

**Agency Issue Paper:**

**Use of New Approach Methodologies to Derive  
Extrapolation Factors and Evaluate Developmental  
Neurotoxicity for Human Health Risk Assessment**

**July 2020**

## Table of Contents

<b>1.0</b>	<b>Introduction</b> .....	5
<b>1.1</b>	<b>Regulatory Context</b> .....	5
<b>1.2</b>	<b>Scope of the Draft Issue Paper</b> .....	6
<b>2.0</b>	<b>Fit-for-purpose alternatives to the <i>in vivo</i> DNT study</b> .....	7
<b>2.1</b>	<b>Background &amp; History on the DNT</b> .....	7
<b>2.2</b>	<b>Background &amp; History of the OPs</b> .....	8
<b>2.3</b>	<b>New approach methodologies (NAMs) for DNT</b> .....	10
<b>2.3.1</b>	<b>Overview of the DNT-NAMs</b> .....	11
<b>2.3.2</b>	<b>DNT-NAM development as an integrated process</b> .....	19
<b>2.3.3</b>	<b>DNT-NAM experimental methodology for OPs</b> .....	22
<b>2.3.4</b>	<b>DNT-NAM experimental validation using assay positive controls</b> .....	26
<b>2.3.5</b>	<b>DNT-NAM experimental results for OPs</b> .....	37
<b>2.3.6</b>	<b>DNT-NAM administered equivalent doses versus rat benchmark doses</b> .....	50
<b>2.3.7</b>	<b>Future research directions to increase utility of DNT-NAMs</b> .....	64
<b>2.3.8</b>	<b>Summary and conclusions on the DNT-NAMs</b> .....	65
<b>3.0</b>	<b>Development of DDEFs for Interspecies and Intraspecies Extrapolation</b> .....	67
<b>3.1</b>	<b>Background</b> .....	67
<b>3.2</b>	<b>Methods</b> .....	69
<b>3.2.1</b>	<b>Laboratory Experiments</b> .....	69
<b>3.2.2</b>	<b>Statistical Approach</b> .....	70
<b>3.3</b>	<b>Results</b> .....	72
<b>3.3.1</b>	<b>Interspecies data analyses</b> .....	72
<b>3.3.2</b>	<b>Intraspecies data analyses</b> .....	73
<b>3.4.2</b>	<b>Intraspecies Pharmacodynamic DDEFs</b> .....	76
<b>4.0</b>	<b>Summary &amp; Next Steps</b> .....	76
<b>5.0</b>	<b>References</b> .....	78

## Executive Summary

To supplement or replace the existing toxicity tests used to support pesticide registration, EPA's OPP is actively engaged in numerous activities with respect to reducing laboratory animal use and implementing new approach methodologies (NAMs). NAM is a broad term referring to any non-animal technology, methodology, approach, or combination thereof that can be used to provide information on chemical hazard and risk assessment. As NAMs can provide human relevant information that may be challenging to test in whole animals, EPA's OPP is also interested in using NAMs to reduce the reliance on default assumptions for risk assessment. In line with this, several years ago, the Office of Pesticide Programs (OPP) began collaborative work with the Office of Research and Development (ORD), academia, and industry to use the organophosphate pesticides (OPs) as a case study for the development of NAMs to inform extrapolation/uncertainty and safety factors in lieu of reliance on default factors. Two approaches related to these efforts are presented.

The first approach presents work completed by ORD towards developing a battery of NAMs for fit-for-purpose evaluation of developmental neurotoxicity (DNT), using OPs as a case study. This includes a microelectrode array network formation assay (MEA NFA) and high-content imaging (HCI) assays of neural cells to understand key processes relevant to neurodevelopment. These assays were found to be robust and reproducible for fit-for-purpose application to identify putative DNT-related bioactivity, with reasonable assay performance controls established. The testing results for 27 OPs found differential activity in the MEA NFA and HCI assay suite indicating a difference in DNT potential across individual chemicals.

Common to all OPs is the ability to inhibit the AChE enzyme, which prevents the breakdown of acetylcholine leading to neurotoxicity. AChE inhibition is the basis of current OP human health risk assessments. In order to compare the relative sensitivity of the MEA NFA and HCI assay results to doses that inhibit AChE in laboratory animals, *in vitro* to *in vivo* extrapolation (IVIVE) approaches using high-throughput toxicokinetic (HTTK) models were utilized to approximate NAM-derived administered equivalent doses (AEDs). These comparisons demonstrated that NAM-derived AEDs were greater than or in some cases approximated doses that inhibit AChE; however, additional information is still needed for to make conclusions regarding the relative sensitivity of individual chemicals.

The IVIVE approach used for these comparisons is driven by the chemical-specific data that are available. Due to the lack of rat clearance data for many of the OPs tested, the use of human data in the rat model was evaluated. Predictions were not found to be uniformly higher or lower when using human rather than rat data. Therefore, for chemicals that are lacking rat clearance data and have AEDs approximating the BMD/BMDL values, it is unknown whether the predictions are over- or underestimates. Pending the recommendations from the SAP and the overall weight of evidence evaluation of DNT potential based on the currently available data, the OPP may determine that rat *in vitro* clearance data for particular OPs are needed to better characterize these comparisons and OPP may reach out to registrants in the future to generate these data.

The MEA NFA and HCI assay suite evaluates the majority of the critical processes of neurodevelopment, including neural network formation and function, cell proliferation, apoptosis, neurite outgrowth, and synaptogenesis, and represents a significant advancement toward developing a NAM battery for fit-for-purpose DNT evaluation. Additional assays are currently under development by researchers funded by the European Food Safety Authority (EFSA) that evaluate processes not covered by the MEA NFA and HCI

assays, such as migration and differentiation. The relative contribution of these assays for a fit-for-purpose NAM battery will be considered once data are available. Additionally, any OP data from these assays may be considered in the future in combination with the results of the MEA NFA and HCI assays as part of an overall weight of evidence evaluation of the DNT potential for individual OPs. Given the gaps in rat clearance data and additional DNT NAMs under development, EPA cannot make any conclusions about the impact of these approaches on the OP human health risk assessments at this time.

The second approach presents *in vitro* data generated by academia on behalf of pesticide registrants to calculate pharmacodynamic data derived extrapolation factors (DDEFs) for 16 OPs in accordance with the EPA's 2014 guidance on DDEFs. EPA is soliciting comment from the SAP on the study design and methods utilized to generate the *in vitro* data, statistical analyses employed to calculate DDEFs, and analyses performed for a small subset of OPs to evaluate contributions of experimental and intrinsic variability. For both interspecies and intraspecies DDEFs, there were no statistical issues identified for 9 chemicals. EPA's preliminary proposal is to use the DDEFs calculated for these 9 chemicals (bensulide oxon, chlorethoxyfos, DDVP, ethoprop, fenamiphos, methamidophos, phorate oxon sulfoxide, phosmet oxon, and terbufos oxon sulfone) in lieu of the default pharmacodynamic uncertainty factors, pending review by the SAP. For the remaining chemicals, SAP recommendations regarding warning statements and model fit issues will be considered before determining the potential use of the calculated DDEFs. Ultimately, the selection and application of DDEFs for each chemical is a policy decision that will be made by EPA.

## 1.0 Introduction

In 2007, the National Research Council's report on *Toxicity Testing in 21<sup>st</sup> Century: A Vision and a Strategy* was published which encouraged a new paradigm in testing that reduces the reliance on using laboratory animals while moving towards *in vitro* and *in silico* predictive approaches that are more efficient and human relevant. EPA's Office of Pesticide Programs (OPP) has developed a strategic vision for implementing the 2007 NRC report on Toxicity Testing in the 21st Century. This strategic vision has multiple components involving a combination of computational and predictive modeling approaches, *in vitro* techniques, and limited, targeted *in vivo* testing. To supplement or replace the existing toxicity tests used to support pesticide registration, EPA's OPP is actively engaged in numerous activities with respect to reducing laboratory animal use and implementing *in vitro* and computational approaches. Development and implementation of non-animal testing has been supported by the EPA, with the EPA Administrator signing a directive in 2019 that prioritizes efforts to reduce animal testing and a goal to reduce requests for and funding of mammalian studies by 30% by 2025. As new approach methodologies (NAMs) can provide human relevant information that may be challenging to test in whole animals, EPA's OPP is also interested in using NAMs to reduce the reliance on default assumptions for risk assessment, including the application of 10X default uncertainty factors each for interspecies and intraspecies extrapolations.

In line with this, several years ago, OPP began collaborative work with EPA's Office of Research and Development (ORD), academia, and industry to use the organophosphate pesticides (OPs), as a case study, for the development of NAMs for purposes of using NAM data to inform extrapolation/uncertainty and safety factors in lieu of reliance on default factors. This research is at a point where external peer review and public comment would be useful to specifically inform human health risk assessment for the OP class of insecticides and also as part of confidence building activities needed to implement NAMs in human health risk assessment more broadly. If appropriate, EPA's OPP may use such NAM information as part of a weight of evidence evaluation for the 10X Food Quality Protection Act (FQPA) Safety Factor. Additionally, *in vitro* acetylcholinesterase (AChE) inhibition data has been generated for OP compounds. The OPP is considering the potential use of these data to develop interspecies and/or intraspecies data-derived extrapolation factors (DDEFs) in accordance with EPA's 2014 *Guidance for Applying Quantitative Data to Develop DDEFs for Interspecies and Intraspecies Extrapolation*<sup>1</sup>.

## 1.1 Regulatory Context

Pesticides are regulated in the United States under both the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) and the Federal Food, Drug and Cosmetics Act (FFDCA), which were both amended in 1996 by the Food Quality Protection Act. Under FIFRA, EPA reviews each registered pesticide at least every 15 years to determine whether it continues to meet the standard for registration. The current 15-year cycle for Registration Review of OPs is due for completion by October 1, 2022. OPs are a class of insecticides with numerous uses including application in agriculture on many crops and by public health organizations for mosquito control. Twenty-two OPs are undergoing Registration Review. Under the FQPA, EPA must consider aggregate risk from exposure to a pesticide from multiple sources (food, water, residential and other non-occupational sources) and consider cumulative exposure to pesticides that have common mechanisms of toxicity. In addition, as part of FQPA, EPA must consider potential for susceptibility of children to pesticides.

---

<sup>1</sup> <https://www.epa.gov/risk/guidance-applying-quantitative-data-develop-data-derived-extrapolation-factors-interspecies-and>

Common to all OPs is the ability to inhibit the AChE enzyme, which prevents the breakdown of acetylcholine leading to neurotoxicity. EPA conducts human health risk assessments for each individual OP in aggregate assessments in addition to the class of OPs in the cumulative risk assessment. OPP has historically used 10% AChE inhibition in red blood cells (RBCs) or brain as the biological response for human health risk assessment. In the early 2000s, scientific evidence from epidemiology, laboratory animals, and mechanistic studies began to emerge that OPs, particularly chlorpyrifos, may adversely affect brain development. Developmental neurotoxicity (DNT) studies are used to evaluate the developing nervous system and are conditionally required under the 40 CFR Part 158 toxicology data requirements for pesticides<sup>2</sup>. EPA issued data call-ins for *in vivo* DNT studies for OPs; however, none of the 18 submitted DNT studies for OPs were found to be more sensitive than AChE inhibition and were not used to derive points of departure (PODs) for human health risk assessment. OPP has spent more than 10 years evaluating the scientific evidence regarding the potential impact of OP exposure on neurodevelopment, including convening multiple Scientific Advisory Panel (SAP) meetings (2008, 2012, 2016). The SAP reports have provided numerous recommendations for additional research and sometimes conflicting advice for how EPA should consider (or not consider) the epidemiology data in conducting EPA's Registration Review for human health risk assessments. The Agency is *not* soliciting comment on the epidemiological and animal behavior studies as part of this SAP. The OPs are being presented as a case study for the development of a battery of NAMs to evaluate DNT for inclusion in an overall weight of evidence evaluation.

## 1.2 Scope of the Draft Issue Paper

This draft document includes two major sections:

**Section 2** begins with a description of the *in vivo* DNT guideline, its attributes, and uncertainties followed by a summary of experiments and statistical analyses of work done by ORD to develop a battery of NAMs for evaluating DNT. Available results for 27 OPs will be presented. This work also involves *in vitro* to *in vivo* extrapolation (IVIVE) to estimate NAM-derived administered equivalent doses (AEDs), which are compared with rat benchmark dose (BMD) and associated lower confidence bound (BMDL) values estimated from AChE inhibition data used in EPA's risk assessments. This Issue Paper does not propose any chemical-specific or class-based FQPA Safety Factor. Moreover, EPA is *not* soliciting comment on the magnitude of the FQPA Safety Factor, as this is a policy decision that encompasses information beyond the scope of the NAM evaluation (e.g., completeness of exposure and toxicology data). EPA is soliciting comment from the SAP on the use of *in vitro* assays developed by ORD for evaluating neurodevelopmental endpoints and the ability of the currently available battery of assays to cover critical processes in neurodevelopment. Additionally, EPA is soliciting comment on the process taken to compare AEDs to BMD/BMDL values based on AChE inhibition in order to compare their relative sensitivity.

**Section 3** provides background on the EPA's 2014 guidance on the development of DDEFs and describes experiments conducted for the purposes of developing interspecies and intraspecies pharmacodynamic DDEFs. The experiments were performed by academia on behalf of pesticide registrants and evaluated *in vitro* AChE inhibition constants in rats and humans for 17 OPs, with inhibition constants generated for 16 of the 17 OPs. This section summarizes these experiments, associated statistical analyses, and provides EPA's interpretation and recommendations for their appropriate use. EPA is soliciting comment from the SAP on the study design and methods utilized to generate the *in vitro* data, statistical analyses employed

---

<sup>2</sup> [https://www.ecfr.gov/cgi-bin/text-idx?tpl=/ecfrbrowse/Title40/40cfr158\\_main\\_02.tpl](https://www.ecfr.gov/cgi-bin/text-idx?tpl=/ecfrbrowse/Title40/40cfr158_main_02.tpl)

to calculate DDEFs, and analyses performed for a small subset of OPs to evaluate contributions of experimental and intrinsic variability.

## 2.0 Fit-for-purpose alternatives to the *in vivo* DNT study

### 2.1 Background & History on the DNT

Exposure to xenobiotics during critical stages of development may result in altered neural development leading to potential lifelong ramifications (Barone et al., 2000; Bondy and Campbell, 2005; Rodier, 1995). Consequently, regulatory authorities throughout the world have promulgated testing guidelines for DNT (OECD, 2007; USEPA, 1998a). DNT refers to any adverse effect of exposure to a toxic substance on the normal development of nervous system structures and/or functions (USEPA, 1998b). The basic purpose of DNT guideline testing is to act as an assessment and screen for the potential of chemicals to cause adverse neurodevelopmental outcomes. This is achieved through a series of evaluations that measure the functional and/or structural integrity of the developing nervous system. Tests for evidence of neurobehavioral deficits in laboratory animals include auditory startle habituation, motor activity, functional observational battery (FOB), learning and memory, and clinical observations. In addition, neuropathologic examinations including simple morphometrics analyses are also part of the DNT study.

To be most useful, studies should have sufficient sensitivity and specificity to reliably detect effects of chemicals at human relevant doses while, to the extent possible, reducing false positives or negatives. As is the case with any testing approach designed to assess complex biological processes, the DNT guideline study has strengths and limitations that must be taken into consideration when evaluating a substance's ability to elicit neurodevelopmental effects. One of the strengths of the DNT guideline study is that it evaluates multiple functional domains. However, the reliable detection, measurement, and interpretation of treatment-related DNT effects depends on appropriate study design and conduct that adequately controls for confounding factors such as variability (*e.g.*, due to dosing regimen, age at treatment and assessment, or inherent measure variability), impact of systemic maternal and/or offspring toxicity, experimental procedures, environmental conditions, etc. Hence, a weight of evidence approach is critical to the assessment of a substance's DNT potential.<sup>3</sup> Moreover, the DNT guideline infers DNT effects on the basis of apical endpoints with little or no information on the underlying biological processes responsible for the observed phenotype. In general, the interpretation of the DNT guideline study is hampered by a number of limitations including high variability, low precision, and being resource intensive (from a financial, labor and animal use perspective). For instance, it is not uncommon for coefficients of variation to be comparable or at times even higher than the mean values for a number of endpoints (*e.g.* motor activity and auditory startle). Another challenge in data interpretation is the issue of isolated findings where a change in one endpoint is not substantiated by other endpoints. For example, a small magnitude change in a morphometric measurement in a single brain region that is not accompanied by behavioral changes presents the challenge of deciding whether the change will eventually lead to an adverse health outcome particularly when considering neuroplasticity. Finally, an additional challenge to the interpretation of DNT data is the issue of correlating behavioral and/or neuropathological effects in the animal model to the myriad of complex neurological deficits seen in the human population ranging from subtle learning disabilities to neural tube defects.

---

<sup>3</sup> NAFTA TWG Developmental Neurotoxicity Guidance Document (2016).  
[https://www.epa.gov/sites/production/files/2017-02/documents/developmental\\_neurotoxicity\\_study\\_internal\\_guidance\\_document\\_final\\_0.pdf](https://www.epa.gov/sites/production/files/2017-02/documents/developmental_neurotoxicity_study_internal_guidance_document_final_0.pdf)

It is within these challenges associated with the DNT guideline that the advantages of NAMs become evident. Integrating *in vitro* and computational information with any available *in vivo* data as part of an overall weight of evidence evaluation can address some of the limitations encountered in the standard testing paradigm (e.g. high variability, low throughput, high cost, or confounding factors) and help elucidate some of the uncertainties these limitations introduce into the assessments, including human relevance, mechanistic underpinning of biological processes leading to apical effects, and impact of confounding factor such as maternal systemic toxicity.<sup>4</sup> Incorporating a battery of NAMs to the evaluation of DNT would also aid in the data interpretation by providing multiple lines of evidence that may help elucidate the biological processes underpinning the apical endpoints affected in the guideline studies.

Since the 1998 adoption of the US EPA's DNT test guideline, OPP has received approximately 100 DNT guideline studies. Notably, only 24 of these studies have been used to establish PODs in risk assessments. Of these 24 studies, nine were selected on the basis of changes in brain morphology (neuropathology and/or morphometrics) and five on the basis of behavioral changes; the remaining studies were selected on the basis of non-neurotoxicity endpoints (e.g., pup weight decrements). For known neurotoxicant chemical classes such as the OPs, *N*-methyl carbamates (NMCs), and pyrethroids, EPA has 32 DNT guideline studies, but none have provided the most sensitive endpoint for human health risk assessment. As a result, EPA has shifted its testing focus from the DNT guideline study to more targeted testing based on commonly accepted modes of action (MOA). In the case of the OPs and NMCs, AChE inhibition was selected as the critical endpoint and the focus of special testing on lifestage susceptibility, including the use of comparative cholinesterase assays (CCA). For pyrethroids, *in vitro* studies and physiologically based pharmacokinetic (PBPK) modeling based on the known MOA, which involves interaction with voltage-gated sodium channels leading to neurotoxicity, have been used to evaluate lifestage susceptibility for this class of chemicals<sup>5</sup>. For chemicals toxic to the thyroid system, the Agency's testing strategy has shifted from the DNT to the Comparative Thyroid Assay (CTA) again using a MOA-driven approach<sup>6</sup>. Hence, while the Agency has already begun implementing alternative approaches to test for lifestage susceptibility, the inclusion of NAMs that take advantage of newer technologies is the logical next step in the Agency's efforts to implement more human relevant and efficient approaches.

## 2.2 Background & History of the OPs

OP pesticides are a group of closely related insecticides that affect the functioning of the nervous system. Following the passage of the FQPA, OPs were among the Agency's first priority groups of pesticides to be reviewed. In 1999, EPA determined that the OPs form a common mechanism group based on their shared ability to bind to and phosphorylate AChE in both the central (brain) and peripheral nervous systems (USEPA, 1999) (USEPA, 1999). Some OPs must be metabolized (activated) to an oxon metabolite, which is the active AChE inhibiting moiety. Inhibition of AChE leads to accumulation of acetylcholine and ultimately to neurotoxicity (see Figure 1). AChE inhibition is consistently observed in the OP toxicology databases in multiple species, durations, lifestages, and routes.

