofton K, Garcia-Reyero N, Hartung T, Worth A & Patlewicz G (2014) Applying Adverse Outcome Pathways (AOPs) to support Integrated Approaches to Testing and Assessment (IATA). *Regul. Toxicol. Pharmacol.* 70: 629–640

Tropsha A, Gramatica P & Gombar V (2003) The Importance of Being Earnest: Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. *QSAR Comb. Sci.* 22: 69–77

Tukker AM, de Groot MW, Wijnolts FM, Kasteel EE, Hondebrink L & Westerink RH (2016) Is the time right for in vitro neurotoxicity testing using human iPSC-derived neurons? *ALTEX* 33: 261–271

Uchida Y, Ohtsuki S, Katsukura Y, Ikeda C, Suzuki T, Kamiie J & Terasaki T (2011) Quantitative targeted absolute proteomics of human blood-brain barrier transporters and receptors. *J. Neurochem.* 117: 333–345 Available at: http://doi.wiley.com/10.1111/j.1471-4159.2011.07208.x [Accessed March 7, 2018]

Uchida Y, Tachikawa M, Obuchi W, Hoshi Y, Tomioka Y, Ohtsuki S & Terasaki T (2013) A study protocol for quantitative targeted absolute proteomics (QTAP) by LC-MS/MS: application for inter-strain differences in protein expression levels of transporters, receptors, claudin-5, and marker proteins at the blood--brain barrier in ddY, FVB, an. *Fluids Barriers CNS* 10: 21 Available at: http://fluidsbarrierscns.biomedcentral.com/articles/10.1186/2045-8118-10-21 [Accessed March 7, 2018]

Valdivia P, Martin M, LeFew WR, Ross J, Houck KA & Shafer TJ (2014) Multi-well microelectrode array recordings detect neuroactivity of ToxCast compounds. *Neurotoxicology* 44: 204–217

Vassallo A, Chiappalone M, De Camargos Lopes R, Scelfo B, Novellino A, Defranchi E, Palosaari T, Weisschu T, Ramirez T, Martinoia S, Johnstone AFM, Mack CM, Landsiedel R, Whelan M, Bal-Price A & Shafer TJ (2017) A multi-laboratory evaluation of microelectrode array-based measurements of neural network activity for acute neurotoxicity testing. *Neurotoxicology* 60: 280–292

Villeneuve DL, Crump D, Garcia-Reyero N, Hecker M, Hutchinson TH, LaLone CA, Landesmann B, Lettieri T, Munn S, Nepelska M, Ottinger MA, Vergauwen L & Whelan M (2014) Adverse Outcome Pathway (AOP) Development I: Strategies and Principles. *Toxicol. Sci.* 142: 312–320

van Vliet E, Stoppini L, Balestrino M, Eskes C, Griesinger C, Sobanski T, Whelan M, Hartung T & Coecke S (2007) Electrophysiological recording of re-aggregating brain cell cultures on

multi-electrode arrays to detect acute neurotoxic effects. *Neurotoxicology* 28: 1136–1146

Wang S, Bates J, Li X, Schanz S, Chandler-Militello D, Levine C, Maherali N, Studer L, Hochedlinger K, Windrem M & Goldman SA (2013) Human iPSC-Derived Oligodendrocyte Progenitor Cells Can Myelinate and Rescue a Mouse Model of Congenital Hypomyelination. *Cell Stem Cell* 12: 252–264 Available at: https://www.sciencedirect.com/science/article/pii/S1934590912007072 [Accessed March 12, 2018]

Wilson HK, Canfield SG, Hjortness MK, Palecek SP & Shusta E V (2015) Exploring the effects of cell seeding density on the differentiation of human pluripotent stem cells to brain microvascular endothelial cells. *Fluids Barriers CNS* 12: 13

Wilson WW, Shapiro LP, Bradner JM & Caudle WM (2014) Developmental exposure to the organochlorine insecticide endosulfan damages the nigrostriatal dopamine system in male offspring. *Neurotoxicology* 44: 279–287

Wolff A, Antfolk M, Brodin B & Tenje M (2015a) In Vitro Blood-Brain Barrier Models-An Overview of Established Models and New Microfluidic Approaches. *J Pharm Sci* 104: 2727–2746

Wolff A, Antfolk M, Brodin B & Tenje M (2015b) In Vitro Blood-Brain Barrier Models-An Overview of Established Models and New Microfluidic Approaches. *J. Pharm. Sci.* 104: 2727–46

Xue Q, Liu Y, Qi H, Ma Q, Xu L, Chen W, Chen G & Xu X (2013) A novel brain neurovascular unit model with neurons, astrocytes and microvascular endothelial cells of rat. *Int. J. Biol. Sci.* 9: 174–89 Available at: http://www.ncbi.nlm.nih.gov/pubmed/23412420 [Accessed March 7, 2018]

Yan A, Liang H, Chong Y, Nie X & Yu C (2013) In-silico prediction of blood-brain barrier permeability. *SAR QSAR Env. Res* 24: 61–74

Yan Y, Bejoy J, Xia J, Guan J, Zhou Y & Li Y (2016) Neural patterning of human induced pluripotent stem cells in 3-D cultures for studying biomolecule-directed differential cellular responses. *Acta Biomater* 42: 114–126

El Yazal J, Rao SN, Mehl A & Slikker W (2001) Prediction of organophosphorus acetylcholinesterase inhibition using three-dimensional quantitative structure-activity relationship (3D-QSAR) methods. *Toxicol. Sci.* 63: 223–232

Zagoura D, Canovas-Jorda D, Pistollato F, Bremer-Hoffmann S & Bal-Price A (2017a) Evaluation of the rotenone-induced activation of the Nrf2 pathway in a neuronal model derived from human induced pluripotent stem cells. *Neurochem Int* 106: 62–73

Zagoura D, Canovas-Jorda D, Pistollato F, Bremer-Hoffmann S & Bal-Price A (2017b) Evaluation of the rotenone-induced activation of the Nrf2 pathway in a neuronal model derived from human induced pluripotent stem cells. *Neurochem. Int.* 106: 62–73

Zhang L, Li L, Liu H, Prabhakaran K, Zhang X, Borowitz JL & Isom GE (2007) HIF-1alpha activation by a redox-sensitive pathway mediates cyanide-induced BNIP3 upregulation and mitochondrial-dependent cell death. *Free Radic Biol Med* 43: 117–127

Zhao X, Wang R, Xiong J, Yan D, Li A, Wang S, Xu J & Zhou J (2017) JWA antagonizes paraquat-induced neurotoxicity via activation of Nrf2. *Toxicol Lett*

Zurich MG, Stanzel S, Kopp-Schneider A, Prieto P & Honegger P (2013) Evaluation of aggregating brain cell cultures for the detection of acute organ-specific toxicity. *Toxicol. Vitr.* 27: 1416–1424

APPENDICES

Appendices A to D can be found in the online version of this output ('Supporting information' section): http://dx.doi.org/10.2903/j.efsa.2018.1410

Appendix A. List of Compounds and compound classes identified from two publications by Grandjean and Landrigan (20061 and 20142) and an EFSA supporting publication by Choi et al. (20163).

Appendix B. List of Natural Compounds (NCs) identified from the Handbook of Neurotoxicity.

Appendix C. MoA analysis sheet (layout)

Appendix D. Detailed information on Mode of Actions (MoAs) of all compounds and compound classes (including natural neurotoxins) with an identified (partial) MoA from Appendix A and B and association to a respective MoA/endpoint category.

Appendix E. Search strings

Neurotoxicity search

PubMed:

(("1990/01/01"[Date - Publication] : "2017/04/31"[Date - Publication])) AND (((((toxic*[Title/Abstract]) AND ((((((brain[Title/Abstract]) OR CNS[Title/Abstract]) OR PNS[Title/Abstract]) OR nervous system[Title/Abstract]) OR neurological*[Title/Abstract]) OR neural*[Title/Abstract]))) OR neurotoxi*[Title/Abstract]))

WoS:

(toxic* AND (Brain OR CNS OR PNS OR nervous system OR neurological* OR neural*)) OR neurotoxi*

Compound search (for Chlorpyrifos as an example compound)

PubMed:

"Zidil"[Title/Abstract] OR "Grofo"[Title/Abstract] OR "Brodan"[Title/Abstract] OR "Suscon"[Title/Abstract] OR "Durmet"[Title/Abstract] OR "Terial"[Title/Abstract] OR "XRM 429"[Title/Abstract] OR "Dursban"[Title/Abstract] OR "Lorsban"[Title/Abstract] OR "Pyrinex"[Title/Abstract] OR "Bonidel"[Title/Abstract] OR "Coroban"[Title/Abstract] OR "Lentrek"[Title/Abstract] OR "Lock-On"[Title/Abstract] OR "Spannit"[Title/Abstract] OR "Tafaban"[Title/Abstract] OR "HSDB 389"[Title/Abstract] OR "OMS-0971"[Title/Abstract] OR "XRM 5160"[Title/Abstract] OR "Piridane"[Title/Abstract] OR "Danusban"[Title/Abstract] OR "2921-88-2"[Title/Abstract] OR "AI3-27311"[Title/Abstract] OR "Dowco 179"[Title/Abstract] OR "ENT 27311"[Title/Abstract] OR "Detmol ua"[Title/Abstract] OR "Dursban F"[Title/Abstract] OR "Dhanusban"[Title/Abstract] OR "Dursban R"[Title/Abstract] OR "Geodinfos"[Title/Abstract] OR "CCRIS 7144"[Title/Abstract] OR "Dursban 4E"[Title/Abstract] OR "Dursban 44"[Title/Abstract] OR "Terial 40L"[Title/Abstract] OR "Killmaster"[Title/Abstract] OR "BRN 1545756"[Title/Abstract] OR "Detmol"[Title/Abstract] OR "suSCon Blue"[Title/Abstract] OR "Lorsban 50SL"[Title/Abstract] OR "Dursban 10CR"[Title/Abstract] OR "Chlorpyrifos"[Title/Abstract] OR "suSCon Green"[Title/Abstract] OR "Chlorpyriphos"[Title/Abstract] OR "UNII-JCS58I644W"[Title/Abstract] OR "Trichlorpyrphos"[Title/Abstract] OR "EINECS 220-864-4"[Title/Abstract] OR "Caswell No. 219AA"[Title/Abstract] OR "Radar"[Title/Abstract] OR "Chlorpyrifos ethyl"[Title/Abstract] OR "Chlorpyrifos-ethyl"[Title/Abstract] OR "Chlorpyriphos-ethyl"[Title/Abstract] OR "Ethyl chlorpyriphos"[Title/Abstract] OR "Chlorpyriphos"[Title/Abstract] OR "Chlorpyrifos"[Title/Abstract] OR "EPA Pesticide Chemical Code 059101"[Title/Abstract] OR "O,O-Diaethyl-O-3,5,6-trichlor-2-pyridylmonothiophosphat"[Title/Abstract] OR "O,O-Diethyl O-3,5,6-trichloro-2-pyridyl phosphorothioate"[Title/Abstract] OR "O,O-Diethyl O-(3,5,6-trichloro-2-pyridyl) phsophorothioate"[Title/Abstract] OR "O,O-Diethyl O-(3,5,6-trichloro-2-pyridinyl)phosphorothioate"[Title/Abstract] OR "O,O-Diaethyl-O-3,5,6-trichlor-2-pyridylmonothiophosphat"[Title/Abstract] OR "Phosphorothioic acid, O,O-diethyl O-(3,5,6-trichloro-2-pyridyl) ester"[Title/Abstract] OR "Phosphorothioic acid, O,O-diethyl O-(3,5,6-trichloro-2-pyridinyl) ester"[Title/Abstract] OR "2-Pyridinol, 3,5,6-trichloro-, O-ester with O,O-diethyl phosphorothioate"[Title/Abstract]

WoS:

"Zidil" OR "Grofo" OR "Brodan" OR "Suscon" OR "Durmet" OR "Terial" OR "XRM 429" OR "Dursban" OR "Lorsban" OR "Pyrinex" OR "Bonidel" OR "Coroban" OR "Lentrek" OR "Lock-On" OR "Spannit" OR "Tafaban" OR "HSDB 389" OR "OMS-0971" OR "XRM 5160" OR "Piridane" OR "Danusban" OR "2921-88-2" OR "AI3-27311" OR "Dowco 179" OR "ENT 27311" OR "Detmol ua" OR "Dursban F" OR "Dhanusban" OR "Dursban R" OR "Geodinfos" OR "CCRIS 7144" OR "Dursban 4E" OR "Dursban 44" OR "Terial 40L" OR "Killmaster" OR "BRN 1545756" OR "Detmol" OR "suSCon Blue" OR "Lorsban 50SL" OR "Dursban 10CR" OR "Chlorpyrifos" OR "suSCon Green" OR "Chlorpyriphos" OR "UNII-JCS58I644W" OR "Trichlorpyrphos" OR "EINECS 220-864-4" OR "Caswell No. 219AA" OR "Chlorpyrifos ethyl" OR "Chlorpyrifos-ethyl" OR "Chlorpyriphos-ethyl" OR "Ethyl chlorpyriphos" OR "Chlorpyriphos" OR "Chlorpyrifos" OR "EPA

Pesticide Chemical Code 059101" OR "O,O-Diaethyl-O-3,5,6-trichlor-2-pyridylmonothiophosphat" OR "O,O-Diethyl O-3,5,6-trichloro-2-pyridyl phosphorothioate" OR "O,O-Diethyl O-(3,5,6-trichloro-2-pyridyl) phsophorothioate" OR "O,O-Diethyl O-(3,5,6-trichloro-2-pyridinyl)phosphorothioate" OR "O,O-Diaethyl-O-3,5,6-trichlor-2-pyridylmonothiophosphat" OR "Phosphorothioic acid, O,O-diethyl O-(3,5,6-trichloro-2-pyridyl) ester" OR "Phosphorothioic acid, O,O-diethyl O-(3,5,6-trichloro-2-pyridinyl) ester" OR "2-Pyridinol, 3,5,6-trichloro-, O-ester with O,O-diethyl phosphorothioate"

Alternative method search 1 (combined by 'OR')

In vitro:

PubMed

((((((model*[Title/Abstract] OR test[Title/Abstract] OR assay*[Title/Abstract] OR method*[Title/Abstract] OR technique*[Title/Abstract] OR set up[Title/Abstract] OR experiment*[Title/Abstract] OR endpoint*[Title/Abstract] OR prioritization*[Title/Abstract] OR system*[Title/Abstract] OR evaluation*[Title/Abstract] OR exposure*[Title/Abstract] OR testing[Title/Abstract] OR tests[Title/Abstract])))) AND in vitro[Title/Abstract])))) OR ((culture*[Title/Abstract] OR brain slice*[Title/Abstract] OR cell based[Title/Abstract] OR cell line*[Title/Abstract] OR cell model*[Title/Abstract] OR cell system*[Title/Abstract] OR cellular model*[Title/Abstract] OR cellular assay*[Title/Abstract] OR cellular system*[Title/Abstract] OR cellular method*[Title/Abstract] OR cellular technique*[Title/Abstract] OR cellular endpoint*[Title/Abstract] OR cellular exposure*[Title/Abstract] OR immortalised[Title/Abstract] OR immortalized[Title/Abstract] OR IPS cell*[Title/Abstract] OR primary cell*[Title/Abstract] OR In Vitro Techniques[MeSH] OR tumor cell line*[Title/Abstract] OR Cells, Cultured[MeSH] OR Astrocyte*[Title/Abstract] OR ESC[Title/Abstract] OR glial cell*[Title/Abstract] OR iPSC[Title/Abstract] OR nerve cell*[Title/Abstract] OR neural cell*[Title/Abstract] OR neuroblastoma [Title/Abstract] OR neuronal cell*[Title/Abstract] OR oligodendrocyte*[Title/Abstract] OR pheochromocytoma*[Title/Abstract] OR pluripotent cell*[Title/Abstract] OR schwann cell*[Title/Abstract] OR stem cell*[Title/Abstract] OR teratocarcinoma*[Title/Abstract] OR tumor cell[Title/Abstract] OR microglia[Title/Abstract]))

WoS

(("model*" OR "test OR assay*" OR "method*" OR "technique*" OR "set up" OR "experiment*" OR "endpoint*" OR "prioritization*" OR "system*" OR "evaluation*" OR "exposure*" OR "testing" OR "tests") AND "in vitro") OR "culture*" OR "brain slice*" OR "cell based" OR "cell line*" OR "cell model*" OR "cell system*" OR "cellular model*" OR "cellular assay*" OR "cellular system*" OR "cellular method*" OR "cellular technique*" OR "cellular endpoint*" OR "cellular exposure*" OR "immortalised" OR "immortalized" OR "IPS cell*" OR "primary cell*" OR "tumor cell line*" OR "Astrocyte*" OR "ESC" OR "glial cell*" OR "iPSC" OR "nerve cell*" OR "neural cell*" OR "neuroblastoma " OR "neuronal cell*" OR "oligodendrocyte*" OR "pheochromocytoma*" OR "pluripotent cell*" OR "schwann cell*" OR "stem cell*" OR "teratocarcinoma*" OR "tumor cell" OR "microglia*"

Cell free

PubMed

acellular assay*[Title/Abstract] OR biochemical assay*[Title/Abstract] OR biomimetic*[Title/Abstract] OR biosensor*[Title/Abstract] OR cell free[Title/Abstract] OR cellfree[Title/Abstract] OR enzyme assay*[Title/Abstract] OR lab on a chip*[Title/Abstract] OR non cell assay*[Title/Abstract] OR receptor assay*[Title/Abstract] OR reporter assay*[Title/Abstract] OR binding assay*[Title/Abstract] OR Cell-Free System[MeSH]

WoS

"acellular assay*" OR "biochemical assay*" OR "biomimetic*" OR "biosensor*" OR "cell free" OR "cellfree" OR "enzyme assay*" OR "lab on a chip*" OR "non cell assay*" OR "receptor assay*" OR "reporter assay*" OR "binding assay*"