---

<sup>4</sup> Report of the Workshop on a Framework for the Development and Use of Integrated Approaches to Testing and Assessment. 2015. OECD Series on Testing and Assessment No. 215

<sup>5</sup> <https://www.epa.gov/ingredients-used-pesticide-products/2019-evaluation-fqpa-safety-factor-pyrethroids>

<sup>6</sup> [https://www.epa.gov/sites/production/files/2015-06/documents/thyroid\\_guidance\\_assay.pdf](https://www.epa.gov/sites/production/files/2015-06/documents/thyroid_guidance_assay.pdf)

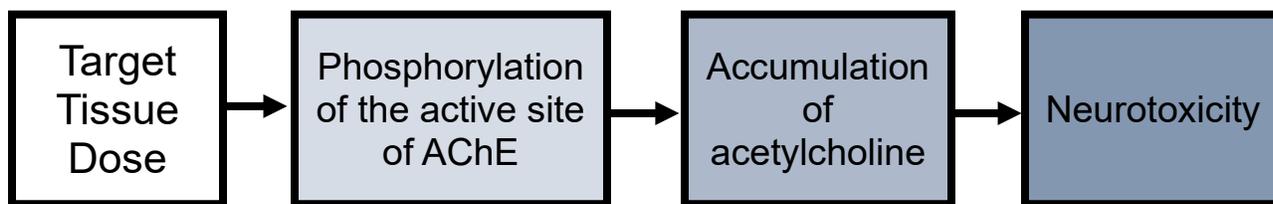


Figure 1. AOP for OPs.

OPs also exhibit a phenomenon known as steady-state AChE inhibition. After repeated dosing at the same dose level, the degree of inhibition comes into equilibrium with the production of new, uninhibited enzyme. At this point, the amount of AChE inhibition at a given dose remains consistent across duration. In general, AChE inhibition from OP exposure reaches steady-state within 2-3 weeks, but this can vary among OPs. As such, risk assessments currently evaluate acute and steady-state exposures from registered uses of OPs.

In more recent years, a series of FIFRA SAP meetings have been held for chlorpyrifos regarding the use of epidemiological data to inform the chlorpyrifos human health risk assessment, as well as its relevance to other OPs. The Agency has used and continues to use inhibition of AChE as the POD for OP human health risk assessment. This science policy is based on decades of work which shows that AChE inhibition is the initial event in the pathway to acute cholinergic neurotoxicity. The use of AChE inhibition data for deriving PODs was supported by the FIFRA SAP in 2008a and 2012<sup>7</sup> for the OP pesticide, chlorpyrifos, as the most robust source of dose-response data for extrapolating risk. A PBPK-pharmacodynamic (PD or PBPK-PD) model has also been developed for chlorpyrifos and its oxon and reviewed by the FIFRA SAP<sup>8</sup>. This model has been used as the underlying structure for developing PBPK-PD models for other OP chemicals, such as dimethoate and malathion. Newer lines of research on OPs in the areas of potential adverse outcome pathways (AOPs), *in vivo* animal studies, and notably epidemiological studies in mothers and children, have raised some uncertainty about the Agency's risk assessment approach with regard to the potential for neurodevelopmental effects in fetuses and children. At this time, a MOAs/AOPs have not been established for neurodevelopmental outcomes. The growing body of literature does demonstrate, however, that OPs are biologically active on a number of processes that have the potential to affect the developing brain. Moreover, there is a large body of *in vivo* laboratory studies, including DNT studies, which show long-term behavioral effects from early life exposure, albeit at doses which cause AChE inhibition with limited exceptions (Carr et al., 2017; Carr et al., 2014). Interpretation of these data has been challenging particularly in cases where conflicting results were observed across studies. The SAP reports have provided numerous recommendations for additional research and sometimes conflicting advice for how EPA should consider (or not consider) the epidemiology data in conducting human health risk assessments for OPs. Given the large amount of data available for OPs and widespread acceptance of the neurotoxic MOA for OPs based on AChE inhibition, the OPs were identified as a group of chemicals that could be utilized as a case study for the development of a battery of NAMs to evaluate DNT. Data from these *in vitro* assays can be considered as part of an overall weight of evidence to determine the DNT potential for individual OPs. If appropriate, this information could also be used to inform the FQPA Safety Factor in the future. Additionally, the well-established MOA provides an opportunity to utilize *in vitro* data for deriving DDEFs for OPs reducing the Agency's reliance on default uncertainty factors for risk assessment. Overall, these approaches are consistent with the Agency's efforts to reduce laboratory animal use and provide opportunities to utilize human relevant data to inform human health risk assessments.

<sup>7</sup>2008 SAP: <https://www.epa.gov/sap/fifra-scientific-advisory-panel-historical-meetings>

2012 SAP: <https://www.regulations.gov/docket?D=EPA-HQ-OPP-2012-0040>

<sup>8</sup> <https://www.regulations.gov/docket?D=EPA-HQ-OPP-2012-0040>

### 2.3 New approach methodologies (NAMs) for DNT

Researchers within the Office of Research and Development (ORD) have been developing new NAMs for DNT for over a decade as part of an international effort to increase the number of chemicals for which DNT-relevant information can be obtained in a screening mode. This international effort recognized limitations of the *in vivo* guideline studies, and through a series of meetings with scientists, regulators and stakeholders (Bal-Price et al., 2018a; Crofton et al., 2011; Fritsche et al., 2018; Lein et al., 2007), ORD developed a strategy to devise NAMs that address some of these issues. This strategy focuses on development of *in vitro* assays that assess processes critical to development of the nervous system, and has resulted in a battery of *in vitro* tests, many of which are now believed to be capable of providing data that would be useful for public health applications (Bal-Price et al., 2018a; Bal-Price et al., 2018b). This battery of tests includes assays that evaluate chemical effects on critical neurodevelopmental processes including proliferation of neuroprogenitor cells, differentiation of neuroprogenitors into glial and neuronal subtypes, apoptosis, migration of neurons and oligodendrocytes, neurite outgrowth, synaptogenesis and neural network formation. Currently, assays developed by ORD using two primary technology platforms from these ongoing efforts are ready for review regarding their relevance to DNT screening and evaluation: microelectrode arrays (MEAs) with neuronal cell types to understand neuronal network formation, and high-content imaging assays of neural cells to understand key processes relevant to neurodevelopment.

These DNT-NAMs are multidimensional assay technologies, with multiple assay endpoints, and are amenable to analysis with the ToxCast Data Pipeline (Filer et al., 2017) for standardizing the concentration response analysis of heterogeneous *in vitro* bioactivity screening data. Briefly, as these assay technologies are described in detail below, the total data set for the MEA evaluation of neuronal network formation assay includes 36 assay endpoints that indicate general neuronal activity, bursting, and connectivity. The data set for the high content imaging of neural cells encompasses 5 related assays: cell proliferation with a human neural progenitor cell line; apoptosis and viability assessment using a human neural progenitor cell line; neurite outgrowth initiation with a human neuronal lineage cell line; primary rat cortical cell neurite outgrowth initiation; primary rat cortical cell neurite maturation and synaptogenesis. From these 5 related assays, 21 assay endpoints can be analyzed. Thus, in total, the ORD DNT-NAM battery currently available in the ToxCast database (see Supplemental File 1 that contains all source data and R software) is comprised of 57 assay endpoints in human and rat neural cells that evaluate proliferation, apoptosis, neurite outgrowth and maturation, synaptogenesis, and neural network formation. Each of these assays incorporate measures of viability.

These DNT-NAMs have been developed for use in DNT screening and prioritization, but data should be interpreted with the following considerations in mind: 1) these assays may not include all of the processes and targets necessary for full evaluation of DNT (Sachana et al., 2019) (see Section 2.3.2); 2) the battery of assays considered here do not address disruption of thyroid hormone levels, which can also have adverse neurodevelopmental outcomes; and, 3) a consideration for assay development was the availability of scalable and economic human cell models, with rat models being utilized when human models were not available or feasible. ORD has a separate, specific effort aimed at developing *in vitro* methods to assess chemical effects on the thyroid axis that may relate to DNT. The DNT-NAM assays currently available represent a significant milestone in the rapid generation of data to inform understanding of putative DNT-related bioactivity. The utility of these DNT-NAM assays for identifying the potential for chemicals to affect DNT, as part of a broader high-throughput screening (HTS) screening effort, is

evident in that assays that recapitulate neurodevelopmental function and processes fill a previous gap in the biological space covered by programs including ToxCast/Tox21 (Paul Friedman et al., 2019; Shafer et al., 2019; Harrill et al., 2018). The DNT-NAMs may also have further fit-for-purpose utility in identifying putative DNT activity within a chemical safety assessment context.

In this Issue Paper, data from these DNT-NAMs are leveraged in a case study for application to chemical safety assessment. The bioactivity data from these DNT-NAMs are used with an IVIVE approach to indicate a threshold bioactivity concentration that may have relevance to DNT evaluation, specifically using a single class of substances: OPs. For this case study with OPs, there are specific bounds on the context of use of the DNT-NAM data. For instance, the registration of OPs has required *in vivo* testing in animal models, and as such these OP chemicals have been evaluated in these models for thyroid-related toxicity in adult animals, as well as acute and developmental neurotoxicity (see Sections 2.1 and 2.2). Further, within the regulatory context for OPs, AChE inhibition data are currently used as the basis for endpoints used for human health risk assessment, and additional data has been and/or can be generated. The specific problem formulation for DNT-NAM data is to evaluate if these HTS data can, at this time, provide additional information regarding the potential for DNT from OP exposure, and if the doses predicted for this putative DNT-related bioactivity differ from the doses currently associated with AChE inhibition *in vivo*.

The subsequent sections (Sections 2.3.1-2.3.6) will discuss the following:

- Section 2.3.1 provides a brief overview of the DNT-NAMs and how they function.
- Section 2.3.2 describes how these DNT-NAMs have been developed within an international scientific community, with plans for further research to understand the utility of existing and developing DNT-NAMs for evaluating DNT potential.
- Section 2.3.3 provides the specific DNT-NAM assay methodology employed to screen the OPs in these assays
- Section 2.3.4. describes assay reproducibility and performance for the DNT-NAMs as a means of indicating their reliability and robustness.
- Section 2.3.5 presents the DNT-NAM results for 27 OP chemicals.
- Section 2.3.6 describes an example implementation of IVIVE of AEDs that relate to bioactivity in the DNT-NAMs and comparisons of these AEDs to BMD values from studies of inhibition of AChE in rats are provided.

## **2.3.1 Overview of the DNT-NAMs**

### ***2.3.1.1 Microelectrode arrays***

Neuronal cell communication proceeds through chemical and electrical signaling. MEAs can be used to understand the effects of *in vitro* chemical exposure on the normal electrical functions of neuronal cells that have traditionally been measured via neurophysiological assessment (Atchison, 1988; Gramowski-Voss et al., 2015; Johnstone et al., 2010; Robinette et al., 2011; Shafer, 2019). Many different adverse outcome pathways may lead to a disruption of the electrical excitability pattern of neuronal cells, and as such, a HTS approach to identifying putative disruption of neuronal action potentials regardless of the specific molecular initiating event (MIE) has been developed (Johnstone et al., 2010). Primary cortical cell cultures grown on multi-well MEAs can demonstrate the electrical functions of neuronal cell networks, and this can be measured via recording the action potentials of these cells. Spontaneous

electrical spikes and groups of spikes (i.e., bursts) are associated with the action potentials of cortical cells grown in a network on MEAs. The rates of electrical activity (e.g., mean firing rate, burst rate), specific pattern of electrical activity (e.g., bursting activity), and the synchronization of these electrical activities across a network, known as network connectivity, can all be measured in HTS to identify responses resultant to *in vitro* chemical exposure. Disruptions in electrical activity, either via increased excitation or suppression of electrical activity, as indicated by effects on multiple parameters measured in MEAs, may indicate neurotoxicity potential from perturbation of one or more MIEs. For instance, though anti-convulsant drugs are associated with slightly different MIEs, they appear to generally reduce bursting activity and network connectivity measures (Johnstone et al., 2010; Schroeder et al., 2008). By contrast, convulsant drugs often increase general firing activity, burst rate and/or duration and increase network connectivity measures (Bradley et al., 2018; Bradley and Strock, 2019; Kreir et al., 2018).

In addition to the use of MEAs to evaluate acute effects on the electrical function of neural cells, MEAs have demonstrated utility for evaluation of cortical neural network formation, a process critical for neurodevelopment. The effects of *in vitro* chemical exposure in a MEA network formation assay (MEA NFA) can be evaluated via measurement of parameters that indicate general activity (mean firing rate, burst rate, etc.); parameters that indicate bursting activity (burst duration, interburst interval, number of spikes in bursts, etc.) and parameters that describe the connectivity of the network (network spikes, correlation coefficient, normalized mutual information (Ball et al., 2017)). While some descriptors are directly measured (mean firing rate, burst rate, number of network spikes) other descriptors (inter spike interval, interburst interval, etc.) are derived from these directly measured descriptors (Brown et al., 2016; Cotterill et al., 2016) (Table 1). Table 1 provides a reference on the measures available from the MEA NFA, including the type of activity evaluated, the name of the assay endpoint, the assay component and endpoint identifiers from the ToxCast Pipeline (Filer et al., 2017)(tcpl, used for analysis), and a description of what is being measured. The MEA network formation assay (MEA-NFA) within EPA ORD has been used to screen >205 unique substances to date (Brown et al., 2016; Frank et al., 2017; Shafer et al., 2019), along with parallel cytotoxicity assays, to increase the number of chemicals for which some information relevant to putative DNT effects is available.

In the MEA-NFA, primary rat cortical neurons grown on arrays of microelectrodes demonstrate increasing neural activity over 12 days in culture; this activity begins with random, uncoordinated sporadic firing of individual neurons and progresses to synchronous network bursts by day 12 in culture (Cotterill et al., 2016). This progressive increase in complexity of neural activity is a result of neurite outgrowth, synaptogenesis, and development of functional network activity, suggesting that the MEA-NFA assay (Cotterill et al., 2016; Frank et al., 2017; Robinette et al., 2011) recapitulates critical aspects of neurodevelopment. This is further supported by the fact that the spiking, bursting and coordinated network activity measured by the MEA-NFA are conserved across mammalian species- from rodents to primates (Mochizuki et al., 2016) and both *in vitro* and *in vivo* (Khazipov and Luhmann, 2006). Small networks of interconnected neurons are critical to nervous system function. These networks often exhibit synchronous and oscillatory behavior (Salinas and Sejnowski, 2001; Uhlhaas et al., 2009; Uhlhaas and Singer, 2006) and their function is disrupted in pathological disease states, including schizophrenia, epilepsy, autism and neurodegenerative diseases (Uhlhaas and Singer, 2006). The formation of functional neural networks requires a number of cell types and processes, including neurite outgrowth, synaptogenesis, maturation of glia, excitatory and inhibitory neuron signaling, neurotransmitter recycling, and maintenance of cellular polarity and electrochemical gradients (Frank et al., 2017). Confirmatory imaging of MEA-NFA assay cultures at day 12 in culture has previously demonstrated the presence of excitatory and inhibitory neurons, astrocytes, and some microglia (Frank et al., 2017), suggesting that

beyond measurement of assay parameters of neural network function, the components needed for appropriate function are indeed present.

Similar deployments of MEA technology to measure cortical and hippocampal neural network formation and function (Bisio et al., 2014; Charlesworth et al., 2015; Chiappalone et al., 2006; de Groot et al., 2014; Frega et al., 2017; Kondo et al., 2017; Lenk et al., 2016) further support the established use of this assay technology for evaluating effects on the ontogeny of neural function, as recently reviewed (Shafer, 2019). Indeed, MEA assays using rat primary cortical cells in already formed neural networks and an acute chemical exposure has been proposed as a reliable screening assay for neurotoxicity and seizure-inducing compounds during drug discovery (Bradley et al., 2018; Bradley and Strock, 2019). Neural networks grown on MEAs have been used to determine the acute activity of a large number of different types and classes of compounds, including agrochemicals (Alloisio et al., 2015) (Baskar and Murthy, 2018; Meyer et al., 2008; Mohana Krishnan and Prakhya, 2016; Shafer et al., 2008), mixtures of pyrethroids (Johnstone et al., 2010; Scelfo et al., 2012), nanoparticles (Gramowski-Voss et al., 2015; Strickland et al., 2018), tricresyl phosphate (Duarte et al., 2017), illicit drugs (Hondebrink et al., 2016), glufosinate (Lantz et al., 2014), antiepileptic drugs (Colombi et al., 2013), excitotoxicants (Frega et al., 2012), components of harmful algae (Alloisio et al., 2016), neuroactive toxins (Kasteel and Westerink, 2017; Pancrazio et al., 2014) and metals (Dingemans et al., 2016; Huang et al., 2016). The widespread deployment of this technology to evaluate a broad variety of compounds and disruptions of nervous system function support the concept that neural networks grown on MEAs provide useful, biologically-relevant information about interactions of chemicals with the nervous system.

**Table 1. Network activity parameters obtained from MEA NFA recordings**

The recordings from MEA NFA can be categorized into 3 activity types: general, bursting, and network connectivity. Further delineation of the measurement type (empirical = Emp or derived = Der), the parameter and its tcpl assay component name (in *italics*), the level of the measurement (electrode or well), the assay component identification (ACID) and a description are provided. Two assay endpoint identifiers are provided (AEIDs), where the odd numbered AEID is for endpoint analyzed in the “up” direction and the even numbered AEID is for the endpoint analyzed in the “down” direction. For features computed by electrode, such as burst rate, one well level value per well was obtained by taking the average across electrodes within a well.

Activity Type	Type	Parameter	Level	ACID	AEIDs	Description
General Activity	Emp	Mean Firing Rate, <i>NHEERL_MEA_dev_firing_rate_mean</i>	Electrode	2471	2494,2495	The mean firing rate on each electrode was calculated, with the well level value equal to the mean across all active electrodes
	Emp	Burst Rate, <i>NHEERL_MEA_dev_burst_rate</i>	Electrode	2472	2496,2497	The number of bursts per minute. Max-interval method used with parameters: ISI to start =0.1s, ISI to end =0.25s, min IBI =0.8, min duration =0.05s, min no. spikes = 5
	Emp	Number of Active Electrodes, <i>NHEERL_MEA_dev_active_electrodes_number</i>	Electrode	2473	2498,2499	Number of electrodes firing at or above 5 spikes per minute.
	Emp	Number of Actively Bursting Electrodes, <i>NHEERL_MEA_dev_bursting_electrodes_number</i>	Electrode	2474	2500,2501	Number of electrodes with burst rates of above 0.5 bursts per minute
Bursting Activity	Der	Interspike Interval (ISI) within a burst, <i>NHEERL_MEA_dev_per_burst_interspike_interval</i>	Electrode	2475	2502,2503	Time interval between spikes within a burst (ms)
	Emp	Percentage of Spikes in Burst, <i>NHEERL_MEA_dev_per_burst_spike_percent</i>	Electrode	2476	2504,2505	The number of spikes within a burst over total spike count x 100
	Der	Mean Burst Duration, <i>NHEERL_MEA_dev_burst_duration_mean</i>	Electrode	2477	2506,2507	Mean duration of a burst (ms)
	Der	Mean interburst interval, <i>NHEERL_MEA_dev_interburst_interval_mean</i>	Electrode	2478	2508,2509	Mean time interval between bursts (sec)

Network Connectivity	Emp	Number of Network Spikes, <i>NHEERL_MEA_dev_network_spike_number</i>	Well	2479	2510,2511	Number of spikes in network spikes
	Der	Network Spike Peak, <i>NHEERL_MEA_dev_network_spike_peak</i>	Well	2480	2512,2513	The number of electrodes active at peak of network spike
	Der	Network Spike Duration, <i>NHEERL_MEA_dev_spike_duration_mean</i>	Well	2481	2514,2515	The average duration (ms) of a network spike
	Der	SD of Network Spike Duration, <i>NHEERL_MEA_dev_network_spike_duration_std</i>	Well	2482	2516,2517	Standard deviation of network spike duration
	Der	ISI in Network Spike, <i>NHEERL_MEA_dev_per_network_spike_interspike_interval_mean</i>	Well	2483	2518,2519	Mean inter-spike interval for spikes in network spikes
	Der	Mean number of Spikes in Network Spikes, <i>NHEERL_MEA_dev_per_network_spike_spike_number_mean</i>	Well	2484	2520,2521	Number of spikes in network spike
	Emp	% Spikes in Network Spike, <i>NHEERL_MEA_dev_per_network_spike_spike_percent</i>	Well	2485	2522,2523	Ratio of spikes in network spikes over total spikes x 100
	Emp	Mean Correlation, <i>NHEERL_MEA_dev_correlation_coefficient_mean</i>	Well	2486	2524,2525	The average of all pairwise correlation between all electrodes
	Emp	Normalized Mutual Information, <i>NHEERL_MEA_dev_mutual_information_norm</i>	Well	2487	2526,2527	Normalized mutual Information between all electrodes in the well.