Alternative organism

PubMed

(invertebrate*[Title/Abstract] OR non mammal*[Title/Abstract] OR caenorhabditis[Title/Abstract] OR C. elegans[Title/Abstract] OR nematod*[Title/Abstract] OR rerio[Title/Abstract] OR roundworm*[Title/Abstract] OR sea urchin*[Title/Abstract] OR zebra fish[Title/Abstract] OR zebrafish[Title/Abstract] OR Xenopus[Title/Abstract] OR tadpole[Title/Abstract] OR clawed frog*[Title/Abstract] OR Zebrafish[MeSH] OR Caenorhabditis[MeSH] OR Sea Urchins[MeSH] OR Xenopus[MeSH] OR Brachydanio[Title/Abstract] OR Brachydanio rerio[Title/Abstract] OR Danio [Title/Abstract] OR Caenorhabditis elegan*[Title/Abstract] OR Fruitfly[Title/Abstract] OR Fruit fly[Title/Abstract] OR Drosophila[Title/Abstract] OR Drosophila melanog*[Title/Abstract] OR Echinoid*[Title/Abstract] OR fruit flies[Title/Abstract] OR fruitflies[Title/Abstract] OR Drosophila[MeSH]

WoS

"invertebrate*" OR "non mammal*" OR "caenorhabditis" OR "C. elegans" OR "nematod*" OR "rerio" OR "roundworm*" OR "sea urchin*" OR "zebra fish" OR "zebrafish" OR "Xenopus" OR "tadpole" OR "clawed frog*" OR "Brachydanio" OR "Brachydanio rerio" OR "Danio " OR "Caenorhabditis elegan*" OR "Fruitfly" OR "Fruit fly" OR "Drosophila" OR "Drosophila melanog*" OR "Echinoid*" OR "fruit flies" OR "fruitflies"

Alternative method related (general)

PubMed

alternative approach*[Title/Abstract] OR alternative assay*[Title/Abstract] OR alternative method*[Title/Abstract] OR alternative model*[Title/Abstract] OR alternative to animal*[Title/Abstract] OR alternatives to animal*[Title/Abstract] OR analogue approach*[Title/Abstract] OR animalfree[Title/Abstract] OR animal-free[Title/Abstract] OR assay characterisation[Title/Abstract] OR assay characterization[Title/Abstract] OR assay development[Title/Abstract] OR assay performance[Title/Abstract] OR assay validation[Title/Abstract] OR bio assay*[Title/Abstract] OR bioassay*[Title/Abstract] OR Biological Assay*[Title/Abstract] OR in vitro screening*[Title/Abstract] OR method characterization[Title/Abstract] OR method characterization[Title/Abstract] OR method development[Title/Abstract] OR method validation[Title/Abstract] OR model characterisation[Title/Abstract] OR model characterization[Title/Abstract] OR model development[Title/Abstract] OR model validation[Title/Abstract] OR non animal alternativ*[Title/Abstract] OR non animal alternative*[Title/Abstract] OR non testing method[Title/Abstract] OR sceening tool*[Title/Abstract] OR screening assay*[Title/Abstract] OR screening method*[Title/Abstract] OR screening system*[Title/Abstract] OR screening test*[Title/Abstract] OR test assay*[Title/Abstract] OR test batteries[Title/Abstract] OR test battery[Title/Abstract] OR test method*[Title/Abstract] OR test strategies[Title/Abstract] OR test strategy[Title/Abstract] OR test system*[Title/Abstract] OR testing assay*[Title/Abstract] OR testing batteries[Title/Abstract] OR testing battery[Title/Abstract] OR testing method*[Title/Abstract] OR testing strategies[Title/Abstract] OR testing strategy[Title/Abstract] OR three dimensional cell culture[Title/Abstract] OR three dimensional model[Title/Abstract] OR throughput screening[Title/Abstract] OR Biological Assay[MeSH]

WoS

"alternative approach*" OR "alternative assay*" OR "alternative method*" OR "alternative model*" OR "alternative to animal*" OR "alternatives to animal*" OR "analogue approach*" OR "animalfree" OR "animal-free" OR "assay characterisation" OR "assay characterization" OR "assay development" OR "assay performance" OR "assay validation" OR "bio assay*" OR "bioassay*" OR "Biological Assay*" OR "in vitro screening*" OR "method characterization" OR "method characterization" OR "method development" OR "method validation" OR "model characterisation" OR "model characterization" OR "model development" OR "model validation" OR "non animal alternativ*" OR "non animal alternative*" OR "non testing method" OR "sceening tool*" OR "screening assay*" OR "screening method*" OR "screening system*" OR "screening test*" OR "test assay*" OR "test batteries" OR "test battery" OR "test method*" OR "test strategies" OR "test strategy" OR "test system*" OR "testing assay*" OR "testing batteries" OR "testing battery" OR "testing method*" OR "testing strategies" OR "testing strategy" OR "three dimensional cell culture" OR "three dimensional model" OR "throughput screening"

In silico search

PubMed

("(Q)SAR"[Title/Abstract] OR computational*[Title/Abstract] OR in silico*[Title/Abstract] OR physico chemical propert*[Title/Abstract] OR physicochemical propert*[Title/Abstract] OR QSAR[Title/Abstract] OR read across*[Title/Abstract] OR structural alert*[Title/Abstract] OR structure activit*[Title/Abstract] OR SAR[Title/Abstract] OR SARs[Title/Abstract] OR "(Q)SARs"[Title/Abstract] OR QSARs[Title/Abstract] OR structure toxicity relationship*[Title/Abstract] OR QSTR*[Title/Abstract] OR Structure-Activity Relationship[Majr] OR Computer Simulation[Majr] OR ((Computational Biology[Majr]) AND Computational Biology[mh:noexp]))

WoS

("(Q)SAR" OR "computational*" OR "in silico*" OR "physico chemical propert*" OR "physicochemical propert*" OR "QSAR" OR "read across*" OR "structural alert*" OR "structure activit*" OR "SAR" OR "SARs" OR "(Q)SARs" OR "QSARs" OR "structure toxicity relationship*" OR "QSTR*" OR "Structure-Activity Relationship" OR "Computer Simulation")

Alternative method search 2

PubMed

(sensitivity[Title/Abstract] AND specificity[Title/Abstract]) OR (bioassay*[Title/Abstract] AND (in vitro[Title/Abstract] OR culture*[Title/Abstract]) OR (alternative approach*[Title/Abstract] OR alternative assay*[Title/Abstract] OR alternative method*[Title/Abstract] OR alternative model*[Title/Abstract] OR alternative to animal*[Title/Abstract] OR alternatives to animal*[Title/Abstract] OR analogue approach*[Title/Abstract] OR animalfree[Title/Abstract] OR animal-free[Title/Abstract] OR assay characterisation[Title/Abstract] OR assay characterization[Title/Abstract] OR assay development[Title/Abstract] OR assay performance[Title/Abstract] OR assay validation[Title/Abstract] OR bio assay*[Title/Abstract] OR Biological Assay*[Title/Abstract] OR in vitro screening*[Title/Abstract] OR method characterization[Title/Abstract] OR method characterization[Title/Abstract] OR method development[Title/Abstract] OR method validation[Title/Abstract] OR model characterisation[Title/Abstract] OR model characterization[Title/Abstract] OR model development[Title/Abstract] OR model validation[Title/Abstract] OR non animal alternativ*[Title/Abstract] OR non testing method*[Title/Abstract] OR screening tool*[Title/Abstract] OR screening assay*[Title/Abstract] OR screening method*[Title/Abstract] OR screening system*[Title/Abstract] OR screening test*[Title/Abstract] OR test assay*[Title/Abstract] OR test batteries[Title/Abstract] OR test battery[Title/Abstract] OR test method*[Title/Abstract] OR test strategies[Title/Abstract] OR test strategy[Title/Abstract] OR test system*[Title/Abstract] OR testing assay*[Title/Abstract] OR testing batteries[Title/Abstract] OR testing bat-

tery[Title/Abstract] OR testing method*[Title/Abstract] OR testing strategies[Title/Abstract] OR testing strategy[Title/Abstract] OR throughput screening[Title/Abstract] OR Biological Assay[MeSH] OR toxicological screening[Title/Abstract] OR screening battery[Title/Abstract] OR screening batteries[Title/Abstract] OR screening model*[Title/Abstract] OR alternative testing method[Title/Abstract] OR alternative toxicity testing[Title/Abstract] OR alternative testing[Title/Abstract] OR alternative test method[Title/Abstract] OR novel method*[Title/Abstract] OR novel model*[Title/Abstract] OR novel system*[Title/Abstract] OR reproducibilitly[Title/Abstract] OR assay capacity[Title/Abstract] OR method capacity[Title/Abstract] OR content assay*[Title/Abstract] OR content method*[Title/Abstract] OR content system*[Title/Abstract] OR content model*[Title/Abstract] OR content screening[Title/Abstract] OR toxicity screening[Title/Abstract] OR neurotoxicity screening[Title/Abstract] OR NT screening[Title/Abstract] OR screening model*[Title/Abstract] OR throughput model*[Title/Abstract] OR throughput method*[Title/Abstract] OR throughput system*[Title/Abstract] OR throughput assay*[Title/Abstract] OR throughput test*[Title/Abstract])