### 2.3.1.2 High content imaging (HCI) for cellular events

The HCI cellular event assays presented herein have been used previously to model critical processes of nervous system development including neuroprogenitor proliferation, apoptosis, neurite outgrowth and synaptogenesis. These processes, when affected by chemical treatment (Baumann et al., 2016; Culbreth et al., 2012; Harrill et al., 2015; Harrill et al., 2018; Krug et al., 2013; Mundy et al., 2015) may contribute to DNT-relevant outcomes such as cognitive impairments and/or learning disabilities, autism spectrum disorder, attention deficit hyperactivity disorder, though the specific neurodevelopmental mechanisms for these DNT outcomes is the subject of ongoing research efforts. Other deployments of HCI include screening small molecule libraries for effects on neurogenesis (Wu and Li, 2018), neuroplasticity and CNS disorders relevant for pharmaceutical development (Cheng et al., 2017), as well as other general drug development applications (including drug target optimization and preclinical safety screening) involving the CNS and its development (Gorshkov et al., 2018; Kepiro et al., 2018; Sharma et al., 2012). These uses support a broader acceptance of HCI technology for identifying and characterizing chemical perturbation of neural cell biology. Further, orthogonal assays have been developed in central and peripheral nerve models (Delp et al., 2018; Nyffeler et al., 2017), as well as 3-dimensional neural models known as neurospheres (Baumann et al., 2016) for the assessment of chemical effects on neurodevelopment.

The HCI cellular event assays described in this Issue Paper, as a multi-dimensional set, have been used to distinguish perturbation of the processes of neuroprogenitor proliferation, neurite outgrowth, and synaptogenesis from cytotoxicity (Table 2). Table 2 provides an overview of the relevant process addressed, HCI assay names, assay component and endpoint identifiers from tcpl (used for analysis), and what is measured in each assay. Apoptosis and cytotoxicity are also relevant to DNT if neural cell death occurs *in vivo* during critical periods of nervous system development and at concentrations that would indicate this apoptotic or cytotoxic event might occur at lower concentrations than some overt toxicity. As several critical processes of *in vivo* neurodevelopment are encompassed by these phenotypic screens, they serve as a broad approach to detect chemical perturbations of a variety of neurodevelopmental processes *in vitro*. Due to the temporal and biological complexity of nervous system development involving the interplay of many different cell types, it is generally accepted that no single *in vitro* screening assay can recapitulate all the critical processes of neurodevelopment or affirmatively identify all chemicals that may produce DNT. Therefore, the need to use an assay battery approach for *in vitro* DNT screening is evident (Bal-Price et al., 2018). Some chemicals may affect only a single neurodevelopmental process as an indicator of DNT activity. For instance, a chemical may affect *in vitro* synaptogenesis and fail to affect cellular events that occur earlier in neural network development such as neural progenitor cell proliferation or neurite outgrowth (Harrill et al., 2018). If either one of the latter two assay types were the only assay used for DNT hazard screening, then the DNT potential of the chemical in question would not have been detected. Conversely, some chemicals may disrupt early processes of nervous system development such as neural progenitor proliferation. *In vivo* this type of effect would also indirectly impact processes such as neurite outgrowth and synapse formation that are temporally downstream of the initial process. *In vitro*, the biological activity of this type of chemical would not be detected if using a single assay that models a late event in neurodevelopment.

**Table 2. HCI Cellular Event Assay Suite.**

The activity type, cell type/species, technology, assay name from tcpl, endpoint names from tcpl, and a description of the measures are provided. The assay component ID (ACID) and assay endpoint ID (AEID) are also provided; these assays were only analyzed in the “down” direction.

Neurodevelopmental Process	Assay Name (Cell Type)	ACID	AEID	Endpoint Name	Description
Neural Progenitor Cell Proliferation	MUNDY_HCI_hNP1_Pro  (human hNP1 neuroprogenitors)	2711	2795	MeanAvgInten_loss	Intensity of BrdU labeling in the nucleus of each cell, averaged across all cells in a well. A decrease as compared to control is indicative of decreased cell proliferation.
		2709	2796	ResponderAvgInten_loss	Percentage of cells with intensity of BrdU labeling $\geq 3X$ background. A decrease as compared to control is indicative of decreased cell proliferation
		2710	2797	ObjectCount_loss	The number of nuclei per well. A decrease as compared to control is indicative of cytotoxicity.
Neural Progenitor Apoptosis / Viability	MUNDY_HCI_hNP1  (human hNP1 neuroprogenitors)	2691	2793	Casp3_7_gain <sup>a</sup>	Intensity of luminescent signal produced by caspase 3/7 cleavage of a detection reagent. The signal produced is proportional to the number of apoptotic cells. An increase as compared to control is indicative of increased apoptosis.
		2700	2794	CellTiter_loss <sup>a</sup>	Intensity of luminescent signal produced by detection of cellular ATP. The signal produced is proportional to the number of viable cells. A decrease as compared to control is indicative of cytotoxicity.
Neurite outgrowth (NOG) initiation <sup>b</sup>	MUNDY_HCI_hN2_NOG  (human hN2 neural cells)	2695	2789	BPCount_loss	Morphology of $\alpha_{III}$ -tubulin labeled neurons as measured using automated microscopy. Measurements of neurite length (Neurite_Length), the number of neurites (NeuriteCount) and the number of neurite branch points (BPCount) per cell are calculated for each assay well. Decreases in any of these measures are associated with inhibition of neurite outgrowth.
		2694	2790	NeuriteCount_loss	
		2693	2791	Neurite_Length_loss	
		2692	2792	NeuronCount_loss	
	MUNDY_HCI_Cortical_NOG  (1° rat cortical cultures)	2699	2777	BPCount_loss	The number of neurons per well (NeuronCount) is also measured. Decreases in the number of neurons per well as compared to control is indicative of cytotoxicity.
		2698	2778	NeuriteCount_loss	
		2697	2779	Neurite_Length_loss	

Neurodevelopmental Process	Assay Name (Cell Type)	ACID	AEID	Endpoint Name	Description
		2696	2780	NeuronCount_loss	
Neurite Maturation and Synaptogenesis <sup>b</sup>	MUNDY_HCI_Cortical_Synap& Neur_Mature <i>(1° rat cortical cultures)</i>	2707	2781	BPCount_loss	Morphology of MAP2 <sup>c</sup> and synapsin labeled neurons as measured using automated microscopy. Measurements of neurite length (NeuriteLength), the number of neurites (NeuriteCount) and the number of neurite branch points (BPCount) per cell are calculated for each assay well. Decreases in any of these measures are associated with inhibition of neurite maturation.
		2702	2782	NeuriteCount_loss	
		2706	2783	NeuriteLength_loss	
		2705	2784	CellBodySpotCount_loss	In addition, the number of pre-synaptic puncta in the cell body compartment (CellBodySpotCount) and the neurite compartment (NeuriteSpotCountPerNeuron) are counted in each assay well. The number of cell body and neurite-associated puncta are combined to calculate the total number of synapses (SynapseCount). The number of neurite-associated puncta are also quantified per unit length of neurite measured (NeuriteSpotCountPerNeuriteLength). Decreases in any of these features are associated with inhibition of synaptogenesis.
		2704	2785	NeuriteSpotCountPerNeuron_loss	
		2703	2786	NeuriteSpotCountPerNeuriteLength_loss	
		2701	2787	SynapseCount_loss	
		2708	2788	NeuronCount_loss	The number of neurons per well (NeuronCount) is also measured. Decreases in the number of neurons per well as compared to control is indicative of cytotoxicity.

<sup>a</sup> These endpoints are measured using a luminescent plate reader. All other endpoints for all other assays are measured using high-content imaging (HCI). <sup>b</sup> The rat cortical neurite outgrowth and rat cortical neurite maturation and synaptogenesis assays, respectively, are performed in the same *in vitro* cell model. However, the timing and duration of chemical exposures differs across the respective assays in order to different phases of *in vitro* neuronal development. <sup>c</sup> MAP2 is a cytoskeletal protein that localizes specifically to dendrites: i.e., specialized neurites that receive incoming signals from other neurons in an integrated neuronal network. In primary rat cortical cultures, dendrites develop slower than axons and measurement of dendritic morphology is an indicator of neurite maturation.

### 2.3.2 DNT-NAM development as an integrated process

The objective of a DNT-NAM battery is to identify putative DNT-related bioactivity from a set of cellular processes, and optimally a DNT-NAM battery would include assays that recapitulate as many of these processes involved in nervous system development and function as practicable, as extensively reviewed previously (Bal-Price et al., 2018b; Fritsche et al., 2018; Sachana et al., 2019). In Figure 2, adapted from Aschner et al. 2016, the main processes relevant for DNT are illustrated as a network. In the two DNT-NAM assay technologies described herein, some but not all, of these processes are recapitulated to observe chemical-mediated effects.

International efforts to build such a DNT-NAM battery are ongoing with participation of US EPA ORD contributors. As a result of the October 2016 meeting in Brussels, Belgium, EFSA funded a project to evaluate a battery of *in vitro* DNT assays. ORD investigators have developed a collaboration with the investigators funded by EFSA. Together, assays developed by ORD and the EFSA-funded researchers comprise a series of assays that form a battery which covers critical processes in neurodevelopment (Figure 2). The combined effort will evaluate these assays with a test set of ~120 compounds, with results anticipated in 2020-2021. The list of tested compounds was developed in collaboration with the US EPA-ORD, and many, but not all, of the compounds were supplied by EPA's ToxCast program. To maximize resources, assay specific lists of compounds to test were developed, as some compounds had already been tested in some assays. In developing the overall (120 compound) list, compounds with a variety of different characteristics were selected, including those with: 1) evidence of possible *in vivo* DNT hazard based on papers from Mundy et al. (2015) and Aschner et al. (2017); 2) Guideline DNT Studies (OECD TG 426 and/or EPA 870.6300); 3) putative negative DNT compounds; 4) compounds of interest for developing case-studies for Integrated Approaches to Testing and Assessment (IATAs) for DNT (see OECD DNT Expert Group below); and, 5) other compounds for which some *in vitro* data existed in DNT-NAMs. When available, these data will be analyzed using the ToxCast Data Pipeline (tcpl). Following the availability of data from all of the assays, the value added by each assay and assay endpoint for identification of positive and negative reference chemicals will be evaluated, likely using several different approaches in the 2020-2023 timeframe. The proposed assays for this DNT-NAM battery, developed by international consortium, are presented in Table 3 with references and in Figure 2 as a network diagram.

Additionally, in March 2019, the Organisation for Economic Cooperation and Development (OECD) charged an expert group with development of a guidance document for the use of DNT-NAMs for Integrated Approaches to Testing and Assessment (IATAs) for DNT. Development of this guidance document will be informed by a series of IATA case studies intended to illustrate the capabilities of DNT-NAMs.

Another model of DNT has been proposed as complementary to DNT-NAMs: assessment of behavioral effects of chemical exposure during nervous system development in zebrafish embryos/larvae (de Esch et al., 2012; Nishimura et al., 2015; Peterson et al., 2008), based on the conservation of neurodevelopmental and functional processes across species (Guo, 2009; Kokel et al., 2010). The zebrafish model also complements the *in vitro* testing systems by having the advantage of assaying for behavioral changes in an intact, integrated model that is capable of hepatic metabolism (de Souza Anselmo et al., 2018; Goldstone et al., 2010), possesses a thyroid axis that influences brain development, and has a complete genome that is easily modulated. As part of the OECD DNT expert group efforts, a consortium of eight laboratories (including ORD) is working on a case study to develop and define a unified experimental protocol for

behavioral analysis using a zebrafish model of neurodevelopment. This unified protocol will include a developmental exposure followed by locomotor activity assessment in a light and a dark photoperiod. With experimental results anticipated in Fall 2020, a unified data analysis protocol (Filer et al., 2017; Reif et al., 2016; Zhang et al., 2017) will have to be developed for efficient, reproducible interpretation of these studies, followed by evaluation of their utility for DNT evaluation. Again, it is likely this work will proceed initially using a variety of approaches over the course of 2020-2023. For this Issue Paper, a zebrafish behavioral model is not yet mature enough for incorporation into the DNT-NAMs considered at this time.

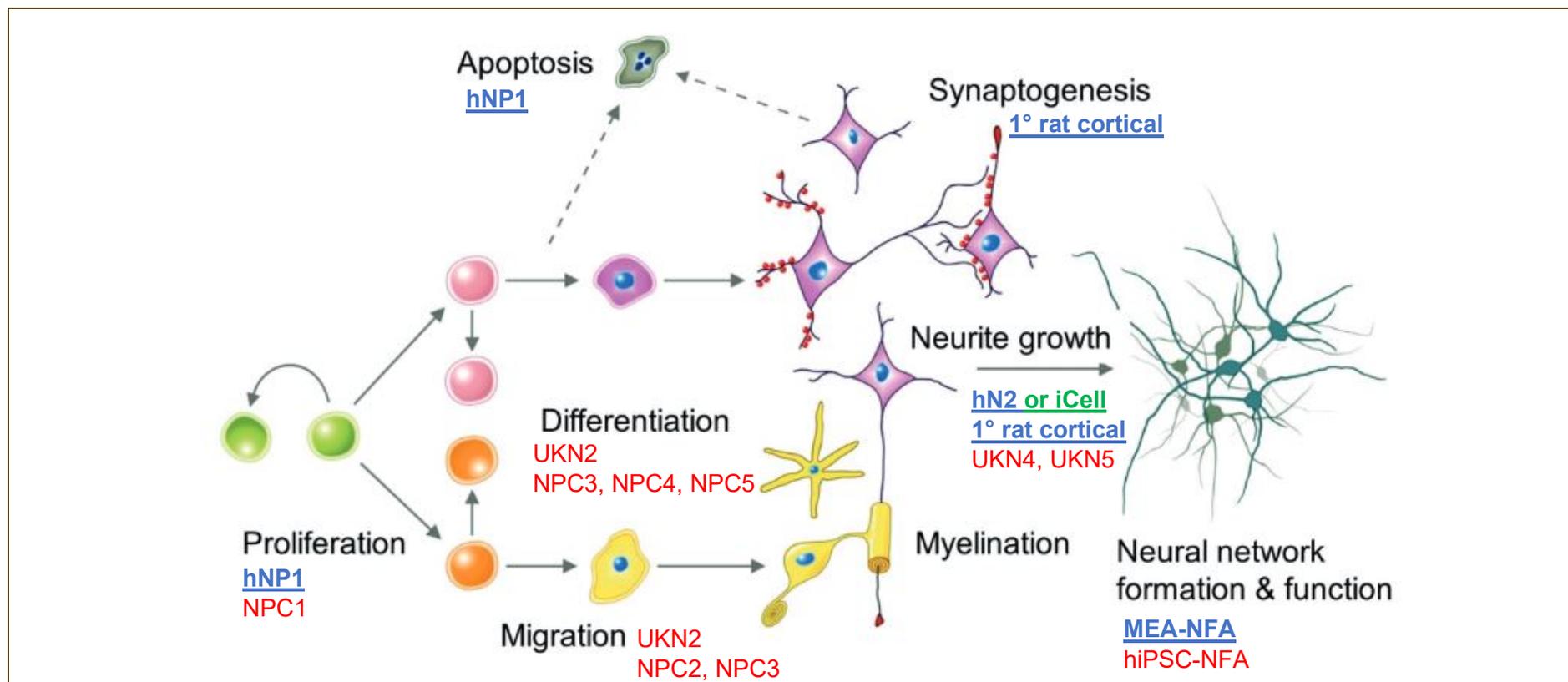
**Table 3. DNT-NAMs for the network of activities relevant for neurodevelopment.**

This table was adapted and expanded based on Sachana et al. (2019). Refer to Figure 2 for additional illustration of the DNT-NAM battery being developed by international collaborators.

Biological Process	Assays	Data available in this Issue Paper	References
Proliferation	hNP1	Yes	(Harrill et al., 2018)
	NPC1	No	(Barenys et al., 2017; Baumann et al., 2016)
Apoptosis	hNP1	Yes	(Druwe et al., 2015; Harrill et al., 2018)
Migration	NPC2	No	(Barenys et al., 2017; Baumann et al., 2016)
Neuron differentiation	UKN2	No	(Nyffeler et al., 2017)
	NPC3	No	(Barenys et al., 2017; Baumann et al., 2016)
Oligodendrocyte differentiation and maturation	NPC5/6	No	(Barenys et al., 2017; Baumann et al., 2016)
Synaptogenesis	1° rat cortical neuron synaptogenesis	Yes	(Harrill et al., 2018; Harrill et al., 2011b)
Neurite outgrowth	1° rat cortical neuron or hN2	Yes	(Harrill et al., 2018; Harrill et al., 2010; Harrill et al., 2011a)
Neuronal network formation	1° rat cortical neuron	Yes	(Brown et al., 2016; Frank et al., 2017; Shafer et al., 2019)
	human IPSC-derived neural networks	No	Unpublished Method

**Figure 2. Integrated network of DNT processes and coverage by relevant DNT-NAMs.**

DNT-relevant processes including proliferation, differentiation, neurite outgrowth, synaptogenesis, migration, neural network formation and function, and apoptosis, are all related. Blue bolded and underlined text (hNP1, hN2, 1° rat cortical, MEA-NFA) indicate DNT-NAM technologies described in this Issue Paper and with currently available data as analyzed using the ToxCast Data pipeline. Green bolded and underlined text (for iCell neurons in neurite growth assays) indicate a cell-assay combination that has been evaluated by EPA, but without data available for this Issue Paper. DNT processes annotated with assay abbreviations in red text indicate DNT-NAMs that have been preliminarily developed but that do not have available screening data for evaluation (NPC1-5, UKN2, UKN4, UKN5, hiPSC-NFA). Cell line abbreviations are decoded as follows: hNP1 = human neural progenitor cell line; hN2 = human neuronal cell line; 1° rat cortical = primary rat cortical neurons; NPC1-4 = human primary neuroprogenitor cells (Lonza) in neurosphere cultures; NPC5 = human glial cells in neurosphere cultures; UKN2 = human neural crest from h9 embryonic stem cells; UKN4 = Lund human mesencephalic human embryonic neuronal precursor (LUHMES) cells; UKN5 = human peripheral nervous system cells (immature dorsal root ganglion) cells from h9 embryonic cells. *Figure adapted from Aschner et al. (2016).*



### 2.3.3 DNT-NAM experimental methodology for OPs

The experimental methodologies described in the following sections pertain to data collected using OP insecticides and summarized in this Issue Paper. While the data on the OP insecticides presented here have not been published to date, data presented here on other chemicals have been made publicly available, embedded within larger datasets, through the peer-reviewed literature. All of the data for assay controls and OP insecticides discussed in this Issue Paper are provided as Supplemental Files.

*Chemical procurement.* The OP insecticides were procured for both the MEA NFA and HCI cellular event assays simultaneously. Aliquots of each OP or OP metabolite, solubilized in dimethyl sulfoxide (DMSO), were supplied with a target concentrations of 20 or 100 mM in 50  $\mu$ L aliquots in sealed, round-bottom 96-well plates by Evotec (Princeton, NJ) as part of the ToxCast Chemical Inventory (Richard et al., 2016). These plates were wrapped in parafilm and stored at 20°C until use.

#### 2.3.3.1 MEA NFA experiments

*Tissue culture and cell source.* The MEA NFA experiments summarized in this work used primary cultures of rat cortical neurons. These primary cultures of cortical neurons were prepared as described in Brown et al. (2016). Briefly, frontal cortex from postnatal 0–1 day Long-Evans rat pups were removed and dissociated, then plated in 25  $\mu$ L droplets containing 150,000 cells into each well of 48-well MEA plates (M-768-KAP-48; Axion Biosystems, Atlanta, GA). Cells were allowed to attach to the substrate for 2 hr after which 500  $\mu$ L of media was added that contained the appropriate concentration of test material or solvent. Complete media changes were done on day *in vitro* (DIV) 5 and 9, and chemical was refreshed along with the media. For the vast majority of chemicals, test concentrations ranged from 0.03  $\mu$ M to 100  $\mu$ M in half-log increments, with the lower concentration bound reflective of practical limitations (e.g., the number of wells on the plate), laboratory resources (e.g., the cost and time of running additional concentrations) and the typical concentration range within which activity is observed, and the upper concentration reflective of DMSO limitations in the assay and a desire to test within a relevant concentration range.

*Electrophysiological assessment of network activity.* Assessment of electrical activity using MEAs was conducted as described in Shafer et al. (2019), Frank et al. (2017), and Brown et al. (2016). Briefly, spontaneous electrical activity of primary rat cortical cells was recorded using the Axion Biosystems Maestro 768-channel amplifier, Middle-man data acquisition interface, and Axion Integrated Studio (AxIS) software v1.9 or later (Axion Biosystems, Atlanta, GA). Spontaneous activity was measured using a gain of 1200x and a sampling frequency of 12.5 kHz. The signal was passed through a Butterworth band-pass filter (0.1–5000Hz) in AxIS and on-line spike detection (threshold = 8 x root mean square (rms) noise) was done with the AxIS adaptive spike detector. All recordings were conducted at 37°C. Activity was recorded for a period of 15 min on DIV 5, 7, 9, and 12 to monitor the development of network activity. Recordings were always made prior to media changes to avoid potential impacts of physical perturbations on network activity. Each biological replicate was comprised of cells from several animals. Each biological replicate screened for a given chemical consisted of triplicate technical replicates (the entire concentration range at 1 well/concentration tested across 3 separate plates from the same culture [n=3/concentration, where n is in reference to technical replicates]).

*Cell viability measurements.* After recording on DIV 12, cell viability was assessed using lactate dehydrogenase (LDH) and alamar blue (AB) assays (Promega, Madison, WI) as described in (Brown et al., 2016; Frank et al., 2017; Shafer et al., 2019).

*Data analysis.* Analysis of the development of spontaneous electrical activity was conducted as described in Shafer et al. (2019) and Frank et al. (2017) using the R programming language (R Core Team, 2016)

for statistical computing and the ToxCast Data Pipeline (tcpl), with some revisions as described herein. Briefly, 17 different parameters (referred to as assay components in tcpl, see Table 1) describing spiking, bursting, and coordinated network activity were determined by analysis of each recording on each DIV. In area under the curve (AUC) determination, curve-fitting and EC<sub>50</sub> determinations were conducted as described in the original published manuscripts (Brown et al., 2016; Frank et al., 2017; Shafer et al., 2019). For the analyses presented herein, trapezoidal AUC (R trapz() function) was used to integrate responses across time and concentration; these AUC values were then supplied as multi-concentration level 0 (mc0) data to tcpl (version 2.0.2). For the cytotoxicity assays, the raw viability assay responses were supplied to tcpl. Then, tcpl was used to normalize data by assay component, and generate concentration-response curves and make hit-calls for each assay endpoint. Analysis and curve-fitting in the up and down directions was performed for the neuronal network formation components, resulting in 34 assay endpoints for neuronal network formation (17 x 2) and 2 assay endpoints for cell viability assessment (LDH and AB) for an overall total of 36 assay endpoints. The tcpl methods applied for the MEA-NFA technology are listed in Table 4. Cutoffs were set based on three times the baseline median absolute deviation (BMAD), where BMAD is an approximation of baseline noise using the vehicle control wells and the two lowest concentrations of test chemical wells on each plate. The BMAD values for assay endpoints in the MEA NFA ranged from 6.9 to 26.4%, resulting in cutoff values of 21 to 79% change.

### 2.3.3.2 HCI assay experiments

*Screening design.* Each chemical was tested across the complete concentration range with well replicates (n=3, where n is in reference to technical replicates) performed within the same experiment across multiple plates (Malo et al., 2006). Some chemicals were tested twice in the rat cortical models for a total of n=6 well replicates. Assay-positive control chemicals for proliferation (aphidicolin, 10 µM), apoptosis (staurosporine, 1 µM), neurite outgrowth (Rac 1 inhibitor (i.e., NSC 23766 trihydrochloride), 10 µM; lithium chloride, 30 mM or bisindolylmaleimide I, 3 µM) or synaptogenesis (sodium orthovanadate, 10 µM or bisindolylmaleimide I, 3 µM) were included on each plate (n = 3 wells / plate) to verify assay performance. After dilution of chemical stocks into media, the final concentration of DMSO or water vehicle in each assay was 0.1%, which had no effect on any cellular endpoint compared with media controls (data not shown).

*hNPI Proliferation Assay.* The assay for hNPI proliferation was performed as developed previously (Harrill et al., 2018; Mundy et al., 2010). Briefly, cells were seeded in clear 96-well plates pre-coated with poly-L-ornithine and laminin at a density of 10,000 cells/well. Two hours after plating, cells were exposed to chemicals and returned to the incubator. Twenty hours after chemical treatment, bromodeoxyuridine (BrdU; Millipore) was added directly to the media in each well to a final concentration of 50 µM and the cells incubated for a further 4 hours. Cells were then fixed with 4% paraformaldehyde for 15 min at room temperature, followed by permeabilization and blocking steps. Dividing cells that incorporated BrdU were labeled using secondary antibodies (Harrill et al., 2018). Hoechst 33342 dye (3 µg/ml) was included in the last step to label all nuclei. Quantitative fluorescence imaging was performed with Cellomics ArrayScan VTI HCS system (ThermoFisher Scientific) using a 10x objective. Fluorescence images were obtained in two channels: channel 1 was used to image Hoechst-labeled nuclei and channel 2 was used to image nuclei exhibiting BrdU labeling. Cells with BrdU labeling intensity >3-fold background were categorized as responders. The number of responders divided by the total number of nuclei in a well was used to calculate percent responders. Images were analyzed using the Cellomics Target Activation Bioapplication (v4). A sufficient number of images were analyzed from each well to count at least 300 cells. Cell viability was based on number of cells per field.

*hNPI Apoptosis Assay.* Apoptosis in hNPI cells was assessed as described previously (Druwe et al., 2015; Harrill et al., 2018). Briefly, cells were seeded in white opaque 96-well plates pre-coated with poly-

L-ornithine and laminin at a density of 10,000 cells/well. Two hours after plating, cells were exposed to chemicals and returned to the incubator for 24 hr. Apoptosis was measured based on detection of activated caspase-3/7 using the Caspase-Glo 3/7 assay kit (Promega). Cell viability was also determined in sister plates using the CellTiter-Glo luminescent assay (Promega) which measures the amount of ATP present in each well. Cells were seeded in white opaque 96-well plates pre-coated with poly-L-ornithine and laminin, allowed to adhere for 2 hr, and then exposed to chemicals for a 24 hr period before conducting the CellTiter-Glo assay per Promega directions and as described previously (Druwe et al., 2015; Harrill et al., 2018).

*hN2 and Rat Cortical Neurite Outgrowth Assays.* Initiation and the early phase of neurite outgrowth was measured as described in (Harrill et al., 2011a). Briefly, rat primary cortical neurons or human hN2 neurons were seeded in clear 96-well plates pre-coated with poly-L-lysine (rat primary cortical neurons) or poly-L-lysine and laminin (human hN2 neurons) at a density of 10,000 cells/well. Two hours after plating, cells were exposed to chemicals and returned to the incubator. Forty-eight hours after chemical treatment, cells were fixed with warm (37° C) 4% paraformaldehyde containing 1.5 µg/ml Hoechst 33342 for 20 min followed by permeabilization and blocking steps. Cell bodies and neurites (i.e., axons and dendrites) expressing the cytoskeletal protein  $\alpha$ <sub>III</sub>-tubulin were labeled using primary and secondary antibodies as described previously (Harrill et al., 2018). The Cellomics ArrayScan VTi HCS system was used for automated image acquisition and analysis of neurite outgrowth. Images were acquired using a 20x objective in two channels: channel 1 was used to image Hoechst-labeled nuclei; channel 2 was used to image cell bodies and neurites. Images were analyzed using the Cellomics Neuronal Profiling BioApplication (v4) to measure neurite morphology in each cell model (human and rat). Parameters for analysis of neurite outgrowth were described previously (Harrill et al., 2011a). In each well, multiple images were acquired until at least 300 neurons were counted. Four morphological features were quantified: 1) total neurite length per neuron; 2) number of neurites per neuron; 3) number of neurite branch points per neuron; and 4) number of cells (neurons) per field. Neurites were defined as processes > 10 µm in length. The number of cells per field was used as an indicator of cell viability at the time of fixation. Well-level population averages were used as the statistical unit of measure.

*Rat Cortical Synaptogenesis Assay.* Maturation of dendrites and formation of synapses was measured as described (Harrill et al., 2011b). Briefly, rat primary cortical neurons were seeded in clear 96-well plates pre-coated with poly-L-lysine at a density of 10,000 cells/well. At 5 DIV half of the media in each well was replaced with fresh media containing cytosine arabinoside for at a final concentration of 1 µM. On DIV 7, a complete media replacement was performed with fresh media (no cytosine arabinoside) followed by chemical exposure. Cells were then cultured for 5 days with no further media exchange. On DIV 12, cells were fixed with warm (37° C) 4% paraformaldehyde containing 1.5 µg/ml Hoechst 33342 for 20 min followed by permeabilization and blocking steps. Cell bodies and dendrites expressing microtubule-associated protein 2 (MAP2) and presynaptic terminals containing synapsin I protein were labeled using primary and secondary antibodies as previously described (Harrill et al., 2018). The Cellomics ArrayScan VTi HCS system was used for automated image acquisition and analysis of dendrite growth and synapse number. Images were acquired using a 20x objective in three channels: channel 1 was used to image Hoechst-labeled nuclei; channel 2 was used to image cell bodies and dendrites; channel 3 was used to image presynaptic puncta. Images were analyzed using the Cellomics Neuronal Profiling BioApplication (v4). Parameters for analysis of the cell nucleus, cell body, dendrite and synaptophysin puncta identification and selection, as well as cell body masking and dendrite tracing parameters were described previously (Harrill et al., 2011b). In each well, multiple images were acquired until at least 300 neurons were counted. Seven morphological features were quantified: 1) total neurite length per neuron; 2) number of neurites per neuron; 3) number of neurite branch points per neuron; 4) the number of cell

body-associated presynaptic puncta; 5) the number of neurite-associated presynaptic puncta; 6) total number of synaptic puncta per neuron; 7) number of synaptic puncta per unit neurite length; and 7) number of cells (i.e., neurons) per field. Well-level population averages were used as the statistical unit of measure.

*Data analysis.* For all HCI assays, complete concentration-response data were generated within a 96-well plate using one well per concentration, and experiments repeated across three culture plates (n=3). Concentration-response data were analyzed and curve-fit using R software for statistical computing and the tcpl package (version 2.0.2) (Filer et al., 2017). The raw values from the HCI assays were used as mc0 data for tcpl. All of the tcpl methods applied to the HCI assays are detailed in Table 4. All data were normalized on a plate-by-plate basis to the median of the vehicle control wells and two lowest concentrations of test chemical wells. The cutoff for a positive response or hit-call was set as the greater of 30% or 3 times the BMAD, where BMAD is an approximation of baseline noise using the vehicle control wells and the two lowest concentrations of test chemical wells. BMAD ranged from 1.5 to 14.9%, and thus the cutoffs for a positive response in each assay endpoint ranged from 30 to 44.7%.

**Table 4. ToxCast Data Pipeline for MEA-NFA and HCI assays.**

<b>ToxCast Data Pipeline Level</b>	<b>MEA-NFA: Methods Applied</b>	<b>HCI assays: Methods Applied</b>
mc0: pre-processed data input	Data are pre-processed to obtain AUC values by assay component	Data are raw input
mc1: mapping to well and column indexes	Auto	
mc2: transformation	No transformation	
mc3: normalization	Baseline value (bval) was calculated as the median value for the vehicle control wells (DMSO) on a by-plate basis; No positive control value was used in normalization (pval=0); the response was calculated as percent of DMSO vehicle control. The response was multiplied by -1 for the “up” endpoints such that all endpoints are curve-fit in the positive direction.	Baseline value (bval) was calculated as the median value for the vehicle control wells (DMSO) on a by-plate basis; No positive control value was used in normalization (pval=0); the response was calculated as percent of DMSO vehicle control.
mc4: BMAD calculation type for curve-fitting	An approximation of noise around the baseline signal, the baseline median absolute deviation, was calculated based on the vehicle control wells and the 2 lowest concentrations of the test wells on each plate.	An approximation of noise around the baseline signal, the baseline median absolute deviation, was calculated based on the vehicle control wells and the 2 lowest concentrations of the test wells on each plate.
mc5: Hitcall and potency determination	The cutoff for a positive response in each assay endpoint was set as 3*BMAD.	The cutoff for a positive response was the greater of 30% or 3*BMAD.
mc6: caution flags on fitting	Flags for single point hit at maximum concentration (6), flags for single point hit not at the maximum concentration screened (7), inactives with multiple median responses above baseline (8), noisy curves relative to the assay (10), actives with borderline efficacy (11), inactives with borderline efficacy (12), low concentration gain-loss curve-fits (15), possibly overfitting (16), hitcalls with less than 50% efficacy (17) were assigned to all; additionally cell viability assays were assigned “viability gain-loss fit” (19)	

### 2.3.4 DNT-NAM experimental validation using assay positive controls

The MEA NFA and HCI assays have been previously published, but the reproducibility of the assays, as demonstrated by variability in the vehicle-control wells assayed in each assay component, has not been reported previously. In this section, the reproducibility of the assays, and the performance of assay controls as probe substrates known to disrupt key neurodevelopmental processes *in vitro*, are discussed. Although both the MEA NFA and HCI assays are comprised by multiple individual assay components (data before direction-dependent analysis) and assay endpoints (data after direction-dependent analysis), it should be noted that the MEA NFA and HCI assay components/endpoints are considered collectively; the components/endpoints encompassed by the MEA NFA or HCI assays are being applied as a group to understand putative *in vitro* DNT-related hazard.

#### 2.3.4.1 Assay reproducibility

##### 2.3.4.1.1 Assay reproducibility based on DMSO vehicle control wells

Several quantitative indicators of assay reproducibility were calculated to understand the performance of MEA NFA and HCI assays. First, the reproducibility of the vehicle control (i.e., DMSO) treated wells was evaluated using the median, the median absolute deviation (MAD), and coefficient of variation (CV) to understand the dispersion of the distribution of raw results for DMSO-treated wells in the MEA NFA (Table 5) and HCI assays (Table 6). For the MEA NFA, the average CV was 16% across all 18 assay components, with a range of 6.57-25.2%. For the HCI assays, the average CV was 8.74% across all 21 assay components, with a range of 3.66-16.6%. The CV of the vehicle control wells in these assays are generally less than 20% and suggest that these assays are reproducible. The variability in baseline is accounted for in determination of assay endpoint positives, i.e., chemical-induced changes considered positive must be separated from any estimate of the baseline response. For determination of positive assay endpoint hit-calls, in the MEA NFA a cutoff was set at activity three times the BMAD of the responses from the vehicle control and two lowest test concentrations on the plate (Table 5); these cutoff values then ranged from 21 to 79% changes. For the HCI assays, the minimum of a 30% change or three times the BMAD was used as a cutoff for a positive hit-call, resulting in cutoff values that ranged from 30 to 44.7% changes (Table 6).

**Table 5. Assay reproducibility using DMSO vehicle control wells for the MEA NFA**

The median, median absolute deviation (MAD), and coefficient of variation (CV, in units of %) were calculated for the pre-processed area-under-the-curve values (ToxCast multi-concentration level 0) from the DMSO-treated wells of each assay component. For the MEA NFA, an area-under-the-curve was used to compress the effects observed over time at each concentration.

ACID	Assay component name	Median	MAD	CV
2471	NHEERL_MEA_dev_firing_rate_mean	10.84	2.12	20.5
2472	NHEERL_MEA_dev_burst_rate	20.04	3.48	18.8
2473	NHEERL_MEA_dev_active_electrodes_number	98.12	6.3	7.63
2474	NHEERL_MEA_dev_bursting_electrodes_number	78.88	7.41	9.20
2475	NHEERL_MEA_dev_per_burst_interspike_interval	0.3	0.07	25.2
2476	NHEERL_MEA_dev_per_burst_spike_percent	457.46	38.17	9.57
2477	NHEERL_MEA_dev_burst_duration_mean	4.18	0.81	23.1

ACID	Assay component name	Median	MAD	CV
2478	NHEERL_MEA_dev_interburst_interval_mean	207.8	48.58	22.8
2479	NHEERL_MEA_dev_network_spike_number	343.88	55.23	21.3
2480	NHEERL_MEA_dev_network_spike_peak	79.02	5.68	7.31
2481	NHEERL_MEA_dev_spike_duration_mean	1.38	0.17	13.1
2482	NHEERL_MEA_dev_network_spike_duration_std	0.49	0.09	22.1
2483	NHEERL_MEA_dev_per_network_spike_interspike_interval_mean	199.07	53.52	24.0
2484	NHEERL_MEA_dev_per_network_spike_spike_number_mean	324.38	48.42	15.1
2485	NHEERL_MEA_dev_per_network_spike_spike_percent	85.13	12.34	17.0
2486	NHEERL_MEA_dev_correlation_coefficient_mean	1.53	0.22	14.8
2487	NHEERL_MEA_dev_mutual_information_norm	0.05	0.01	19.6
2488	NHEERL_MEA_dev_LDH	1.05	0.07	7.66
2489	NHEERL_MEA_dev_AB	23014.9	1505.58	6.57

**Table 6. Assay reproducibility using DMSO vehicle control wells for the HCI assays**

The median, median absolute deviation (MAD), and coefficient of variation (CV, in units of %) were calculated for the raw values (ToxCast multi-concentration level 0) from the DMSO-treated wells of each assay component. Gray rows indicate assay components that are indicative of cytotoxicity.

ACID	Assay component name	Median	MAD	CV
2711	MUNDY_HCI_hNP1_Pro_MeanAvgInten	87.7	11.5	12.7
2710	MUNDY_HCI_hNP1_Pro_ResponderAvgInten	33.7	4.16	11.9
2709	MUNDY_HCI_hNP1_Pro_ObjectCount	92.9	8.45	8.51
2691	MUNDY_HCI_hNP1_Casp3_7	65300	2230	3.8
2700	MUNDY_HCI_hNP1_CellTiter	714000	18200	1.96
2695	MUNDY_HCI_hN2_NOG_BPCount	0.56	0.0741	13.5
2694	MUNDY_HCI_hN2_NOG_NeuriteCount	1.66	0.0593	3.63
2693	MUNDY_HCI_hN2_NOG_NeuriteLength	75.2	4.76	7.15
2692	MUNDY_HCI_hN2_NOG_NeuronCount	23.4	3.56	13.5
2699	MUNDY_HCI_Cortical_NOG_BPCount	3.47	0.282	8.36
2698	MUNDY_HCI_Cortical_NOG_NeuriteCount	2.95	0.0445	1.76
2697	MUNDY_HCI_Cortical_NOG_NeuriteLength	128	10.5	8.42
2696	MUNDY_HCI_Cortical_NOG_NeuronCount	26.3	1.81	7.6
2707	MUNDY_HCI_Cortical_Synap&Neur_Matur_BPCount	10.3	0.778	7.39
2706	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteCount	4.88	0.17	3.66
2705	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteLength	396	28.5	9.41
2702	MUNDY_HCI_Cortical_Synap&Neur_Matur_CellBodySpotCount	14.1	1.69	13.9
2703	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteSpotCountPerNeuron	64.2	12.3	16.2
2708	MUNDY_HCI_Cortical_Synap&Neur_Matur_SynapseCount	77.7	16.6	16.6
2704	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteSpotCount PerNeuriteLength	0.18	0.0148	8.24
2701	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuronCount	23.4	1.24	5.41

#### **2.3.4.1.2 Assay reproducibility based on repeated test chemical screening**

Additionally, in the MEA NFA assays, 21 test chemicals were screened across more than one experimental block using more than one sample of the chemical. Eight of these compounds were OP chemicals. This provides a means of understanding hit-call and potency reproducibility for test chemicals. No chemicals had been screened more than once in the HCI assays for this analysis.