WoS

("alternative approach*" OR "alternative assay*" OR "alternative method*" OR "alternative model*" OR "alternative to animal*" OR "alternatives to animal*" OR "analogue approach*" OR "animalfree" OR "animal-free" OR "assay characterisation" OR "assay characterization" OR "assay development" OR "assay performance" OR "assay validation" OR "bio assay*" OR "Biological Assay*" OR "in vitro screening*" OR "method characterization" OR "method characterization" OR "method development" OR "method validation" OR "model characterisation" OR "model characterization" OR "model development" OR "model validation" OR "non animal alternativ*" OR "non testing method*" OR "screening tool*" OR "screening assay*" OR "screening method*" OR "screening system*" OR "screening test*" OR "test assay*" OR "test batteries" OR "test battery" OR "test method*" OR "test strategies" OR "test strategy" OR "test system*" OR "testing assay*" OR "testing batteries" OR "testing battery" OR "testing method*" OR "testing strategies" OR "testing strategy" OR "throughput screening" OR "Biological Assay" OR "toxicological screening" OR "screening battery" OR "screening batteries" OR "screening model*" OR "alternative testing method" OR "alternative toxicity testing" OR "alternative testing" OR "alternative test method" OR "novel method*" OR "novel model*" OR "novel system*" OR "reproducibilitly" OR "assay capacity" OR "method capacity" OR "content assay*" OR "content method*" OR "content system*" OR "content model*" OR "content screening" OR "toxicity screening" OR "neurotoxicity screening" OR "NT screening" OR "screening model*" OR "throughput model*" OR "throughput method*" OR "throughput system*" OR "throughput assay*" OR "throughput test*" OR ("bioassay*" AND ("in vitro" OR "culture*")) OR ("sensitivity" AND "specificity")

BBB search (combined by 'AND')

BBB search

PubMed

BBB[Title] OR brain barrier*[Title] OR brain blood barrier*[Title] OR BBB model*[Title/Abstract] OR brain barrier model*[Title/Abstract] OR brain blood barrier model*[Title/Abstract] OR model BBB[Title/Abstract] OR model biological barrier*[Title/Abstract] OR model brain blood barrier*[Title/Abstract] OR model blood brain barrier*[Title/Abstract]

WoS

"BBB" OR "brain barrier*" OR "brain blood barrier*" OR "BBB model*" OR "brain barrier model*" OR "brain blood barrier model*" OR "model BBB" OR "model biological barrier*" OR "model brain blood barrier*" OR "model blood brain barrier*"

Alternative method search 1 for BBB search

PubMed

(culture*[Title/Abstract] OR cell based[Title/Abstract] OR cell line*[Title/Abstract] OR cell model*[Title/Abstract] OR cell system*[Title/Abstract] OR cellular model*[Title/Abstract] OR cellular assay*[Title/Abstract] OR cellular system*[Title/Abstract] OR cellular method*[Title/Abstract] OR cellular technique*[Title/Abstract] OR cellular endpoint*[Title/Abstract] OR cellular exposure*[Title/Abstract] OR in vitro[Title/Abstract] OR invertebrate*[Title/Abstract] OR non mammal*[Title/Abstract] OR caenorhabditis[Title/Abstract] OR C. elegans[Title/Abstract] OR nematod*[Title/Abstract] OR rerio[Title/Abstract] OR roundworm*[Title/Abstract] OR sea urchin*[Title/Abstract] OR zebra fish[Title/Abstract] OR zebrafish[Title/Abstract] OR Xenopus[Title/Abstract] OR tadpole[Title/Abstract] OR clawed frog*[Title/Abstract] OR Zebrafish[Title/Abstract] OR Caenorhabditis[Title/Abstract] OR Sea Urchins[Title/Abstract] OR Xenopus[Title/Abstract] OR Brachydanio[Title/Abstract] OR Brachydanio rerio[Title/Abstract] OR Danio [Title/Abstract] OR Caenorhabditis elegan*[Title/Abstract] OR Fruitfly[Title/Abstract] OR Fruit fly[Title/Abstract] OR Drosophila[Title/Abstract] OR Drosophila melanog*[Title/Abstract] OR Echinoid*[Title/Abstract] OR fruit flies[Title/Abstract] OR fruitflies[Title/Abstract] OR Drosophila[Title/Abstract] OR (Q)SAR[Title/Abstract] OR computational*[Title/Abstract] OR in silico*[Title/Abstract] OR physico chemical propert*[Title/Abstract] OR physicochemical propert*[Title/Abstract] OR QSAR[Title/Abstract] OR read across*[Title/Abstract] OR structural alert*[Title/Abstract] OR structure activit*[Title/Abstract] OR SAR[Title/Abstract] OR SARs[Title/Abstract] OR (Q)SARs[Title/Abstract] OR QSARs[Title/Abstract] OR structure toxicity relationship*[Title/Abstract] OR QSTR*[Title/Abstract] OR Cell-Free System[MeSH] OR Zebrafish[MeSH] OR Caenorhabditis[MeSH] OR Sea Urchins[MeSH] OR Xenopus[MeSH] OR Drosophila[MeSH] OR In Vitro Techniques[MeSH] OR Cells, Cultured[MeSH] OR Structure-Activity Relationship[Majr] OR Computer Simulation[Majr])))

WoS

"culture*" OR "cell based" OR "cell line*" OR "cell model*" OR "cell system*" OR "cellular model*" OR "cellular assay*" OR "cellular system*" OR "cellular method*" OR "cellular technique*" OR "cellular endpoint*" OR "cellular exposure*" OR "in vitro" OR "invertebrate*" OR "non mammal*" OR "caenorhabditis" OR "C. elegans" OR "nematod*" OR "rerio" OR "roundworm*" OR "sea urchin*" OR "zebra fish" OR "zebrafish" OR "Xenopus" OR "tadpole" OR "clawed frog*" OR "Zebrafish" OR "Caenorhabditis" OR "Sea Urchins" OR "Xenopus" OR "Brachydanio" OR "Brachydanio rerio" OR "Danio " OR "Caenorhabditis elegan*" OR "Fruitfly" OR "Fruit fly" OR "Drosophila" OR "Drosophila melanog*" OR "Echinoid*" OR "fruit flies" OR "fruitflies" OR "Drosophila" OR "(Q)SAR" OR "computational*" OR "in silico*" OR "physico chemical propert*" OR "physicochemical propert*" OR "QSAR" OR "read across*" OR "structural alert*" OR "structure activit*" OR "SAR" OR "SARs" OR "(Q)SARs" OR "QSARs" OR "structure toxicity relationship*" OR "QSTR*"

Alternative method search 2 for BBB search

PubMed

(sensitivity[Title/Abstract] AND specificity[Title/Abstract]) OR ((in vitro[Title/Abstract]) OR culture*[Title/Abstract]) AND bioassay*[Title/Abstract]) OR (alternative approach*[Title/Abstract] OR alternative assay*[Title/Abstract] OR alternative method*[Title/Abstract] OR alternative model*[Title/Abstract] OR alternative to animal*[Title/Abstract] OR alternatives to animal*[Title/Abstract] OR analogue approach*[Title/Abstract] OR animalfree[Title/Abstract] OR animal-free[Title/Abstract] OR assay characterisation[Title/Abstract] OR assay characterization[Title/Abstract] OR assay development[Title/Abstract] OR assay performance[Title/Abstract] OR assay validation[Title/Abstract] OR bio assay*[Title/Abstract] OR Biological Assay*[Title/Abstract] OR in vitro screening*[Title/Abstract] OR method characterization[Title/Abstract] OR method characterization[Title/Abstract] OR method development[Title/Abstract] OR method validation[Title/Abstract] OR model characterisation[Title/Abstract] OR model characterization[Title/Abstract] OR model development[Title/Abstract] OR model validation[Title/Abstract] OR non animal alternativ*[Title/Abstract] OR non testing method*[Title/Abstract] OR screening tool*[Title/Abstract] OR screening assay*[Title/Abstract] OR screening method*[Title/Abstract] OR screening system*[Title/Abstract] OR screening test*[Title/Abstract] OR test assay*[Title/Abstract] OR test batteries[Title/Abstract] OR test battery[Title/Abstract] OR test method*[Title/Abstract] OR test strategies[Title/Abstract] OR test strategy[Title/Abstract] OR test system*[Title/Abstract] OR testing assay*[Title/Abstract] OR testing batteries[Title/Abstract] OR testing bat-

tery[Title/Abstract] OR testing method*[Title/Abstract] OR testing strategies[Title/Abstract] OR testing strategy[Title/Abstract] OR throughput screening[Title/Abstract] OR Biological Assay[Title/Abstract] OR toxicological screening[Title/Abstract] OR screening battery[Title/Abstract] OR screening batteries[Title/Abstract] OR screening model*[Title/Abstract] OR alternative testing method[Title/Abstract] OR alternative toxicity testing[Title/Abstract] OR alternative testing[Title/Abstract] OR alternative test method[Title/Abstract] OR novel method*[Title/Abstract] OR novel model*[Title/Abstract] OR novel system*[Title/Abstract] OR reproducibilitly[Title/Abstract] OR assay capacity[Title/Abstract] OR method capacity[Title/Abstract] OR content assay*[Title/Abstract] OR content method*[Title/Abstract] OR content system*[Title/Abstract] OR content model*[Title/Abstract] OR content screening[Title/Abstract] OR toxicity screening[Title/Abstract] OR neurotoxicity screening[Title/Abstract] OR NT screening[Title/Abstract] OR screening model*[Title/Abstract] OR throughput model*[Title/Abstract] OR throughput method*[Title/Abstract] OR throughput system*[Title/Abstract] OR throughput assay*[Title/Abstract] OR throughput test*[Title/Abstract] OR performance[Title/Abstract] OR validation[Title/Abstract] OR validity[Title/Abstract] OR prediction[Title/Abstract] OR predict[Title/Abstract] OR valid[Title/Abstract] OR reproducibility[Title/Abstract] OR tansport model[Title/Abstract] OR predicting[Title/Abstract] OR reproducible[Title/Abstract] OR fabrication*[Title/Abstract] OR microtechnolog*[Title/Abstract] OR engeneering*[Title/Abstract])

WoS

"alternative approach*" OR "alternative assay*" OR "alternative method*" OR "alternative model*" OR "alternative to animal*" OR "alternatives to animal*" OR "analogue approach*" OR "animalfree" OR "animal-free" OR "assay characterisation" OR "assay characterization" OR "assay development" OR "assay performance" OR "assay validation" OR "bio assay*" OR "Biological Assay*" OR "in vitro screening*" OR "method characterization" OR "method characterization" OR "method development" OR "method validation" OR "model characterisation" OR "model characterization" OR "model development" OR "model validation" OR "non animal alternativ*" OR "non testing method*" OR "screening tool*" OR "screening assay*" OR "screening method*" OR "screening system*" OR "screening test*" OR "test assay*" OR "test batteries" OR "test battery" OR "test method*" OR "test strategies" OR "test strategy" OR "test system*" OR "testing assay*" OR "testing batteries" OR "testing battery" OR "testing method*" OR "testing strategies" OR "testing strategy" OR "throughput screening" OR "Biological Assay" OR "toxicological screening" OR "screening battery" OR "screening batteries" OR "screening model*" OR "alternative testing method" OR "alternative toxicity testing" OR "alternative testing" OR "alternative test method" OR "novel method*" OR "novel model*" OR "novel system*" OR "reproducibilitly" OR "assay capacity" OR "method capacity" OR "content assay*" OR "content method*" OR "content system*" OR "content model*" OR "content screening" OR "toxicity screening" OR "neurotoxicity screening" OR "NT screening" OR "screening model*" OR "throughput model*" OR "throughput method*" OR "throughput system*" OR "throughput assay*" OR "throughput test*" OR "performance" OR "validation" OR "validity" OR "prediction" OR "predict" OR "valid" OR "reproducibility" OR "tansport model" OR "predicting" OR "reproducible" OR "fabrication*" OR "microtechnolog*" OR "engeneering*" OR ("bioassay*" AND ("in vitro" OR "culture*")) OR ("sensitivity" AND "specificity")

MoA search (only PubMed)

(MoA[Title/Abstract] OR mode of action[Title/Abstract] OR key event[Title/Abstract] OR initiating event[Title/Abstract] OR organ effect[Title/Abstract] OR cell effect[Title/Abstract] OR cellular effect[Title/Abstract] OR mechanism*[Title/Abstract])

Appendix F: Results of 'grey' literature search

Appendix F can be found in the online version of this output ('Supporting information' section): http://dx.doi.org/10.2903/j.efsa.2018.1410

Results of the QSAR screening

QSAR Neurotoxicity	
OECD QSAR toolbox	http://www.qsartoolbox.org/
OECD eChemPortal	http://www.echemportal.org/echemportal/index?pageID=0&request_locale=en
Derek Nexus	https://www.lhasalimited.org/derek_nexus/ neurotox endpoint which also covers cholinesterase inhibition
HazardExpert	http://www.compudrug.com/
PASS	http://195.178.207.233/PASS/index.html
Leadscope	http://www.leadscope.com/
ADME	
Lazar	https://lazar.in-silico.de/predict
ADMET Predictor BBB	http://www.simulations-plus.com/
ACD/Labs	http://www.acdlabs.com/products/percepta/predictors.php
Accelrys Accord BBB	http://accelrys.com/solutions/domains/
Grouping and read across for neurotoxicity	
OECD QSAR toolbox	http://www.qsartoolbox.org/
Toxmatch (JRC)	https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/qsar_tools/toxmatch
ChemIDplus (Toxnet)	https://chem.nlm.nih.gov/chemidplus/
AIM (USEPA)	http://www.epa.gov/tsca-screening-tools/analog-identification-methodologyaim-tool
Databases searchable by the endpoint neurotoxicity	
OECD eChemPortal	http://www.echemportal.org/echemportal/index?pageID=0&request_locale=en

Appendix G: Study selection sheet of title/abstract screening

Appendix G can be found in the online version of this output ('Supporting information' section): http://dx.doi.org/10.2903/j.efsa.2018.1410

Appendix H: Data collection sheet

Appendix H can be found in the online version of this output ('Supporting information' section): http://dx.doi.org/10.2903/j.efsa.2018.1410

Appendix I: Tables I1-I5

Table I1: Test Systems grouped by species and Cell Types with respective numbers of citations.

	# citations
human	**130**
immortalized cells	**49**
CHME-5 (microglia)	3
LUHMES	32
mesencephalic Cells (MESC2.10)	9
ReNcell CX cells	5
primary cells	**11**
NSC derived culture	8
primary astrocytes	2
sigmoid colon tissue	1
stem/progenitor cells	**65**
iPSC derived mixed culture (neurons+glia)	8
iPSC derived neurons	49
NPCs	1
undifferentiated neurospheres	8
cell free	**3**
recombinant AChE	3
primary tissue	**2**
neuronal tau-40 protein in phosphate buffer	2
mouse	**173**
immortalized cells	**27**
2.3D (neuroepithelial cells differentiated into astrocytes and neurons)	1
BV-2	15
CRL-2534, astrocyte type III	2
GT1-7 cells (hypothalamic cell line)	2
HT-22	5
N9 microglia	1
SN4741	1
primary cells	**116**
brain slices	1
cerebellar granule cell	3
cerebellar granule neurons	3
dorsal root ganglia/spinal cord cultures	3
isolated mouse hemidiaphragm muscles	1
mixed neuron and glia culture	15
mixed culture (dopaminergic neurons+astrocytes)	22
mixed culture (neurons+microglia)	1

mixed culture (spinal cord - skeletal muscle)	1
murine brain microvascular endothelial cells	1
neuron/astrocyte contact co-culture	1
primary glia	3
primary neurons	62
stem/progenitor cells	**19**
ESC	3
ESC derived glutamatergic neurons	9
ESC derived neurons	7
cell free	**3**
recombinant AChE	3
primary tissue	**8**
brain homogenate	2
cell membrane	2
mitochondria	3
synaptosomes	1
rat	**567**
Immortalized cells	**42**
E18 neuroblast	2
HAPI	8
N27	27
RBE4	5
primary cells	**479**
astrocyte rich culture	11
brain slices	33
cell membrane	4
cerebellar granule cell	13
cerebellar granule neurons	9
dopaminergic neurons	1
hippocampal CA1 pyramidal neurons	1
Oligodendrocyte progenitors	1
primary glia	55
primary microglia	7
primary neurons)	194
primary oligodendrocytes	1
purkinje neurons	2
re-aggregating brain cell cultures	59
trigeminal ganglion neurons	1
mixed neuron and glia cultures	84
Dorsal root ganglia	3
stem/progenitor cells	**7**
differentiated mesencephalic NPCs	4
differentiated striatal neural NPCs	3
tumor cells	**2**

PC6-3	2
primary tissue	**114**
brain homogenate	17
cell membrane	22
microsome	8
mitochondria	18
mitochondrial suspension from liver	3
synaptosomal mitochondria	5
synaptosomes	41
xenopus	**10**
primary cells	**10**
neurolemma in Xenopus Oocytes	6
sodium channels in Xenopus Oocytes	4
chicken	**18**
primary cells	**16**
cerebellar bergmann glia	5
primary neurons	10
spheroids	1
primary tissue	**2**
brain homogenate	2

Table I2: Total number of citations for endpoint categories grouped according to species. In addition, predictivity analyses were performed for each endpoint category and species giving the true as well as false positives and negatives.