For the MEA NFA repeated test chemical screening, the results of the 21 test chemicals can be parsed into three response groups: 10/21 had repeated strong positive (both replicates with >3 assay endpoints consistently positive) or negative (all replicates with 0 assay endpoints positive) responses, 5/21 had equivocal responses (1 replicate with between 1 and  $\leq 3$  assay endpoints positive and one replicate with no assay endpoints positive), and 6/21 had mixed (1 replicate with positive and 1 replicate equivocal or negative) responses. Thus, from a qualitative perspective, the replicability of overall chemical hit-call was 71% (15/21). In terms of quantitative concordance, for the 9 of the 10 chemicals with repeated positive responses, the average standard deviation in log<sub>10</sub>-AC<sub>50</sub> values (in micromolar units) was less than 0.5 log<sub>10</sub>-micromolar, except for 2,2',4,4'-tetrabromodiphenyl ether, where there was more uncertainty in the potency values (average standard deviation across all AEIDs for this chemical was 1.71 log<sub>10</sub>-micromolar). Complete details on all 21 compounds can be found in Supplemental Appendix B, while information on the OP chemicals can be found in Section 2.3.5. There are several potential reasons why different samples of the same chemical have yielded mixed results, including sample stability, culture preparations, temporal differences and personnel changes.

#### **2.3.4.2 Assay performance with assay controls**

Assay performance for the MEA NFA and HCI assays was evaluated with assay controls known to cause effects on neural network formation and function, or HCI endpoints that inhibit neurodevelopmental processes (e.g., synaptogenesis, neurite outgrowth) *in vitro* (Table 7).

The performance of assay controls is visually summarized in Figures 3 and 4, with more data provided for the MEA NFA in Table 8 and for the HCI assays in Table 9. The approaches between the MEA NFA and HCI assays varied slightly, as the MEA NFA included training compounds in the screening set, but did not use on-plate controls due to throughput limitations, whereas on-plate controls were used for the HCI assays but not in a full concentration-response (thus curve-fitting could not be used to approximate AC<sub>50</sub> values). In the MEA NFA, the positive control substances (loperamide hydrochloride, bisindolylmaleimide I, L-domoic acid, mevastatin, and sodium orthovanadate) work to inhibit (decrease) neuronal network formation. In the HCI assays, similarly, the positive controls (aphidocholin, lithium chloride, bisindolylmaleimide I, sodium orthovanadate, NSC 23766 trihydrochloride, and staurosporine) all work to inhibit key critical neurodevelopmental processes (i.e., neurite outgrowth, synaptogenesis, proliferation) in specific cell types, save staurosporine, which reduces cell viability via apoptosis (Supplemental Appendix A).

**Table 7. Assay performance controls**

The DSSTox unique identifier (DTXSID), chemical name, assay activity anticipated, use of the control in the MEA-NFA or HCI Assays, and notes on the assay performance controls are provided in Table 8.

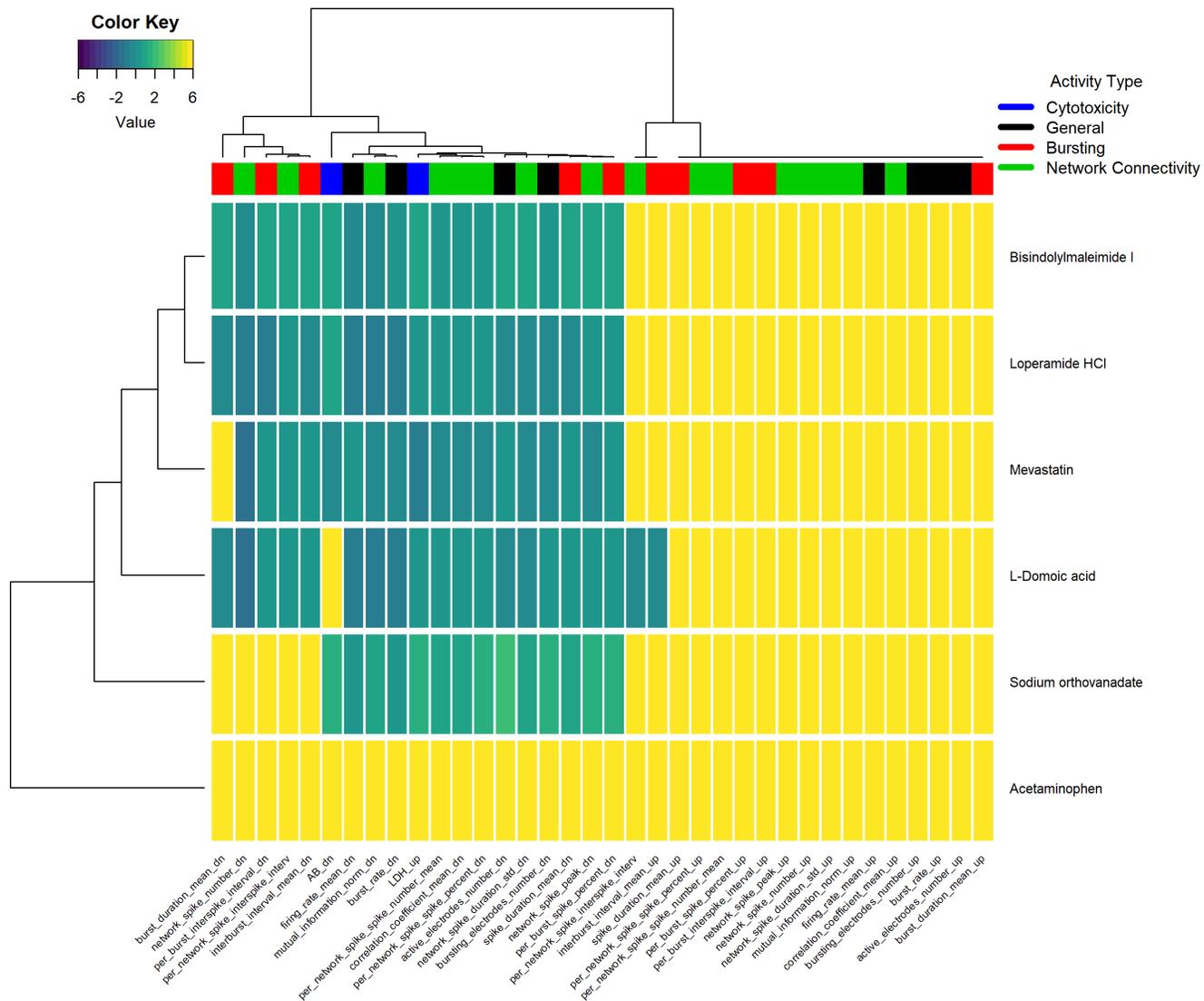
DTXSID	Chemical name	Assay activity	MEA NFA	HCI Assays	Notes
DTXSID2020006	Acetaminophen	Negative control (McConnell et al., 2012)	Yes	No	
DTXSID00880006	4-(4-Chlorophenyl)-4-hydroxy-N,N-dimethyl- $\alpha,\alpha$ -diphenylpiperidine-1-butylamide monohydrochloride [also known as Loperamide hydrochloride]	No evidence of DNT <i>in vivo</i> , cytotoxic and modulates ion transport (Einarson et al., 2000; Gong et al., 2012). Mode of action is agonist of the $\mu$ -opoid receptor (Vandenbossche et al., 2010)	Yes	No	<i>In vivo</i> , loperamide hydrochloride does not reach effective concentrations in the CNS due to rapid removal via p-glycoprotein transporters (Upton, 2007)(Upton, 2007).
DTXSID20274180	L-Domoic acid	Inhibited network activity and bicuculline response in cortical networks following <i>in vitro</i> developmental exposure (Hogberg et al., 2011)	Yes	No	Domoic acid is an agonist for glutamate receptors, most specifically with the AMPA receptor (Hampson and Manalo, 1998).
DTXSID4040684	Mevastatin	Inhibits synaptogenesis <i>in vitro</i> (Harrill et al., 2011a)	Yes	No	
DTXSID50157932	Bisindolylmaleimide I	Inhibits neurite outgrowth, synaptogenesis and ontogeny of network activity (Robinette et al., 2011)	Yes	Yes	Protein kinase C and GSK3 $\alpha$ inhibitor (Hers et al., 1999; Toullec et al., 1991). Selective inhibition of <i>in vitro</i> neurite outgrowth and synaptogenesis below the threshold for cytotoxicity (Harrill et al., 2011a; Robinette et al., 2011).
DTXSID2037269	Sodium orthovanadate	Inhibits neurite outgrowth and synaptogenesis <i>in vitro</i> (Harrill et al., 2011b)	Yes	Yes	Broad spectrum phosphatase inhibitor (Gordon, 1991; Pugazhenthil et al., 1996). Selective inhibition of neurite outgrowth and synaptogenesis below the threshold for cytotoxicity (Harrill et al., 2013; Harrill et al., 2011b).
DTXSID90328386	NSC 23766 trihydrochloride	Affects primary rat neurite outgrowth (Harrill et al., 2018)	No	Yes	Rac1 inhibitor (Gao et al., 2004). Selective inhibition of neurite outgrowth below the threshold for cytotoxicity (Druwe et al., 2016).
DTXSID2025509	Lithium chloride	Affects human (hN2) neurite outgrowth (Harrill et al., 2018)	No	Yes	GSK3b inhibitor (Kirshenboim et al., 2004; O'Brien et al., 2011; Ryves and Harwood, 2001; Stambolic et al., 1996). Selective inhibition of neurite outgrowth below the threshold for cytotoxicity (Harrill et al., 2011a)
DTXSID5036711	Aphidicolin	Affects proliferation (Harrill et al., 2018)	No	Yes	DNA polymerase inhibitor (Cheng and Kuchta, 1993; Goulian et al., 1990; Kota et al., 2012; Krokan et al., 1981; Wright et al., 2006). Prototypical inhibitor of cell proliferation (Kohno et al., 2006; Mundy et al., 2010; Rolls et al., 2007; Walton et al., 2006).
DTXSID6041131	Staurosporine	Induces apoptosis (Harrill et al., 2018)	No	Yes	Broad spectrum kinase inhibitor (Karaman et al., 2008; Ruegg and Burgess, 1989). Prototypical inducer of apoptosis in many different cell models (Bertrand et al., 1994; Druwe et al., 2015; Feng and Kaplowitz, 2002).

The approach to understanding assay performance with control chemicals seeks to evaluate the robustness of the assay performance controls; i.e., for chemicals with known *in vitro* activities (as detailed in Table 7 and Supplemental Appendix A) relevant to neurodevelopment. In a quantitative approach, several measures of how well a signal can be reliably distinguished from noise were quantified. A robust  $Z'$  factor was calculated (Zhang et al., 1999; Paul Friedman et al., 2016) for each assay endpoint. A  $Z'$  of 0.5-1.0 indicates an assay with sufficiently high signal-to-background distinction and an inter-sample variability low enough to consistently distinguish positive and negative test chemicals. Though this measure is commonly used to evaluate high-throughput assay quality, its interpretation herein should be qualified by the following: (1)  $Z'$  is largely influenced by the effect size of the positive control, and so for assay endpoints lacking high efficacy controls, the  $Z'$  will be diminished; (2) none of these assay endpoints are typically used individually as these assay endpoints are used collectively to understand the *in vitro* behavior of chemicals on neurodevelopmental processes or neuronal network formation. Thus, the  $Z'$  can be informative in understanding which assay endpoints have very efficacious, reproducible controls that result in large effect sizes. The strictly standardized mean difference (SSMD) (Bray and Carpenter, 2017) is another metric often used in high-throughput bioactivity screening quality control to distinguish positive hits from negative hits on the basis of effect size. Larger SSMD absolute values tend to correspond to larger  $Z'$  values. Signal-to-noise (SN) is another metric that indicates the positive control response from the baseline response, considering variability in that baseline. Again, for particular assay controls, the  $Z'$  for the assay endpoint will be largely influenced by the effect size and its variability.

For the MEA NFA, a heatmap summarizes the performance of assay controls (Figure 3) based on the  $AC_{50}$  potency values for each assay performance control across the suite of assay endpoints. Acetaminophen appears to be an optimal negative control, with no positive hit-calls in any assay endpoint. All of the other assay performance controls decrease general activity, bursting activity, and network connectivity measures, rather than increasing any of these activity types. The five chemicals used as assay positive controls (loperamide hydrochloride, bisindolylmaleimide I, L-domoic acid, mevastatin, and sodium orthovanadate) appear to affect all of the activity types in the MEA NFA, which may be due to the selection of these assay controls based on their ability to affect neurite outgrowth and synaptogenesis *in vitro* (Supplemental Appendix A). Evaluating assay performance controls quantitatively across all assay endpoints in the MEA NFA revealed that median  $Z'$  for all the assay performance controls ranged from 0.55 to 0.8, suggesting that these controls demonstrate reproducible, robust results in the MEA NFA. Effect size as indicated by SSMD and the SN tracked with the  $Z'$  values: larger effect sizes and SN ratios corresponded to higher  $Z'$  values, in particular demonstrating that bisindolylmaleimide I and L-domoic acid were generally the most efficacious and reproducible controls in the MEA NFA.

**Figure 3. Heatmap of MEA NFA assay performance control results.**

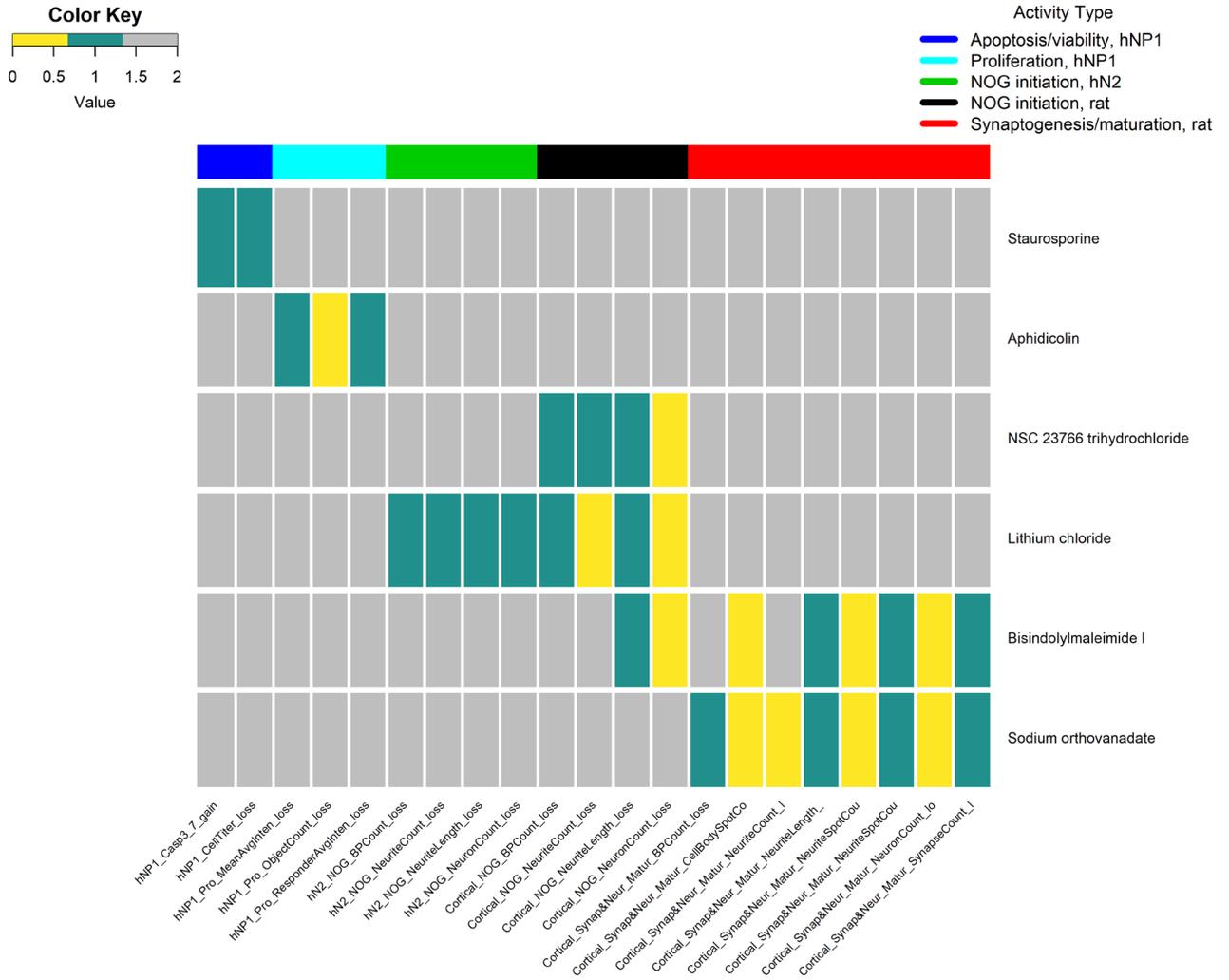
Activity ( $AC_{50}$  values) for the assay training set compounds. Each compound appears as a separate row, with separate columns for each assay parameter (listed along the bottom). Data are shown for each parameter in both the “up” and “down” directions. The type of activity measure is denoted by the different colors (black=general activity, blue=cytotoxicity, red=bursting, green=connectivity).



For the HCI assays, a review of how assay performance controls affect assay endpoint results is visually summarized in Figure 4 based on assay endpoint hit-call, as these assay controls were not screened in multi-concentration response, but rather as on-plate assay controls at a particular concentration. As not all assay controls were used in all assay endpoints/activity types, assay controls that were not screened in particular assay endpoints appear gray in the heatmap. Additional quantitative data to support Figure 4 is available in Table 9. Starting with the assay endpoints for apoptosis/viability, staurosporine appears to act as an optimal positive control, with robust  $Z'$  values  $>0.75$  for both assay endpoints and large effect sizes, as indicated by SSMD, that can be well distinguished from the assay endpoint baselines as indicated by SN ratio. For the proliferation assay component endpoints, aphidicolin produced effects but with very small effect sizes as indicated by SSMD and more variability as indicated by low SN and  $Z'$ , making aphidicolin a less optimal probe substrate but one that clearly affected hNP1 proliferation. For neurite outgrowth in hN2 cells, lithium chloride robustly affected the number of branch points and neurites, with smaller and less reproducible effect sizes for neurite length. In contrast, for the rat cortical neurite outgrowth assays, lithium chloride most robustly affects neurite length. Two other assay controls were used for the rat cortical neurite outgrowth assays: bisindolylmaleide I and a Rac GTPase inhibitor, NSC 23766 trihydrochloride (with bisindolylmaleide I only used to evaluate neurite length and count). Bisindolylmaleide I produced moderate effects on neurite length (in the absence of cytotoxic effects on neurite count), and NSC 23766 very robustly affected branch point and neurite count as well as neurite length, with  $Z'$  all  $\leq 0.5$  and large SSMD values. For the assay endpoints relevant to neurite maturation and synaptogenesis, bisindolylmaleide I and sodium orthovanadate were used. Bisindolylmaleide I demonstrated moderate effect sizes for branch point, neurite, and synapse count as well as neurite length, with concomitantly moderate  $Z'$  scores in the 0.3 range, suggesting bisindolylmaleide I is a reasonable control despite limitations on its efficacy at this non-cytotoxic concentration. In general, sodium orthovanadate performed poorly as a control: it demonstrated small effect sizes with more variability, and consequently demonstrated very low  $Z'$  values. Overall, the qualitative and quantitative results suggest the following: (1) for activity types including apoptosis/viability, hN2 neurite outgrowth, and rat primary cortical neurite outgrowth, optimal positive controls with high efficacy and highly reproducible responses have been identified and used to verify assay performance; (2) for activity related to cortical synaptogenesis/neurite maturation, moderate positive controls have been identified for some of the endpoints measured; and, (3) for activity related to hNP1 proliferation, the assay positive control produced effects but of low magnitude, resulting in less robust  $Z'$  values.

**Figure 4. Heatmap of HCI assay performance control results.**

The hit-calls [0=negative (yellow), 1 = positive (teal)] are illustrated for the on-plate assay controls for the HCI assay suite. Assay controls varied across the assay endpoints; controls not screened for a particular assay endpoint are annotated as hit-call=2 (gray). The median response of the assay control was used to derive a median hit-call (median response > COFF, median hit-call = 1) that is used in this figure. The activity type is annotated in the bar across the top of the heatmap (black = neurite outgrowth (NOG) initiation, rat; red = synaptogenesis/maturation in rat cells; green = neurite outgrowth (NOG) initiation in hN2 cells; dark blue = apoptosis and viability in hNP1 cells; and, light blue = proliferation in hNP1 cells). No clustering was performed due to the differences in the controls screened between assay endpoints.



**Table 8. Assessment of assay performance controls for the MEA NFA.**

The assay activity anticipated (Brown et al., 2016), the number of positive MEA NFA endpoints (out of 36 possible), and the minimum, median and maximum AC<sub>50</sub> values (in micromolar,  $\mu\text{M}$ ) are provided. Additionally, the median Z', median SSMD, and the median SN are provided for each assay control across the 17 MEA NFA assay endpoints that indicate decreased activity and the two parallel cytotoxicity assay endpoints. The range of the Z', SSMD, and SN for all assay endpoints is provided in parentheses.

DTXSID	Chemical name	# positive endpoints	Minimum AC <sub>50</sub> ( $\mu\text{M}$ )	Median AC <sub>50</sub> ( $\mu\text{M}$ )	Maximum AC <sub>50</sub> ( $\mu\text{M}$ )	Median Z'	Median SSMD	Median SN
DTXSID2020006	Acetaminophen	0	NA	NA	NA	NA	NA	NA
DTXSID00880006	4-(4-Chlorophenyl)-4-hydroxy-N,N-dimethyl-alpha,alpha-diphenylpiperidine-1-butyramide monohydrochloride [also known as Loperamide hydrochloride]	19	0.11	0.77	4.31	0.55 (0.29-0.89)	7 (4-28)	6.68 (4.23-27.91)
DTXSID50157932	Bisindolylmaleimide I	19	0.01	4.12	8.38	0.8 (0.33-0.94)	15 (4-48)	15.17 (4.47-47.88)
DTXSID20274180	L-Domoic acid	20	0.02	1.08	3.56	0.78 (0.26-0.94)	14 (4-48)	13.56 (4.04-47.88)
DTXSID4040684	Mevastatin	18	0.06	0.92	94.87	0.61 (-0.17-0.86)	8 (3-21)	7.74 (2.57-21.44)
DTXSID2037269	Sodium orthovanadate	14	1.98	15.04	81.26	0.74 (0.08-0.9)	12 (-10-32)	11.53 (-9.64-31.55)

**Table 9. Assessment of assay performance controls for the HCI assays.**

Assay endpoint ID (AEID), assay endpoint name (AENM), cutoff response for a positive hit-call (COFF), median response of the assay performance control chemical (MED.RESP), median hit-call for that assay performance control chemical in the AEID (MED.RESP > COFF, MED.HITC = 1), the assay performance control chemical name (CHEM), the DTXSID, micromolar concentration (CONC.UM), Z', strictly standardized mean difference (SSMD), and signal-to-noise ratio (SN) are reported. Some AEID used different assay performance controls; as such, the Z', SSMD, and SN are presented as the median across all plates for an AEID-CHEM combination. Gray rows with italicized text indicate assay endpoints that if positive indicate cytotoxicity; assay performance controls were intended to be negative for these assay endpoints (i.e., there was an attempt to screen assay performance control at concentrations below cytotoxicity).

Activity type	AEID	AENM	COFF	MED.R ESP	MED.H ITC	CHEM	DTXSID	CONC.UM	Z'	SSMD	SN
Apoptosis / viability	2793	MUNDY_HCI_hNP1_Casp3_7_gain	30	199.41	1	Staurosporine	DTXSID6041131	1	0.8	19	54.59
	2794	<i>MUNDY_HCI_hNP1_CellTiter_loss</i>	<i>30</i>	<i>70.3</i>	<i>1</i>	<i>Staurosporine</i>	<i>DTXSID6041131</i>	<i>10</i>	<i>0.75</i>	<i>18</i>	<i>19.69</i>
Proliferation	2795	MUNDY_HCI_hNP1_Pro_MeanAvgInten_loss	31.03	47.12	1	Aphidicolin	DTXSID5036711	10	0	3	2.99
	2796	<i>MUNDY_HCI_hNP1_Pro_ObjectCount_loss</i>	<i>30</i>	<i>20.83</i>	<i>0</i>	<i>Aphidicolin</i>	<i>DTXSID5036711</i>	<i>10</i>	<i>0</i>	<i>1</i>	<i>1.57</i>
	2797	MUNDY_HCI_hNP1_Pro_ResponderAvgInten_loss	34.29	87.14	1	Aphidicolin	DTXSID5036711	10	0.1	4	4.77
NOG, hN2	2789	MUNDY_HCI_hN2_NOG_BPCount_loss	37.7	79.34	1	Lithium chloride	DTXSID2025509	10000	0	3	3.62
	2790	MUNDY_HCI_hN2_NOG_NeuriteCount_loss	30	43.53	1	Lithium chloride	DTXSID2025509	10000	0.38	8	10.45
	2791	MUNDY_HCI_hN2_NOG_NeuriteLength_loss	30	62.56	1	Lithium chloride	DTXSID2025509	10000	0.58	9	13.25
	2792	<i>MUNDY_HCI_hN2_NOG_NeuronCount_loss</i>	<i>32.45</i>	<i>44.25</i>	<i>1</i>	<i>Lithium chloride</i>	<i>DTXSID2025509</i>	<i>10000</i>	<i>0</i>	<i>2</i>	<i>3.63</i>
NOG, rat cortical	2779	MUNDY_HCI_Cortical_NOG_NeuriteLength_loss	30	42.66	1	Bisindolylmaleimide I	DTXSID50157932	3	0.31	4	4.76
	2780	<i>MUNDY_HCI_Cortical_NOG_NeuronCount_loss</i>	<i>30</i>	<i>18.52</i>	<i>0</i>	<i>Bisindolylmaleimide I</i>	<i>DTXSID50157932</i>	<i>3</i>	<i>0</i>	<i>0</i>	<i>-0.1</i>
	2777	MUNDY_HCI_Cortical_NOG_BPCount_loss	30	48.39	1	Lithium chloride	DTXSID2025509	3	0.06	3	9.07
	2778	MUNDY_HCI_Cortical_NOG_NeuriteCount_loss	30	24.66	0	Lithium chloride	DTXSID2025509	3	0	3	3.58
	2779	MUNDY_HCI_Cortical_NOG_NeuriteLength_loss	30	42.66	1	Lithium chloride	DTXSID2025509	3	0.3	5	6.98

	2780	MUNDY_HCI_Cortical_NOG_NeuronCount_loss	30	18.52	0	Lithium chloride	DTXSID2025509	3	0	2	2.5
	2777	MUNDY_HCI_Cortical_NOG_BPCount_loss	30	48.39	1	NSC 23766 trihydrochloride	DTXSID90328386	3	0.5	8	9.96
	2778	MUNDY_HCI_Cortical_NOG_NeuriteCount_loss	30	24.66	0	NSC 23766 trihydrochloride	DTXSID90328386	3	0.71	14	46.88
	2779	MUNDY_HCI_Cortical_NOG_NeuriteLength_loss	30	42.66	1	NSC 23766 trihydrochloride	DTXSID90328386	3	0.53	8	11.73
	2780	MUNDY_HCI_Cortical_NOG_NeuronCount_loss	30	18.52	0	NSC 23766 trihydrochloride	DTXSID90328386	3	0	1	2.57
Synaptogenesis and neurite maturation, rat cortical	2782	MUNDY_HCI_Cortical_Synap&Neur_Matur_CellBodySpotCount_loss	44.68	-25.1	0	Bisindolylmaleimide I	DTXSID50157932	3	0	-1	-1.41
	2784	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteLength_loss	30	68.81	1	Bisindolylmaleimide I	DTXSID50157932	3	0.3	5	4.91
	2785	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteSpotCountPerNeuriteLength_loss	30	0.16	0	Bisindolylmaleimide I	DTXSID50157932	3	0	0	-0.01
	2786	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteSpotCountPerNeuron_loss	36.58	68.21	1	Bisindolylmaleimide I	DTXSID50157932	3	0.29	5	6.31
	2787	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuronCount_loss	30	-16.98	0	Bisindolylmaleimide I	DTXSID50157932	3	0	-2	-2.88
	2788	MUNDY_HCI_Cortical_Synap&Neur_Matur_SynapseCount_loss	31.17	62.52	1	Bisindolylmaleimide I	DTXSID50157932	3	0.34	5	5.99
	2781	MUNDY_HCI_Cortical_Synap&Neur_Matur_BPCount_loss	30	45.56	1	Sodium orthovanadate	DTXSID2037269	10	0	3	5.14
	2782	MUNDY_HCI_Cortical_Synap&Neur_Matur_CellBodySpotCount_loss	44.68	-25.1	0	Sodium orthovanadate	DTXSID2037269	3	0	0	0.62
	2783	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteCount_loss	30	9.09	0	Sodium orthovanadate	DTXSID2037269	10	0	2	2.87
	2784	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteLength_loss	30	68.81	1	Sodium orthovanadate	DTXSID2037269	3	0	3	5.77
	2785	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteSpotCountPerNeuriteLength_loss	30	0.16	0	Sodium orthovanadate	DTXSID2037269	3	0	1	0.67
	2786	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuriteSpotCountPerNeuron_loss	36.58	68.21	1	Sodium orthovanadate	DTXSID2037269	3	0	2	1.95
	2787	MUNDY_HCI_Cortical_Synap&Neur_Matur_NeuronCount_loss	30	-16.98	0	Sodium orthovanadate	DTXSID2037269	3	0	1	1.99
	2788	MUNDY_HCI_Cortical_Synap&Neur_Matur_SynapseCount_loss	31.17	62.52	1	Sodium orthovanadate	DTXSID2037269	3	0	1	1.28

## 2.3.5 DNT-NAM experimental results for OPs

In this section, the approach along with results and discussion for OPs in the MEA NFA and HCI assays are presented.

### 2.3.5.1 Approach to experimental result analysis

#### 2.3.5.1.1 Dataset

The dataset used here was derived from experiments and data analysis as described in Section 2.3.3. Additionally, the dataset was filtered to remove curve-fit information for less reproducible curve-fits, namely curve-fitting information was not used if the curve-fit was associated with 3 or more caution flags (from level 6 of multi-concentration data analysis using tcpl) and/or a fit category of 36 or 45 (from level 5 of multi-concentration data analysis using tcpl) that would indicate that the calculated AC50 was less than the concentration range screened. For these less reproducible curve-fits, the hit-call was changed to negative (0) and the AC50 value was set to NA.

#### 2.3.5.1.2 Heatmap visualizations

The specific OP chemical sample and its activity in the MEA NFA and HCI assays are illustrated in Figures 5 and 6, respectively. The OP substance results are presented using the  $\log_{10}$ -AC50 ( $\mu\text{M}$ ) values used in the main heatmap images, with columns indicating assay endpoint names and rows indicating OP substance name and sample id (spid). Darker blue corresponds to lower  $\log_{10}$ -AC50 values, and a value of 6 (represented by yellow) has been assigned for substances with a negative hit-call for the assay endpoint. The column assay endpoint names are abbreviated by removing the “NHEERL\_MEA\_dev\_” or “MUNDY\_HCI\_” from the beginning and trimming to 50 characters for ease of reading. The activity type for each assay endpoint is also annotated for the columns. For row annotations, OP substances that have an oxon structural moiety have been annotated by row (gray). Hierarchical clustering was performed using Ward.D2 methodology. In these visualizations, any trends for OP chemicals in terms of the assay endpoints affected, the potency with which assay endpoints were affected, and the impact of the oxon structural moiety and/or performance of parent/metabolite pairs can be examined.

#### 2.3.5.1.3 Potency, selectivity, and a metric of relative activity

The effects of OP chemical samples across both the MEA NFA and HCI assay technologies were considered in more detail via examination of the potency ranges, presence of any assay endpoint effects at concentrations lower than cytotoxicity (referred to as selectivity), and a metric (referred to as the scaled AUC sum) to consider the relative potency and efficacy across the set of 27 OP substances are presented in Tables 12 and 13 for the MEA NFA and HCI assays, respectively.

*Potency.* For the MEA NFA and HCI assays, OP chemical potency was indicated by the minimum and median AC50 values for any positive assay endpoints in the assay technology. This potency range indicates at the highest level the concentration range that OP chemicals would be anticipated to be positive in the DNT-NAM battery in this Issue Paper.

*Selectivity.* For the MEA NFA assays, a selectivity score was calculated for the entire assay technology using the minimum cytotoxicity potency value (minimum of  $\log_{10}$ -AC50 ( $\mu\text{M}$ ) from LDH and AB assay endpoints) minus the minimum potency value ( $\log_{10}$ -AC50 ( $\mu\text{M}$ )) for all positive assay endpoints. Selectivity scores of  $> 0.3$  likely indicate some selective activity of the OP chemical in the assay, i.e., the lowest concentration-related effects on general activity, bursting activity, and/or network connectivity occurred at lower concentrations than cytotoxicity. If there were no positive assay endpoints, no selectivity score could be calculated. If the cytotoxicity assay endpoints were negative, a selectivity score

was inferred using the maximum log<sub>10</sub>-concentration of the OP chemical screened minus the minimum potency value (log<sub>10</sub>-AC<sub>50</sub> (μM)) for all positive assay endpoints.

For the HCI, separate selectivity values were calculated for: neurite outgrowth (NOG) with hN2 cells; NOG with rat primary cortical cells; neurite maturation with rat primary cortical cells; synaptogenesis with rat primary cortical cells; and, proliferation with hNP1 cells. For NOG, neurite maturation, and synaptogenesis assay endpoints, the cytotoxicity potency used to calculate selectivity was taken from the assay endpoints that evaluated “neuron count.” For the proliferation assays in hNP1 cells, the potency value in the hNP1\_Pro\_ObjectCount\_loss assay was used as the cytotoxicity potency. Then, the minimum log<sub>10</sub>-AC<sub>50</sub> value for assay endpoints for each of these activity types was subtracted from the corresponding cytotoxicity log<sub>10</sub>-AC<sub>50</sub> value for a selectivity score. Just like the MEA NFA, selectivity scores of > 0.3 likely indicate some selective activity of the OP chemical in the HCI assay subset. If the appropriate cytotoxicity assay endpoint was negative, the selectivity score was inferred using the maximum log<sub>10</sub>-concentration of the OP chemical screened minus the minimum potency value (log<sub>10</sub>-AC<sub>50</sub> (μM)). Selectivity could not be calculated for the apoptosis/cell viability assays in hNP1.

*Scaled area under the curve (AUC) sum.* The objective of the scaled AUC sum was to provide a metric that could be used to consider the relative *in vitro* bioactivity, including both potency and efficacy, of the 27 OP chemicals in this set. No other reference chemicals have been used to derive this relative comparison. First, the area under the curve (AUC) for each positive assay endpoint in the MEA NFA and the HCI assays was calculated using the curve-fitting information from tcpl. A scaled AUC sum was calculated for each OP sample for (1) the MEA NFA assay technology and (2) the HCI assay technology via the following equation (Eq. 1):

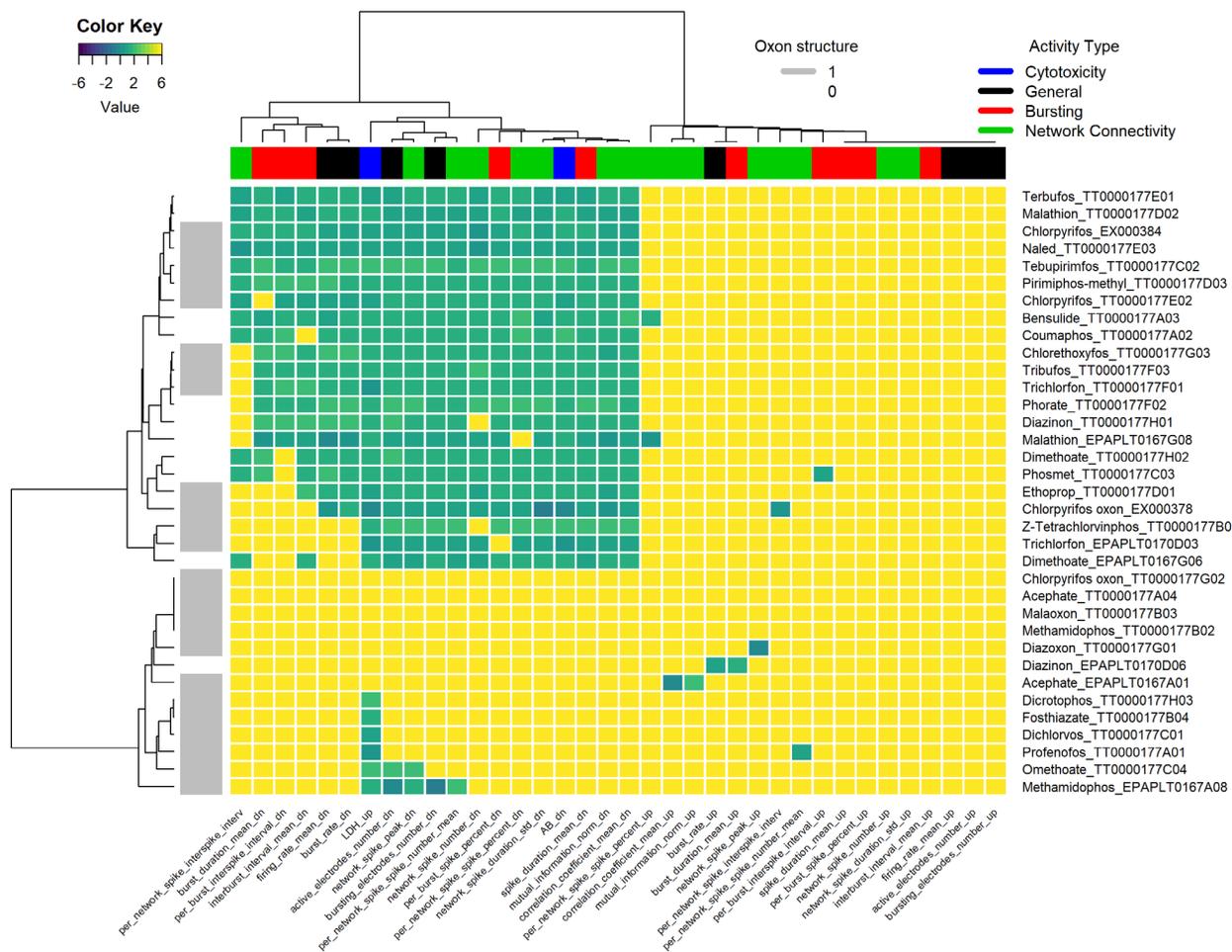
$$Eq\ 1.: \quad scaled\ AUC\ sum_{MEA\ NFA,\ HCI} = \frac{\log_2(\sum AUC)}{95th\ percentile(\log_2(\sum AUC))}$$

The scaled AUC sum for each assay technology (MEA NFA or HCI) is scaled only to the other OP samples in this dataset for that assay technology; the 95<sup>th</sup> percentile AUC sum value from the OP dataset for the assay technology (MEA NFA or HCI) was used to scale the values so that they would fall roughly from 0 to 1.

### 2.3.5.2 Heatmap visualizations of potency results for OPs

**Figure 5. Heatmap of MEA NFA experimental results for OPs.**

The OP substance results are presented using the log<sub>10</sub>-AC<sub>50</sub> (μM) values used in the main heatmap image, with columns indicating assay endpoint names and rows indicating OP substance name and sample id (spid), resulting in 35 rows for 27 unique substances. Darker blue corresponds to lower log<sub>10</sub>-AC<sub>50</sub> values, and a value of 6 (represented by yellow) has been assigned for substances with a negative hit-call for the assay endpoint. The activity type (cytotoxicity = blue, general activity = black, bursting = red, network connectivity = green) for each assay endpoint is also annotated for the columns. For row annotations, OP substances that have an oxon structural moiety have been annotated by row (gray).



For the MEA NFA, OPs generally decreased assay endpoint responses, much like the known assay performance controls. Very few OP chemicals elicited any increased activity, and the disparate observations of increased activity (never more than 2 assay endpoints demonstrated increased activity for a substance) suggest that OP chemicals active in the MEA NFA generally act to decrease network formation and function. However, there was no indication that OP samples affected general activity, bursting activity, network connectivity, or cytotoxicity assay endpoints with any differential pattern. This is analogous to the observations of the assay performance controls. In terms of potency, nearly all of the log<sub>10</sub>-AC<sub>50</sub> values for OP samples ranged from approximately 0 to 2 (1 to 100 μM) in the MEA NFA,

with the exception of one sample of methamidophos and one sample of chlorpyrifos oxon which demonstrated some log<sub>10</sub>-AC<sub>50</sub> values less than 0 (in the 0.1-1 μM range). Approximately three groups are apparent from hierarchical clustering of the data:

- OP samples for which <3 assay endpoints were positive (malaoxon, acephate, diazoxon, dicrotophos, fosthiazate, dichlorvos, profenofos, omethoate);
- OP chemicals with two separate samples with discordant patterns (chlorpyrifos oxon, diazinon, methamidophos); and,
- the largest group, OP samples for which general activity, bursting, and network connectivity were all decreased (bensulide, chlorethoxyfos, chlorpyrifos, coumaphos, dimethoate, ethoprop, malathion, naled, phorate, phosmet, pirimphos-methyl, tebupirimfos, terbufos, tribufos, trichlorfon, z-tetrachlorvinphos).