	f. n.	f. p.	t. n.	t. p.	total
Chicken	**3**			**15**	**18**
Activation of Chloride Channels				1	1
Delayed Neuropathy				1	1
Enzyme Inhibition				10	10
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	2			1	4
Other	1				1
Stimulation of Cholinergic Neurotransmission				1	1
Human	**12**	**8**	**22**	**88**	**130**
Axonopathies	1			1	2
Cytoskeletal Alterations				4	4
Inhibition of Cholinergic Neurotransmission				2	2
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	11			60	72
negative		8	22		30
Neuroinflammation				1	1
Other				1	1
Redox Cycling				14	14
Stimulation of Cholinergic Neurotransmission				4	4
Stimulation of dopaminergic Neurotransmission				1	1
Mouse	**9**		**2**	**162**	**173**
Activation of Sodium Channels				9	9
Altered Calcium Signaling				2	2
Axonopathies				2	2
Effects on Other Neuronal Receptors				2	2
Enzyme Inhibition				1	1
Inhibition of Chloride Channels				1	1
Inhibition of Cholinergic Neurotransmission				16	16
Inhibition of dopaminergic Neurotransmission				1	1
Inhibition of GABAergic Neurotransmission				7	7
Inhibition of Glycinergic Neurotransmission				3	3
Inhibition of Sodium Channels				2	2
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	6			73	79
negative			2		2
Neuroinflammation				6	6
Neurotransmission in General				1	1
Other				4	4
Redox Cycling	2			11	13
Stimulation of Cholinergic Neurotransmission	1			14	15
Stimulation of dopaminergic Neurotransmission				1	1
Stimulation of GABAergic Neurotransmission				1	1
Stimulation of Glutamatergic Neurotransmission				6	6
Rat	**85**	**8**	**22**	**529**	**644**

Activation of Sodium Channels	5			24	29
Altered Calcium Signaling	2			18	20
Axonopathies				15	15
Cytoskeletal Alterations	4			1	5
Effects on Other Neuronal Receptors	3			9	12
Enzyme Inhibition				4	4
Inhibition of Adrenergic Neurotransmission				2	2
Inhibition of Cholinergic Neurotransmission	3	2		4	9
Inhibition of dopaminergic Neurotransmission				8	8
Inhibition of GABAergic Neurotransmission	5			33	38
Inhibition of Glycinergic Neurotransmission				4	4
Inhibition of Sodium Channels				1	1
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	32		2	235	269
Myelin Toxicity				2	2
negative		1	15		16
Neuroinflammation	11			22	33
Neurotransmission in General	1			7	8
Other				4	4
Redox Cycling	7		1	61	69
Stimulation of Cholinergic Neurotransmission	9	3	4	50	66
Stimulation of dopaminergic Neurotransmission				11	11
Stimulation of GABAergic Neurotransmission				2	2
Stimulation of Glutamatergic Neurotransmission	3			12	15
Stimulation of Adrenergic Neurotransmission		1			1
Stimulation of Serotonergic Neurotransmission		1			1
Xenopus				**10**	**10**
Activation of Sodium Channels				6	6
Inhibition of Calcium Channels				1	1
Inhibition of Chloride Channels				1	1
Inhibition of Sodium Channels				2	2
total	**109**	**16**	**46**	**806**	**977**

Table I3: Total number of citations for endpoint categories grouped according to human test systems. Predictivity analyses were performed for each endpoint category, species and test system giving the true as well as false positives and negatives.

	f. n.	f. p.	t. n.	t. p.	total
cell free				**3**	**3**
recombinant AChE				**3**	**3**
Stimulation of Cholinergic Neurotransmission				3	3
Immortalized cells	**2**			**47**	**49**

CHME-5 (microglia)				**3**	**3**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				3	3
LUHMES	**2**			**30**	**32**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	2			30	32
mesencephalic cells (MESC2.10)				**9**	**9**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				2	2
Redox Cycling				6	6
Stimulation of dopaminergic Neurotransmission				1	1
ReNcell CX cells				**5**	**5**
Redox Cycling				5	5
primary cells				**11**	**11**
NSC derived culture				**8**	**8**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				8	8
primary glia				**2**	**2**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				2	2
sigmoid colon tissue				**1**	**1**
Stimulation of Cholinergic Neurotransmission				1	1
primary tissue				**2**	**2**
neuronal tau-40 protein in phosphate buffer				**2**	**2**
Cytoskeletal Alterations				2	2
stem/progenitor cells	**10**	**8**	**22**	**25**	**65**
iPSC derived mixed culture (neurons+glia)	**2**			**6**	**8**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	2			5	7
Neuroinflammation				1	1
iPSC derived neurons	**5**	**8**	**22**	**14**	**49**
Axonopathies	1			1	2
Cytoskeletal Alterations				2	2
Inhibition of Cholinergic Neurotransmission				2	2
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	4			5	9
negative		8	22		30
Other				1	1
Redox Cycling				3	3
NPCs				**1**	**1**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				1	1
undifferentiated neurospheres	**3**			**5**	**8**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	3			5	8
total	**12**	**8**	**22**	**88**	**130**

Table I4: Total number of citations for endpoint categories grouped according to mouse test systems. Predictivity analyses were performed for each endpoint category, species and test system giving the true as well as false positives and negatives.

	f. n.	**t. n.**	**t. p.**	**total**

cell free			**3**	**3**
recombinant AChE			**3**	**3**
Stimulation of Cholinergic Neurotransmission			3	3
Immortalized cells	**2**		**25**	**27**
2.3D (neuroepithelial cells differentiated into astrocytes and neurons)			**1**	**1**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			1	1
BV-2	**2**		**13**	**15**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			10	10
Redox Cycling	2		3	5
CRL-2534, astrocyte type III			**2**	**2**
Stimulation of Glutamatergic Neurotransmission			2	2
GT1-7 cells (hypothalamic cell line)			**2**	**2**
Axonopathies			1	1
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			1	1
HT-22			**5**	**5**
Redox Cycling			5	5
N9 microglia			**1**	**1**
Redox Cycling			1	1
SN4741			**1**	**1**
Redox Cycling			1	1
primary cells	**7**	**2**	**108**	**117**
brain slices			**1**	**1**
Stimulation of Glutamatergic Neurotransmission			1	1
cerebellar granule cell			**3**	**3**
Inhibition of GABAergic Neurotransmission			3	3
cerebellar granule neurons			**3**	**3**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			2	2
Stimulation of Cholinergic Neurotransmission			1	1
dorsal root ganglia/spinal cord cultures			**2**	**2**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			2	2
isolated mouse hemidiaphragm muscles			**1**	**1**
Inhibition of Cholinergic Neurotransmission			1	1
mixed culture (dopaminergic neurons+astrocytes)	**5**		**17**	**22**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	5		17	22
mixed culture (neurons+microglia)			**1**	**1**

Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			1	1
mixed culture (spinal cord - skeletal muscle)			**1**	**1**
Inhibition of Cholinergic Neurotransmission			1	1
mixed neuron and glia cultures	**1**		**14**	**15**
Inhibition of Cholinergic Neurotransmission			1	1
Inhibition of Sodium Channels			2	2
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	1		2	3
Neuroinflammation			1	1
Stimulation of Cholinergic Neurotransmission			6	6
Stimulation of dopaminergic Neurotransmission			1	1
Stimulation of GABAergic Neurotransmission			1	1
murine brain microvascular endothelial cells			**1**	**1**
Enzyme Inhibition			1	1
neuron/astrocyte contact co-culture			**1**	**1**
Neuroinflammation			1	1
primary glia			**3**	**3**
Neuroinflammation			3	3
primary neurons	**1**	**2**	**59**	**62**
Activation of Sodium Channels			9	9
Altered Calcium Signaling			1	1
Axonopathies			1	1
Inhibition of Chloride Channels			1	1
Inhibition of GABAergic Neurotransmission			4	4
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			34	34
Redox Cycling			1	1
negative		2		2
Neuroinflammation			1	1
Neurotransmission in General			1	1
Other			2	2
Stimulation of Cholinergic Neurotransmission	1		1	2
Stimulation of Glutamatergic Neurotransmission			3	3
dorsal root ganglia			**1**	**1**
Inhibition of Cholinergic Neurotransmission			1	1
primary tissue			**8**	**8**
brain homogenate			**2**	**2**
Stimulation of Cholinergic Neurotransmission			2	2
cell membrane			**2**	**2**
Effects on Other Neuronal Receptors			2	2
mitochondria			**3**	**3**

Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			3	3
synaptosomes			**1**	**1**
Inhibition of dopaminergic Neurotransmission			1	1
stem/progenitor cells			**19**	**19**
ESC			**3**	**3**
Altered Calcium Signaling			1	1
Other			2	2
ESC derived glutamatergic neurons			**9**	**9**
Inhibition of Cholinergic Neurotransmission			8	8
Stimulation of Cholinergic Neurotransmission			1	1
ESC derived neurons			**7**	**7**
Inhibition of Cholinergic Neurotransmission			4	4
Inhibition of Glycinergic Neurotransmission			3	3
total	**9**	**2**	**163**	**174**

Table I5: Total number of citations for endpoint categories grouped according to rat test systems. Predictivity analyses were performed for each endpoint category, species and test system giving the true as well as false positives and negatives.