In previous work with the MEA for acute studies, it has been suggested that “true positive” responses may only occur for samples that elicit effects for 3 or more assay endpoints (Kosnik et al., 2020), though a statistical analysis of the possible Type I error rate across the assay suite has not been performed to date for the acute MEA or the MEA NFA. Tentatively, the samples with less than 3 assay endpoints positive may be considered equivocal, whereas samples with 0 assay endpoints positive may be considered negative.

For the three OP chemicals that have samples with discordant patterns of response, it is not possible to explain fully the results and re-screening would be required for a definitive result. However, eight OP chemicals were tested with two separate samples, and overall these chemicals reflected the hit-call concordance of the MEA NFA of approximately 80% between replicate samples (Supplemental Appendix B). For four of eight chemicals with replicate samples, results were concordant, with >3 assay endpoints being altered for chlorpyrifos, dimethoate, malathion and trichlorfon, resulting in concordant positive results. One sample of acephate was negative in all assay endpoints, and one sample of acephate was positive for only two assay endpoints, indicating that this chemical was possibly in the equivocal category. In contrast, results with replicate samples of chlorpyrifos oxon, diazinon and, to a lesser extent, methamidophos, were discordant. Many factors may have affected these results, including the experiments being performed in different years by different laboratory technicians, but one hypothesis is that sample stability in DMSO may have contributed to some of the discordant results for the OPs.

**Table 10. OP chemicals screened in at least two screening blocks in the MEA NFA.**

The DTXSID, chemical, sample, number of positive assay endpoints ((+) AEIDs), the minimum log<sub>10</sub>-AC<sub>50</sub> (μM), the mean log<sub>10</sub>-AC<sub>50</sub> (μM), the standard deviation of the log<sub>10</sub>-AC<sub>50</sub> values for both samples of the chemical across all assay endpoints, and the response group are presented. Response group indicates the activity as equivocal (0-2 assay endpoint positive for each sample); positive (3 or more positive assay endpoints for each sample); and, mixed (discordant results between samples).

DTXSID	Chemical	Sample	(+) AEIDs	Min log <sub>10</sub> -AC <sub>50</sub>	Mean log <sub>10</sub> -AC <sub>50</sub>	SD (log <sub>10</sub> -AC <sub>50</sub> ), average	Response group
DTXSID8023846	Acephate	EPAPLT0167A01	2	-0.042	0.968	NA	Equivocal
DTXSID8023846	Acephate	TT0000177A04	0	NA	NA	NA	
DTXSID4020458	Chlorpyrifos	EX000384	19	0.197	1.13	0.151	Positive
DTXSID4020458	Chlorpyrifos	TT0000177E02	18	1.1	1.22	0.151	
DTXSID1038666	Chlorpyrifos oxon	EX000378	16	-0.84	0.596	NA	Mixed
DTXSID1038666	Chlorpyrifos oxon	TT0000177G02	0	NA	NA	NA	
DTXSID9020407	Diazinon	EPAPLT0170D06	2	1.08	1.23	NA	Mixed
DTXSID9020407	Diazinon	TT0000177H01	17	1.51	1.72	NA	
DTXSID7020479	Dimethoate	EPAPLT0167G06	15	1.06	1.27	0.3	Positive
DTXSID7020479	Dimethoate	TT0000177H02	18	1.56	1.7	0.3	
DTXSID4020791	Malathion	EPAPLT0167G08	18	-0.0954	0.798	0.254	Positive
DTXSID4020791	Malathion	TT0000177D02	19	0.864	1.16	0.254	
DTXSID6024177	Methamidophos	EPAPLT0167A08	5	-0.999	0.664	NA	Mixed
DTXSID6024177	Methamidophos	TT0000177B02	0	NA	NA	NA	
DTXSID0021389	Trichlorfon	EPAPLT0170D03	12	0.281	0.85	0.443	Positive
DTXSID0021389	Trichlorfon	TT0000177F01	18	0.548	1.56	0.443	

Further information on sample stability is indicated here:

- Differential activity was observed between two different samples of chlorpyrifos-oxon; one sample (TT0000177G02) demonstrated no effects, whereas the other sample (EX000378) demonstrated suppression of general activity, bursting, and network connectivity. For chlorpyrifos oxon, quality control data on the particular sample used in screening was not available whereas purity issues of separately sourced Tox21 samples were detected both at the time of shipment (75-90% purity) and after 4 months of storage at room temperature (<50% purity) ([https://tripod.nih.gov/tox21/samples/Tox21\\_301063](https://tripod.nih.gov/tox21/samples/Tox21_301063)). Thus, the inactive chlorpyrifos oxon sample may be inactive due to sample degradation, but additional quality control test of the chemical sample or additional screening with a new analytically verified chlorpyrifos oxon sample would likely be needed to resolve the activities. For the work herein, the active chlorpyrifos oxon sample bioactivity data have been used for derivation of AEDs in Section 2.3.6.
- One sample of diazinon (EPAPLT0170D06) altered only three assay endpoints, whereas the other sample (TT0000177H01) decreases numerous endpoints related to general activity, bursting and connectivity. Analysis of related samples of diazinon suggest that diazinon in some samples may

degrade over time, with high purity confirmed at sample shipment but with purity < 50% at 4 months of storage at room temperature ([https://tripod.nih.gov/tox21/samples/Tox21\\_300730](https://tripod.nih.gov/tox21/samples/Tox21_300730)).

- One sample of methamidophos (TT0000177B02) was without activity towards any endpoint, whereas the other sample (EPAPLT0167A08) decreased a subset of 5 endpoints. Quality control data on related and unrelated samples of methamidophos indicate stability in DMSO solution over time. The positive responses from one sample (EPAPLT0167A08) of methamidophos appear to be the result of relatively borderline efficacy.

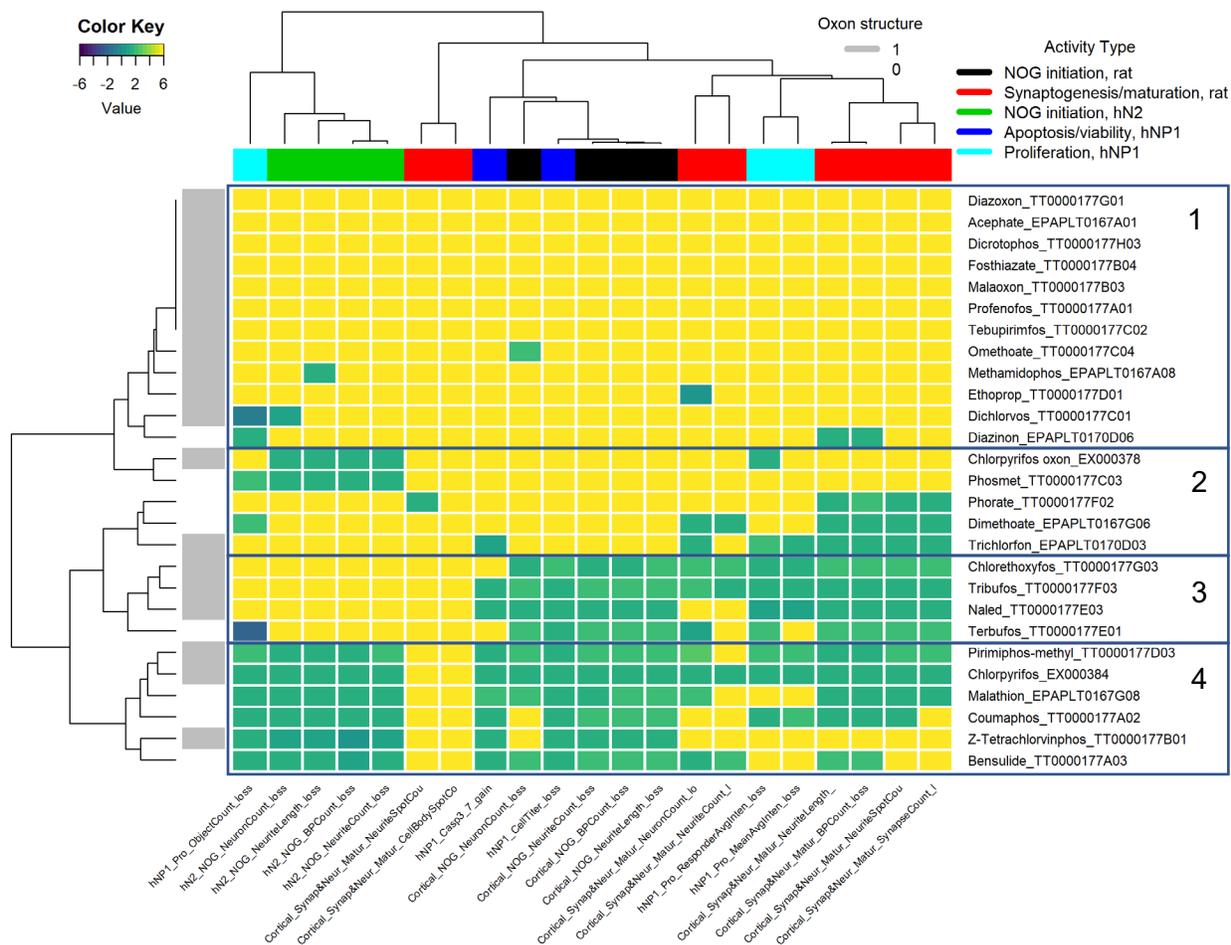
Considering the heatmap of the MEA NFA results, it is unclear if the oxon structural moiety drives differential activity due to the small number of chemicals included in this dataset. Most of the 27 OP chemicals in this dataset contained the oxon structural moiety (19/27, as indicated by the gray row annotation in Figure 3), and these oxon-containing OP chemicals distribute between the main clusters (positive, equivocal, and negative). Of the non-oxon OP chemicals (bensulide, coumaphos, diazinon, dimethoate, malathion, phorate, phosmet, and terbufos), all except for one sample of diazinon appear in the positive cluster. Considering the bioactivity of parent/metabolite pairs may explain some of the differential activity observed: (1) malathion was positive and its metabolite malaaxon was negative; (2) diazinon demonstrated discordant activity across samples (one positive and one equivocal) and its metabolite diazoxon was equivocal; (3) dimethoate was positive and its metabolite omethoate was equivocal; and, (4) chlorpyrifos was positive and its metabolite chlorpyrifos oxon demonstrated discordant activity (positive and negative) across samples.

In terms of potency, the log<sub>10</sub>-AC<sub>50</sub> ranges across the active OP samples ranged from approximately 0 to 2 (1 to 100 μM), with the exception of dichlorvos and terbufos, which demonstrated some log<sub>10</sub>-AC<sub>50</sub> values less than 0 (in the 0.01-1 μM range) in one of the apoptosis-related assay endpoints. In contrast to the MEA NFA, the assays in the HCI technology demonstrated some differential activity across activity types represented by different assay subsets. For the HCI assays, approximately four groups are apparent from hierarchical clustering of the data:

- Cluster 1: OP samples with no effects or disparate patterns of response in one to three assay endpoints (no effects with acephate, diazoxon, dicrotophos, fosthiazate, malaaxon, profenofos, or tebupirimifos, and disparate effects with diazinon, dichlorvos, ethoprop, omethoate, and methamidophos);
- Cluster 2: OP samples with effects on five or more assay endpoints, with chlorpyrifos oxon and phosmet affecting NOG initiation in rat cortical cells and dimethoate, phorate, and trichlorfon appearing to affect synaptogenesis/neurite maturation in rat cortical cells;
- Cluster 3: OP samples with effects on all HCI assay activity types except for NOG initiation in hN2 cells (chlorethoxyfos, naled, terbufos, tribufos); and,
- Cluster 4: OP samples with more widespread effects across activity types, with chlorpyrifos, coumaphos, malathion, and pirimiphos-methyl affecting all activity types (rat and hN2 NOG; synaptogenesis and neurite maturation in rat; proliferation, apoptosis and viability in hNP1 cells). Bensulide affected all activity types without affecting synaptogenesis and neurite maturation, and Z-tetrachlorvinphos affected all activity types but with more limited effects on synaptogenesis and neurite maturation.

**Figure 6. Heatmap of HCI assay results for OPs.**

The OP substance results are presented using the log10-AC50 (µM) values used in the main heatmap image, with columns indicating assay endpoint names and rows indicating OP substance name and sample id (spid), resulting in 27 rows for 27 unique substances (no repeated substances). Darker blue corresponds to lower log10-AC50 values, and a value of 6 (represented by yellow) has been assigned for substances with a negative hit-call for the assay endpoint. The activity type (NOG initiation, rat = black; synaptogenesis/maturation, rat = red; NOG initiation, hN2 = green; apoptosis/viability, hNP1 = dark blue; proliferation, hNP1 = light blue) for each assay endpoint is also annotated for the columns. For row annotations, OP substances that have an oxon structural moiety have been annotated by row (gray). The apparent clusters are numbered for further discussion below.



All of the OPs in the first cluster (no effects or effects in one to three assay endpoints) were oxon-containing OP chemicals, though oxons distributed throughout all of the clusters. The OP chemicals that do not contain an oxon structural moiety were distributed among the three more active clusters. Considering parent-metabolite pairs, as with the MEA NFA results, generally indicate that the parent OP chemical was more active than the oxon metabolite: (1) diazinon demonstrated disparate effects whereas its metabolite diazoxon was negative; (2) malathion demonstrated widespread effects on all activity types whereas its metabolite malaoxon was negative; (3) dimethoate affected synaptogenesis/neurite maturation but omethoate affected only one assay endpoint related to rat NOG; and, (4) chlorpyrifos demonstrated

widespread effects across all activity types whereas chlorpyrifos oxon affected rat NOG initiation only. Though there are only 4 parent-metabolite pairs, the trend of more limited DNT-NAM activity for metabolites in parent-metabolites pairs appears consistent.

Overall, the range of log<sub>10</sub>-AC<sub>50</sub> values for active assay endpoints was similar for the MEA NFA and the HCI assays (0-2 log<sub>10</sub>-μM). Further, negative and equivocal results in MEA NFA (0-2 assay endpoints positive) generally corresponded to Cluster 1 for the HCI assays (0-3 assay endpoints positive). There is only 1 OP chemical (ethoprop) that was positive in MEA NFA and in the HCI cluster 1, indicating that negatives tend to be consistent between MEA NFA and HCI assays. Acephate, diazoxon, dichlorvos, dicrotophos, malaaxon, omethoate, and profenofos demonstrated limited to no activity in the MEA NFA and the HCI assays (Table 11). Considering all chemicals in HCI activity clusters 2, 3, and 4, only 1 OP chemical (methamidophos) was not positive in the MEA NFA, indicating that if activity is observed in the HCI assays, it is likely that the OP chemical will also be active in the MEA NFA.

**Table 11. Comparison of heatmap clustering.**

The cluster for the MEA NFA (negative = Neg; equivocal (0-2 assay endpoints positive) = Equiv; positive (3 or more assay endpoints positive) = Pos) is indicated alongside the HCI assay cluster (1 = limited to no activity; 2 = activity in 5 or more assay endpoints; 3 = activity in all assay? types except NOG initiation in hN2 cells; 4 = widespread activity).

DTXSID	Chemical	MEA NFA			HCI			
		Neg	Equiv	Pos	1	2	3	4
DTXSID8023846	Acephate	X	X		X			
DTXSID9032329	Bensulide			X				X
DTXSID2032344	Chlorethoxyfos			X			X	
DTXSID4020458	Chlorpyrifos			X,X				X
DTXSID1038666	Chlorpyrifos oxon	X		X		X		
DTXSID2020347	Coumaphos			X				X
DTXSID9020407	Diazinon		X	X		X		
DTXSID5037523	Diazoxon		X		X			
DTXSID5020449	Dichlorvos		X		X			
DTXSID9023914	Dicrotophos		X		X			
DTXSID7020479	Dimethoate			X		X		
DTXSID4032611	Ethoprop			X	X			
DTXSID0034930	Fosthiazate		X		X			
DTXSID9020790	Malaoxon	X			X			
DTXSID4020791	Malathion			X				X
DTXSID6024177	Methamidophos	X	X			X		
DTXSID1024209	Naled			X			X	
DTXSID4037580	Omethoate		X		X			
DTXSID4032459	Phorate			X		X		
DTXSID5024261	Phosmet			X		X		
DTXSID0024266	Pirimiphos-methyl			X				X
DTXSID3032464	Profenofos		X		X			
DTXSID1032482	Tebupirimfos			X	X			
DTXSID2022254	Terbufos			X			X	
DTXSID1024174	Tribufos			X			X	
DTXSID0021389	Trichlorfon			X		X		
DTXSID1032648	Z-Tetrachlorvinphos			X				X

### 2.3.5.3 Further consideration of potency, selectivity, and relative bioactivity

It is currently unknown whether cytotoxicity in neural cell types is itself relevant to putative DNT-related activity, or if this cytotoxicity can be considered off-target activity. Thus, for the purposes of comparing potency in the DNT-NAM assays to BMDs from *in vivo* studies (Section 2.3.6), hits in the cytotoxicity and apoptosis assays have been included. However, it is of interest to understand if effects on neural network formation and function in the MEA NFA, or effects on NOG initiation, synaptogenesis and neurite maturation, and/or proliferation occur at concentrations lower than the concentrations that appear to affect apoptosis and/or cell viability. A comparison of the DNT-NAM minimum potency to the potency values observed in the broader ToxCast/Tox21 screening program, and the cytotoxicity burst threshold potency, is provided in Supplemental Appendix C.

Selectivity results for the MEA NFA were mixed across the set of OP substances examined. Bensulide, chlorethoxyfos, diazoxon, dimethoate, malathion, methamidophos, and naled appeared to demonstrate some selectivity (selectivity score > 0.3). Acephate, chlorpyrifos and chlorpyrifos oxon, and methamidophos appeared to demonstrate selectivity depending on the chemical sample screened in the MEA NFA. Some OP substances demonstrated activity in HCI assays that appears selective (selectivity > 0.3), but this observation was dependent on activity type/assay subset. Several chemicals (coumaphos, malathion, naled, phorate, pirimphos-methyl, and tribufos) appeared to demonstrate selectivity in the synaptogenesis and neurite maturation assay endpoints with rat primary cortical neurons. Bensulide and methamidophos were the only chemicals that appeared to demonstrate selectivity for hN2 NOG. Chlorethoxyfos, chlorpyrifos oxon, naled, and tribufos appeared to demonstrate some selectivity for the hNP1 proliferation assays. Selectivity varied considerably by chemical and by the particular assay/activity type.

The scaled AUC sum was calculated to indicate the relative *in vitro* bioactivity of the OP chemicals in this dataset by compressing potency and efficacy into a single value by OP chemical. The scaled AUC sum results suggest that several OPs act with similar magnitude of effect in the MEA NFA (scaled AUC sum > 0.9). Scaled AUC sums < 0.8 appeared to capture OP substances with smaller magnitudes of effect in the MEA NFA (acephate, diazoxon, dichlorvos, dicotophos, fosthiazate, methamidophos, omethoate, profenofos, Z-Tetrachlorvinphos). Indeed, these substances with scaled AUC sum values < 0.8 in the MEA NFA tended to reflect the OP chemicals with equivocal responses. Scaled AUC sum values in the HCI assays of < 0.8 appeared to generally indicate substances that were less active in the HCI assays overall, whereas scaled AUC sum > 0.8 generally indicated the more active substances (e.g., more assay endpoints positive) in the HCI assays. In future applications, a scaled AUC sum or similar representation of relative *in vitro* bioactivity could be used to indicate chemicals with effects similar or dissimilar to known DNT reference chemicals, as potency comparisons alone may not reveal differences in efficacy.