	f. n.	f. p.	t. n.	t. p.	total
Immortalized cells	**4**			**38**	**42**
E18 neuroblast				**2**	**2**
Redox Cycling				2	2
HAPI				**8**	**8**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				6	6
Neuroinflammation				2	2
N27	**4**			**23**	**27**
Altered Calcium Signaling				1	1
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				11	11
Redox Cycling	4			11	15
RBE4				**5**	**5**
Redox Cycling				5	5
primary cells	**65**	**8**	**22**	**384**	**479**
astrocyte rich culture	**1**			**10**	**11**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	1			8	9
Other				2	2
brain slices	**3**		**1**	**29**	**33**
Activation of Sodium Channels				2	2
Cytoskeletal Alterations				1	1

Inhibition of Adrenergic Neurotransmission			1	1
Inhibition of Cholinergic Neurotransmission	3			3
Inhibition of GABAergic Neurotransmission			2	2
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis		1	4	5
Redox Cycling			5	5
Stimulation of Cholinergic Neurotransmission			4	4
Stimulation of dopaminergic Neurotransmission			8	8
Stimulation of Glutamatergic Neurotransmission			2	2
cell membrane			**4**	**4**
Inhibition of GABAergic Neurotransmission			4	4
cerebellar granule cell	**2**		**11**	**13**
Inhibition of GABAergic Neurotransmission			4	4
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	2		4	6
Redox Cycling			3	3
cerebellar granule neurons			**9**	**9**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			7	7
Stimulation of Cholinergic Neurotransmission			1	1
Stimulation of Glutamatergic Neurotransmission			1	1
dopaminergic neurons			**1**	**1**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			1	1
hippocampal CA1 pyramidal neurons			**1**	**1**
Stimulation of Cholinergic Neurotransmission			1	1
oligodendrocyte progenitors			**1**	**1**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			1	1
primary glia	**16**		**39**	**55**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	4		18	22
Neuroinflammation	11		9	20
Redox Cycling	1		12	13
primary microglia	**2**		**5**	**7**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	2		3	5
Neuroinflammation			2	2
primary neurons	**22**	**9**	**163**	**194**
Activation of Sodium Channels			13	13
Axonopathies			5	5
Cytoskeletal Alterations	4			4
Inhibition of Adrenergic Neurotransmission			1	1
Inhibition of Cholinergic Neurotransmission			3	3
Inhibition of dopaminergic Neurotransmission			1	1
Inhibition of GABAergic Neurotransmission			15	15
Inhibition of Glycinergic Neurotransmission			2	2
Inhibition of Sodium Channels			1	1
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	9		88	97

negative			9		9
Neuroinflammation				6	6
Neurotransmission in General				2	2
Other				1	1
Redox Cycling	2			11	13
Stimulation of Cholinergic Neurotransmission	5			9	14
Stimulation of Glutamatergic Neurotransmission	2			5	7
primary oligodendrocytes				**1**	**1**
Myelin Toxicity				1	1
purkinje neurons				**2**	**2**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				2	2
re-aggregating brain cell cultures	**6**	**7**	**4**	**42**	**59**
Axonopathies				2	2
Inhibition of Cholinergic Neurotransmission		2		1	3
Inhibition of GABAergic Neurotransmission				2	2
Inhibition of Glycinergic Neurotransmission				2	2
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	3			8	11
Myelin Toxicity				1	1
Neuroinflammation				2	2
Neurotransmission in General	1			3	4
Other				1	1
Redox Cycling				1	1
Stimulation of Cholinergic Neurotransmission	2	3	4	18	27
Stimulation of Glutamatergic Neurotransmission				1	1
Stimulation of Adrenergic Neurotransmission		1			1
Stimulation of Serotoninergic Neurotransmission		1			1
trigeminal ganglion neurons				**1**	**1**
Stimulation of Cholinergic Neurotransmission				1	1
mixed neuron and glia cultures	**13**	**1**	**8**	**62**	**84**
Activation of Sodium Channels	5			3	8
Altered Calcium Signaling				5	5
Inhibition of dopaminergic Neurotransmission				1	1
Inhibition of GABAergic Neurotransmission	5			6	11
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis			1	32	33
negative		1	6		7
Neuroinflammation				1	1
Redox Cycling			1	7	8
Stimulation of Cholinergic Neurotransmission	2			4	6
Stimulation of dopaminergic Neurotransmission				1	1
Stimulation of GABAergic Neurotransmission				1	1
Stimulation of Glutamatergic Neurotransmission	1			1	2
dorsal root ganglia				**3**	**3**

Axonopathies				3	3
stem/progenitor cells				**7**	**7**
differentiated mesencephalic NPCs				**4**	**4**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				3	3
Redox Cycling				1	1
differentiated striatal neural NPCs				**3**	**3**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				2	2
Redox Cycling				1	1
primary tissue	**16**			**98**	**114**
brain homogenate				**17**	**17**
Enzyme Inhibition				4	4
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				11	11
Stimulation of Cholinergic Neurotransmission				2	2
cell membrane	**5**			**17**	**22**
Altered Calcium Signaling	2				2
Effects on Other Neuronal Receptors	3			9	12
Stimulation of Cholinergic Neurotransmission				8	8
microsome				**8**	**8**
Altered Calcium Signaling				8	8
mitochondria	**6**			**12**	**18**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	6			10	16
Redox Cycling				2	2
mitochondrial suspension from liver	**3**				**3**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	3				3
synaptosomal mitochondria				**5**	**5**
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis				5	5
synaptosomes	**2**			**39**	**41**
Activation of Sodium Channels				6	6
Altered Calcium Signaling				4	4
Axonopathies				5	5
Inhibition of dopaminergic Neurotransmission				6	6
Mitochondrial Dysfunction/Oxidative Stress/Apoptosis	2			9	11
Neurotransmission in General				2	2
Stimulation of Cholinergic Neurotransmission				2	2
Stimulation of dopaminergic Neurotransmission				2	2
Stimulation of GABAergic Neurotransmission				1	1
Stimulation of Glutamatergic Neurotransmission				2	2
total	**85**	**8**	**22**	**527**	**642**

Appendix J: Performance criteria to define the readiness of test methods for hazard evaluation.

Criteria	Description	Examples / Why is it important	Max. score
1 Test system		*Note: here scoring not for 'test method'*	**10**
1a What is modelled	Is there a clear rationale given for what target organ/tissue relevant for human poisoning/pathology the test systems should reflect	Here: question is not for relevance, but whether there is documentation and a rationale at all.	1
1b Relevance	Is the chosen test system known to be a key component in pathogenesis, or why is it thought to reflect a key component, mechanism or tissue	Here: is the tissue/organ modelled important for regulatory toxicology or biomedical research purposes. Is evidence given for the relevance of the model by morphological comparison, gene expression or functional criteria? Are all/sufficient cell types included in the model?	1
1c System uncertainties and human correlate (HC)	(i) Is there a discussion on where the test system differs from the mimicked human tissue, and which gaps of analogy need to be considered? (ii) Do toxicant-altered genes (or other biomarkers) correspond to changes in mimicked human tissue (after poisoning or in relevant pathologies)	(i) E.g. a differentiated cell or a cell line (such as HepG2) does not necessarily reflect all features of the corresponding in vivo tissue/conditions. (ii) This is an additional measure to increase confidence in the test; not mandatory, but helpful.	1
1d Definition of cells	Is the test system sufficiently characterized (source; multiple positive and negative markers for cell identity, number, quality, composition, differentiation state, viability, usual morphology, basic function, basic reaction to stimuli, STR...)	This is especially important for cells that have to be produced regularly, e.g. by differentiation or primary cell isolation.	1
1e Cell composition	For multi-component systems: information on all cellular subpopulations. What is the percentage of contaminating cells or in co-cultures what is the percentage of all subpopulations.	This is important for the test endpoints as it could be that only one cell type may be affected by a toxicant. For primary cells: have cells from different sources (suppliers) been tested (e.g. hepatocytes from different suppliers may differ in purity and quality). For routine use it would be beneficial to have pre-set acceptance criteria for each cell type	1
1f Cellular environment	Information on structuring components of the test system: coating, scaffolds, matrix description, medium (supplements), microfluidic effects, supportive cells, dimensions and positioning/handling of 3D constructs,....	This means a very detailed description of the culture conditions, including temporal and spatial aspects. Cell differentiation and response (quality, quantity, kinetics) may depend on multiple external factors and on the 3D arrangement	1
1g Biological consistency	(i) Has the variation of the test system been assessed, influencing factors identified? (ii) Have acceptance criteria and performance standards for the test system been defined (different from the test!)?	(i) E.g. do medium supplements have an influence on the outcome of the cells; such as batch effects of FCS or serum replacement additives? (ii) e.g. a range of marker expression levels, of biological function (proliferation, protein production,...), of structural features (cell number, organoid size,...),... For lines: what is the optimum passage number of cells? . For routine use it would be beneficial to have pre-set acceptance criteria for the whole model/test system	1
1h Critical components	Have critical components and handling steps been identified and described? Are examples for normal performance and morphology given; are there examples for alerts?	E.g. cell density on a specific day of differentiation could be a critical step; wrong, strange morphology of cells could be an alert.). For routine use it would be beneficial to have pre-set acceptance criteria.	
1i Cell stability	Stability proven over multiple doublings; genetic stability shown; pluripotency/multipotency (for stem cells) shown, cell identity shown	For stem cells, stability needs to be shown over many passages (≥10). For primary cells: stability and identity of supply needs to be shown; stability of function (e.g. xenobiotic metabolism) shown.	1
1j Transgenic cells	Transgene characterized (source, sequence, regulation); insertion characterized; stability of function shown and quantified; cell identity and function related to wt; clonality documented.		*1*
2 Exposure scheme			**3**
2a Description	Complete, detailed, unambiguous.	Medium changes, re-additions, coating, treatment period and timing, incubation conditions (temp. gasing,..)	1
2b Unique identity	Tests with multiple variants of a test need to define very transparently, which variant the data come from	E.g. from which cell type/clone; which time; which plate format; which medium additives...	1
2c Graphical scheme	Complete sequence of events, including endpoint assessment	Supports clarity and data assignment to test variants	1
3 Documentation / SOP			**5**

3a Availability	Method description for test system, test procedure, analytical endpoints and prediction model; public availability of SOP (data bank or test developer upon request)	Normal scientific publications are usually not sufficient, unless it is a specific methods paper. For transferability of the test method it is beneficial to have SOPs or other documents covering each component of test method and the whole testing process	*1*
3b Stage of development	Version history; updated		1
3c For CRO tests	Are full performance standards and corresponding data delivered by the CRO along with test data (in case SOP details are not disclosed)	Non-disclosure of SOP is acceptable, if full performance/readiness criteria are given.	*1*
3d Test components	Documented and available (receipt, storage, handling and disposal documents); quality criteria and checking procedure established	E.g. for media, plates, coating it should be defined, what is acceptable/non-acceptable and how this is controlled. Test chemical identity and purity (certificate of analysis) and safety data sheets from chemicals	1
3e Stocks	Procedure for preparation, storage and quality control of stocks established		1
4 Main endpoint(s)		*Mainly referring to specific/functional endpoints*	**4**
4a Biol. relevance	Is there a rationale given why test endpoint is relevant to adverse outcomes	Helps to interpret the results obtained.	1
4b Toxicological relevance	Are toxicants (≥ 3) known to affect the endpoint	Helps to interpret the results obtained.	1
4c Analytical methods	Methods defined, rationale given; positive controls and acceptability criteria	Positive controls for analytical method may differ from controls for test/endpoint	1
4d Multiple endpoints	Are all endpoints and their relation to one another (priority, preference) defined	E.g. neurite outgrowth / cytotoxicity	1
5 Cytotoxicity		*Here: if cytotoxicity is not main endpoint*	**5**
5a Cytotoxicity within test	Cytotoxicity is preferentially determined within same test compartment as the major endpoint; second choice is under same conditions in parallel	Control of cytotoxicity in a different format (e.g. other types of plates; other time are very problematic). Measuring cytotoxicity under the same test conditions as the main end point help to interpret the mechanism related to the adverse effects for the main end point (specific or cytotoxicity driven mechanism)	1
5b Subpopulation effects	Are subpopulations detected by measure for cytotoxicity or proliferation; are minor changes detected? Has sensitivity been shown?	Usually at least three types of assay required (measurement of viability, measurement of cell death, single cell analysis)	0.5
5c Specificity (compared to cytotox)	A measure needs to be established to distinguish a specific/functional endpoint from cytotoxicity	E.g. neurite outgrowth, migration inhibition in non-cytotoxic concentration ranges	0.5
5d Timing within test	For repeated/prolonged dosing, early death and compensatory growth need to be considered	The test of cytotoxicity only at the end may give false negative data, if cells die early and this is not detectable late, because of compensatory proliferation.	*0.5*
5e Timing after test	For very short endpoints, e.g. electrophysiology measured 30 min after toxicant exposure, delayed measure of cytotoxicity is necessary	Cells cannot die in very short time, even though compound triggers lethal changes. Data for 24h exposure should be given.	*0.5*
5f Curve fitting	Sufficient non-toxic data points (baseline); at least 40% toxicity / change to allow fitting		0.5
5g Non-cytotoxicity	Absence of 'cytotoxicity' does not mean non-cytotoxicity (question of power): has data variation been considered; is a measure of uncertainty given for non-cytotoxicity (e.g. BMCL calculation)?		0.5
5h Bench mark response	Has a rationale been given for setting a threshold value for cytotoxicity (statistical or biological significance)	E.g. statistical: 3x standard deviation; biological: 90% viability; see also: http://invitrotox.uni-konstanz.de/	0.5
5i Apoptosis/ Proliferation	If natural feature of the test system: measure for normal rate required		*0.5*
6 Test method controls			**4**
6a Positive controls (PC)	≥ 3 toxicants required for test definition; preferentially of different mechanisms; preferentially human-relevant toxicants; indicate variation of PC within and across assays	Used to define acceptability criteria, S/N ratio or z'-value of screen	1
6b Negative controls (NC)	≥ 5 negative controls are required to define specificity at ±20% level; concentration of negatives needs to be defined and rationalized	Ways to define negatives: (i) e.g. compound only acting when metabolized, (ii) acting on another organ, (iii) known to be safe for pregnant women, (iv) being selective for another assay, (v) pairs/matches of a specific positive control (e.g. inactive metabolite)	1
6c Unspecific controls (UC)	A type of negative control for functional assays: not inactive, but only cytotoxic	Absolutely essential to define baseline variation and thus the relevant benchmark response for positive hits	1

6d Endpoint-specific controls (EC)	To provide plausibility, and to help initial test setup: EC show that pathways considered to be relevant for test endpoint are indeed affecting the test endpoint. EC help to correlate (by concentration and time) compound effect on pathway (activity measure to be established) and on test endpoint (standard test readout). EC may be chemicals or siRNA; pathways may be defined from literature or experimentally (gene expression)	Example: actin is required for migration, thus an actin inhibitor should affect migration endpoint	1
7 Data evaluation		*Here: referring to main endpoint(s)*	4
7a Outliers	Procedure for handling and documentation should be established		1
7b Concentration - dependence	Higher confidence in concentration-dependent data; no-effect concentrations must be included (full range curve); data need sufficiently dense spacing around benchmark concentration; preferably provide statistical significance for key data points		1
7c Benchmark response	Give rationale for definition (statistical (after FDR correction) or biological). Provide power estimate if conclusions are drawn from negatives.		1
7d Curve fitting	Indicate detailed procedure used for curve fitting; preferentially force fitted curve through 100% at negative control conditions (full function)	E.g. sigmoidal, linear or exponential curve fit	
8 Testing strategy			4
8a Hazard prediction	Which hazard is assessed; which question does the test method answer		1
8b Link to an AOP	Does the test give input to a mechanistic concept, e.g. an AOP	Helps to position in battery; helps to interpret results	1
8c Role in battery	Full score for stand alone tests. For tests that are not stand alone, information on their relation to other tests in a battery is required.	Information is required on how the test data would be used in a battery and under which conditions this is possible.	1
8d Comparison to similar tests	Does the test fill a gap in a battery? Is it providing advantages compared to another test for the same hazard?	Avoid overlapping tests to be performed. Ensure adequate testing battery/strategy	1
9 Robustness			4
9a Reproducibility	Data available on normal variation; Information on factors affecting test variation is given	Historic control data on positive controls show normal range; known artefacts and shortcomings	1
9b Intra-lab	Data available from different operators, different test runs over longer time		1
9c Inter-lab	Data available on transferability / reproducibility in another lab		1
9d Historical controls	Data for PC and NC over time		1
10 Test benchmarks			4
10a Sensitivity (of the test)	Signal noise ratio (S/N) defined. Sensitivity information available	S/N based on adequate data sets. The S/N is used to determine the limit of detection. Additional measures: True positive rate, hit rate; sensitivity to detect a panel of positive controls, etc...	1
10b Specificity (of the test)	Tested with sufficient number and quality of negative controls	Additional measures: true negative rate, etc.	1
10c Acceptance criteria	Clearly defined and documented. Normal range of variation known	E.g. a given positive control has to reduce the main endpoint by at least 25%, otherwise test plate is discarded.	1
10d Response characteristics	Should the response be linear? What are the upper and lower limits?	Additional measures: mono-directional or bi-directional deviation defined; Info on accuracy, precision, limit of quantification, etc.	1
11 Prediction model			4
11a Definition	Information should be available and clear; (including rationale for model, i.e. its particular strengths)Information and rationale should be given for use of sharp thresholds or probabilistic approach.	Information on how many classes of toxicants are predicted. Positives and non-positives; or strong, medium, weak positives. Information on uncertainty of prediction should be given, at least for positives (note that uncertainty of negatives is often not defined). E.g. you can define a sharp threshold all above 4 is positive or you can define above 4 has a 70% likelihood to be positive	1

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

Appendix K: Readiness check (detailed scoring)

Appendix K can be found in the online version of this output ('Supporting information' section): http://dx.doi.org/10.2903/j.efsa.2018.1410