**Table 12. OP potency, selectivity, and magnitude of effect in the MEA NFA.**

The sample id (spid), DTXSID, chemical, hit-call sum (number of positive assay endpoints in the MEA NFA out of 36 total), minimum log<sub>10</sub>-AC50 (μM) for all positive assay endpoints in the MEA NFA, the minimum cytotoxicity log<sub>10</sub>-AC50 (μM) from the MEA NFA LDH and AB assay endpoints, the selectivity score (cytotox – min AC50, in log<sub>10</sub>-μM), the minimum MEA NFA AC50 value in μM units, the median MEA NFA AC50 value in μM units, and the scaled AUC sum for the positive MEA NFA assay endpoints are all provided. When selectivity > 0.3, this indicates that MEA NFA activity was disrupted at concentrations lower than concentrations associated with cytotoxicity.

spid	DTXSID	Chemical	Hitcall sum	Min log <sub>10</sub> -	Cytotox log <sub>10</sub> -AC50 (μM)	Selectivity	Min AC50 (μM)	Med AC50 (μM)	Scaled AUC sum
EPAPLT0167A01	DTXSID8023846	Acephate	2	-0.04	1.48	1.52	0.91	9.28	0.73
TT0000177A04	DTXSID8023846	Acephate	0	NA	2.00	NA	NA	NA	NA
TT0000177A03	DTXSID9032329	Bensulide	20	1.07	1.53	0.47	11.69	31.35	0.96
TT0000177G03	DTXSID2032344	Chlorethoxyfos	18	1.14	1.49	0.35	13.78	40.1	0.91
EX000384	DTXSID4020458	Chlorpyrifos	19	0.20	1.20	1	1.57	14.38	0.92
TT0000177E02	DTXSID4020458	Chlorpyrifos	18	1.10	1.14	0.04	12.68	17.84	0.99
EX000378	DTXSID1038666	Chlorpyrifos oxon	16	-0.84	-0.50	0.34	0.14	7.66	0.99
TT0000177G02	DTXSID1038666	Chlorpyrifos oxon	0	NA	2.00	NA	NA	NA	NA
TT0000177A02	DTXSID2020347	Coumaphos	18	1.39	1.67	0.28	24.61	45.72	0.91
EPAPLT0170D06	DTXSID9020407	Diazinon	2	1.08	1.48	0.4	12.07	17.05	0.54
TT0000177H01	DTXSID9020407	Diazinon	17	1.51	1.55	0.03	32.68	52.06	0.87
TT0000177G01	DTXSID5037523	Diazoxon	1	-0.06	2.00	2.06	0.88	0.88	0.5
TT0000177C01	DTXSID5020449	Dichlorvos	1	0.67	0.67	0	4.63	4.63	0.42
TT0000177H03	DTXSID9023914	Dicrotophos	1	1.96	1.96	0	91.72	91.72	0.11
EPAPLT0167G06	DTXSID7020479	Dimethoate	15	1.06	1.42	0.36	11.54	20.02	0.8
TT0000177H02	DTXSID7020479	Dimethoate	18	1.56	1.60	0.04	35.94	50.49	0.88
TT0000177D01	DTXSID4032611	Ethoprop	16	1.08	1.13	0.05	11.89	26.57	0.95
TT0000177B04	DTXSID0034930	Fosthiazate	1	1.76	1.76	0	57.05	57.05	0.39
TT0000177B03	DTXSID9020790	Malaoxon	0	NA	2.00	NA	NA	NA	NA
EPAPLT0167G08	DTXSID4020791	Malathion	18	-0.10	1.40	1.49	0.8	8.01	0.99

spid	DTXSID	Chemical	Hitcall sum	Min log10-	Cytotox log10-AC50 (µM)	Selectivity	Min AC50 (µM)	Med AC50 (µM)	Scaled AUC sum
TT0000177D02	DTXSID4020791	Malathion	19	0.86	1.30	0.44	7.31	12.5	1
EPAPLT0167A08	DTXSID6024177	Methamidophos	5	-1.00	1.46	2.46	0.1	25.11	0.77
TT0000177B02	DTXSID6024177	Methamidophos	0	NA	2.00	NA	NA	NA	NA
TT0000177E03	DTXSID1024209	Naled	19	0.49	0.82	0.34	3.06	7.5	1.05
TT0000177C04	DTXSID4037580	Omethoate	3	1.81	1.93	0.11	65.31	84.61	0.53
TT0000177F02	DTXSID4032459	Phorate	18	1.50	1.54	0.04	31.57	64.11	0.84
TT0000177C03	DTXSID5024261	Phosmet	19	0.90	1.50	0.6	8.01	33.59	0.94
TT0000177D03	DTXSID0024266	Pirimiphos-methyl	19	1.34	1.46	0.12	21.75	32.3	0.92
TT0000177A01	DTXSID3032464	Profenofos	2	0.15	0.15	0	1.41	2.88	0.6
TT0000177C02	DTXSID1032482	Tebupirimfos	19	1.64	1.81	0.17	44.07	81.17	0.8
TT0000177E01	DTXSID2022254	Terbufos	19	1.10	1.18	0.08	12.73	18.88	0.99
TT0000177F03	DTXSID1024174	Tribufos	18	1.44	1.44	0	27.32	34.16	0.91
EPAPLT0170D03	DTXSID0021389	Trichlorfon	12	0.28	0.28	0	1.91	10.13	0.91
TT0000177F01	DTXSID0021389	Trichlorfon	18	0.55	0.55	0	3.53	42.9	0.94
TT0000177B01	DTXSID1032648	Z-Tetrachlorvinphos	12	1.74	1.74	0	55.5	92.05	0.67

**Table 13. OP potency, selectivity, and magnitude of effect in the HCI assays.**

The sample id (spid), DTXSID, chemical, hit-call sum (number of positive assay endpoints in the HCI assays out of 21 total), the selectivity score (cytotox – min AC50, in log10- $\mu$ M) for hN2 NOG, rat NOG, rat neurite maturation in the synaptogenesis assay, rat synaptogenesis, and hNPI proliferation, the minimum HCI AC50 value in  $\mu$ M units, the median HCI AC50 value in  $\mu$ M units, and the scaled AUC sum for the positive HCI assay endpoints are all provided. When selectivity > 0.3, this indicates that HCI activity was disrupted at concentrations lower than concentrations associated with cytotoxicity.

Spid	DTXSID	Chemical	Hitcall sum	hN2 NOG selectivity	Rat NOG selectivity	Rat Neur. Matur. selectivity	Rat Synap. selectivity	hNPI Prolif. selectivity	Min AC50 ( $\mu$ M)	Med AC50 ( $\mu$ M)	Scaled AUC sum
EX000384	DTXSID4020458	Chlorpyrifos	19	-0.12	0.1	0.07	0.12	0.08	18.26	29.97	1
TT0000177B01	DTXSID1032648	Z-Tetrachlorvinphos	10	0.16	0.5	NA	NA	0	3.85	18.18	1
TT0000177E03	DTXSID1024209	Naled	12	NA	0.02	0.48	0.7	1.07	8.42	29.79	0.95
TT0000177A03	DTXSID9032329	Bensulide	15	0.34	0.04	-0.03	-0.14	0	11.57	60.26	0.92
TT0000177A02	DTXSID2020347	Coumaphos	15	0	0.08	0.41	0.5	0.16	21.75	38.76	0.91
EPAPLT0167G08	DTXSID4020791	Malathion	16	-0.02	0.09	0.29	0.32	0	32.57	40.13	0.9
TT0000177D03	DTXSID0024266	Pirimiphos-methyl	18	-0.05	0.04	0.85	0.93	0	33.88	84.46	0.87
TT0000177E01	DTXSID2022254	Terbufos	12	NA	-0.02	-1.31	-1.3	0	0.01	83.52	0.87
TT0000177F03	DTXSID1024174	Tribufos	14	NA	-0.01	0.43	0.45	0.5	31.53	39.49	0.86
EPAPLT0170D03	DTXSID0021389	Trichlorfon	8	NA	NA	-0.15	-0.01	0.24	14.81	53.08	0.85
TT0000177G03	DTXSID2032344	Chlorethoxyfos	13	NA	0.04	0.06	0.11	0.41	38.59	65.52	0.84
EX000378	DTXSID1038666	Chlorpyrifos oxon	5	-0.09	NA	NA	NA	0.32	28.79	43.29	0.79
EPAPLT0167G06	DTXSID7020479	Dimethoate	7	NA	NA	0.23	0.1	0	35.37	51.33	0.79
TT0000177C03	DTXSID5024261	Phosmet	5	-0.03	NA	NA	NA	0	29.12	32.98	0.79
TT0000177C01	DTXSID5020449	Dichlorvos	2	NA	NA	NA	NA	0	0.11	0.84	0.7
TT0000177F02	DTXSID4032459	Phorate	5	NA	NA	0.25	0.46	NA	34.69	35.14	0.7
EPAPLT0170D06	DTXSID9020407	Diazinon	3	NA	NA	0.23	0.2	0	53.38	58.97	0.67
TT0000177D01	DTXSID4032611	Ethoprop	1	NA	NA	NA	NA	NA	3.84	3.84	0.67
TT0000177C04	DTXSID4037580	Omethoate	1	NA	NA	NA	NA	NA	118.19	118.19	0.43

Spid	DTXSID	Chemical	Hitcall sum	hN2 NOG selectivity	Rat NOG selectivity	Rat Neur. Matur. selectivity	Rat Synap. selectivity	hNP1 Prolif. selectivity	Min AC50 (µM)	Med AC50 (µM)	Scaled AUC sum
EPAPLT0167A08	DTXSID6024177	Methamidophos	1	0.35	NA	NA	NA	NA	44.42	44.42	0.42
EPAPLT0167A01	DTXSID8023846	Acephate	0	NA	NA	NA	NA	NA	NA	NA	NA
TT0000177G01	DTXSID5037523	Diazoxon	0	NA	NA	NA	NA	NA	NA	NA	NA
TT0000177H03	DTXSID9023914	Dicrotophos	0	NA	NA	NA	NA	NA	NA	NA	NA
TT0000177B04	DTXSID0034930	Fosthiazate	0	NA	NA	NA	NA	NA	NA	NA	NA
TT0000177B03	DTXSID9020790	Malaoxon	0	NA	NA	NA	NA	NA	NA	NA	NA
TT0000177A01	DTXSID3032464	Profenofos	0	NA	NA	NA	NA	NA	NA	NA	NA
TT0000177C02	DTXSID1032482	Tebupirimfos	0	NA	NA	NA	NA	NA	NA	NA	NA

### 2.3.6 DNT-NAM administered equivalent doses versus rat benchmark doses

The objective of this section is to demonstrate the comparison between the results of transforming the *in vitro* bioactive concentrations ( $\mu\text{M}$ , that is,  $\mu\text{mol/L}$ ) from the current DNT-NAM battery into AEDs in units of  $\text{mg/kg/day}$  and BMD and BMDL values also in units of  $\text{mg/kg/day}$  from *in vivo* studies of AChE inhibition in rats. This is a proof-of-concept application of the DNT-NAM data to a targeted question within a regulatory context for a set of OPs. The purpose of this proof-of-concept is to evaluate whether currently available DNT-NAM data suggest possible bioactivity of OPs at similar, higher, or lower doses to those observed in *in vivo* AChE studies. It is important to note that the DNT-NAM battery may or may not include AChE inhibition as a MIE, as the AChE expression has not been specifically evaluated in the MEA NFA or HCI key event assays. However, these DNT-NAMs do evaluate neural network formation and functional processes critical to neurodevelopment that could be disrupted by neuroactive xenobiotics. As such, the DNT-NAM battery may provide relevant information on DNT potential, and combined with IVIVE approaches, may provide information for determining at what predicted doses putative DNT effects would occur.

In the sections below, the IVIVE approach is briefly summarized for using high-throughput toxicokinetic (HTTK) data and models to compute AEDs based on the DNT-NAMs. As the general HTTK approach has been previously published (Bell et al., 2018; Pearce et al., 2017a; Pearce et al., 2017b; Rotroff et al., 2010; Wambaugh et al., 2018; Wambaugh et al., 2015; Wetmore, 2015; Wetmore et al., 2012) and reviewed by FIFRA SAP in July 2014 (Jenkins, 2014), only the methodological details pertinent to this Issue Paper analysis are presented. More details, including the data and models available in HTTK for the OPs in this proof-of-concept, are available in Supplemental Appendix D.

#### 2.3.6.1 General assumptions for the IVIVE approach

The reverse dosimetry component of IVIVE in case study relies on a few high-level assumptions:

- (1) that a bioactive nominal *in vitro* assay concentration approximates an *in vivo* plasma concentration that would correspond to a similar effect;
- (2) that *in vivo* plasma concentration can be approximated based on steady-state kinetics; and,
- (3) that a toxicokinetic model to estimate the external exposures (in  $\text{mg/kg/day}$  units) that may have resulted in that plasma concentration can be constructed using estimates of species-specific physiology and Phase I and Phase II enzyme-driven hepatic clearance.

Additionally, there was another uncertainty introduced specific to this work: in the absence of hepatic clearance values from rat hepatocytes, rat liver microsomes, or rat liver Phase I enzymes, would the use of human hepatocyte-derived hepatic clearance values be a reasonable substitute? Thus, the impact of using human-derived HTTK information as a data-gap filling technique when rat-derived HTTK information was unavailable was evaluated in the context of HTTK modeling, with the details of this analysis provided in Supplemental Appendix D. Resultant to this work, in addition to comparing rat-derived  $\text{AED}_{50}$  values to BMD10 and BMDL10 values from rat studies, we also compared AED values from the “humanized-rat” or the huRat, which used human HTTK data in a model parameterized with rat physiology, to BMD10 and BMDL10 values from rat studies.

#### 2.3.6.2 HTTK approach selected

For this proof-of-concept, many HTTK model and data choices could have been selected; among these choices, the simplest HTTK model requiring the least amount of data was selected: the 3-compartment steady state model (3comps), which enables calculation of the dose that would cause an *in vivo* plasma

concentration equivalent to the bioactive *in vitro* AC50 if hepatic intrinsic clearance (Clint, in units of  $\mu\text{L}/\text{min}/10^6$  hepatocytes) data are available. This model was selected so that in future cases, the same approach could be taken with the least amount of data available, with the possibility of utilizing the HHTK physiologically-based toxicokinetic (PBTK) modeling if both Clint and fraction unbound in plasma (Fup)<sup>9</sup> were available. LogP and pKa, used in estimation of tissue partitioning coefficients required for PBTK modeling, can be predicted and so are unlikely to be the limiting factor in model selection. Beyond the options for IVIVE using an HHTK approach, it is possible that other PBPK models could also be actualized to generate plasma concentration predictions, e.g., steady state plasma concentration (C<sub>ss</sub>), when available.

Most HHTK models, including the 3compss model used here, share several simplifying technical assumptions:

- (1) 100% bioavailability (all of an oral dose is received by the liver through the portal vein);
- (2) No extrahepatic metabolism; the liver is the only source of chemical clearance from the body by metabolism;
- (3) Hepatic metabolism is first order (proportional to concentration) and does not saturate;
- (4) Renal clearance is proportional to Fup and glomerular filtration rate (i.e., no active transport); and,
- (5) No biliary excretion or enterohepatic recirculation occurs.

With these assumptions, HHTK models have demonstrated reasonable accuracy in predicting relevant TK endpoints, for example plasma concentrations over time (AUC) ( $R^2 = 0.62$ ) and maximum plasma concentrations (C<sub>max</sub>) ( $R^2 = 0.48$ ) (Wambaugh et al., 2018).

AED values in mg/kg/day units were calculated using the following equation (Eq. 2):

$$\text{Eq. 2: } AED_{50} \left( \frac{\text{mg}}{\text{kg}} \right) = AC50 (\mu\text{M}) * \frac{1 \frac{\text{mg}}{\text{kg}} / \text{day}}{C_{SS50}}$$

Where the C<sub>ss</sub> values for the median individual based on Monte Carlo simulation of species-specific physiological parameters (C<sub>SS50</sub>) (Pearce et al. 2017) were generated using the 3compss model. Values for C<sub>SS50</sub> were generated in a species-specific manner to result in 3 different kinds of AED<sub>50</sub> values: using rat physiology and empirical rat Clint data where available (Rat, C<sub>SS50</sub>-AED), rat physiology and empirical human Clint data where available (huRat, C<sub>SS50</sub>-AED), and then human physiology with empirical human Clint data (Human, C<sub>SS50</sub>-AED). The huRat was used as an additional comparison because human HHTK data, specifically human Clint data for use in the 3 compss model, were available for more of the OPs in this proof-of-concept than the number for which rat Clint data were available. The use of human HHTK data for Clint and Fup in the prediction of C<sub>max</sub> and plasma concentrations over time (area under the curve, AUC) as a data gap-filling technique when rat HHTK data were not available is supported by preliminary results of evaluating the overall impact of using rat or human HHTK data in a PBTK model for 151 substances that have both rat and human HHTK data (Supplemental Appendix D, Table 2 and Figure 3). This analysis in Supplemental Appendix D suggests that plasma C<sub>max</sub> values obtained from the rat PBTK model, using either rat or human HHTK data for Fup and Clint, result in values that are similar (generally within  $\pm 0.5 \log_{10}$ - $\mu\text{M}$ ). The plasma AUC values that result from using rat or human

<sup>9</sup> For the 3 compss model, Fup is used, if available, for estimation of the blood::plasma ratio. If Fup is not available, the 3 compartment steady state model estimates the blood::plasma ratio using the average Fup across the entire dataset within the HHTK R package.

HTTK data in a rat PBTK model generally were within  $\pm 1 \log_{10}$ - $\mu\text{M}$ . This comparison does not suggest that a standard conversion factor is necessary when using human HTTK data for Fup and Clint as a substitute for rat Fup and Clint HTTK data, as the plasma predictions are not uniformly higher or lower when using human rather than rat HTTK in the rat PBTK model.

The availability of HTTK data for the 27 OP substances is provided in Supplemental Appendix D, Table 4. Reliable empirical measurements of human (and rat) Clint for chlorethoxyfos, naled, and Z-tetrachlorvinphos were not possible due to chemical instability and degradation in plasma and hepatocyte media (pH 7.4) within 4-7 hrs of solubilization (unpublished, Wetmore 2020). Due to the instability of these specific OP chemicals in solution, these OP chemicals were excluded from further IVIVE comparison in this Issue Paper.

### **2.3.6.3 AED to BMD/BMDL comparison**

All of the DNT-NAM AC50 ( $\mu\text{M}$ ) potency data from the MEA NFA and HCI assays were included for use in these comparisons (after filtering for less reproducible curve-fits as indicated in Section 2.3.5.1). However, to enable species-specific comparisons, the data from the DNT-NAM battery were parsed according to the species from which the cell type used in the assay originated. The MEA NFA contains 36 assay endpoints (17 assay components associated with neuronal network formation and function, analyzed in up and down directions to yield 34 assay endpoints, plus two assay endpoints for cytotoxicity) that were all assayed using rat primary cortical neurons as described in Section 2.3.3.1. The HCI assays are comprised of 21 assay endpoints that encompass 5 related assays: cell proliferation with a human neural progenitor cell line (hNP1, 3 assay endpoints); apoptosis and viability assessment using a human neural progenitor cell line (hNP1, 2 assay endpoints); neurite outgrowth initiation with a human neuronal lineage cell line (hN2, 4 assay endpoints); primary rat cortical cell neurite outgrowth initiation (rat cortical cell, 4 assay endpoints); and, primary rat cortical cell neurite maturation and synaptogenesis (rat cortical cells, 8 assay endpoints).

Thus, 9 of 57 total assay endpoints in the DNT-NAM battery that evaluate neural cell proliferation, apoptosis and viability, and neurite outgrowth initiation were evaluated using human-derived cells. These data were used in a “human” comparison, whereby *in vitro* potency values from human-derived cells, human-derived HTTK data and a human-parameterized 3compss model, and BMD10/BMDL10 values from rat divided by an uncertainty factor of 10 (default interspecies uncertainty factor applied in risk assessments to account for extrapolation from laboratory animals to humans) were compared. Only 17 OP chemicals had positive values in the human assays, and of these 14 had sufficient HTTK data and modeling to calculate human AED<sub>50</sub> values with the 3 compss model.

Forty-eight of the 57 total assay endpoints in the DNT-NAM battery that evaluate neuronal network formation and function, neurite outgrowth initiation, and neurite maturation and synaptogenesis were evaluated using rat primary cortical cells. These data were used in rat and huRat comparisons, whereby *in vitro* potency values from primary rat cells, combined with rat-derived or human derived Clint data and a rat-parameterized 3 compss model, and BMD10/BMDL10 values from rat were compared. Three OP chemicals (chlorethoxyfos, naled, Z-tetrachlorvinphos) were excluded from the IVIVE comparisons due to chemical instability in the matrices used for HTTK assays (unpublished, Wetmore 2020). Malaoxon was completely negative in all assay endpoints. Thus, 23/27 OP chemicals have enough data and modeling available to derive huRat AED<sub>50</sub> values. Only 9 OP chemicals had enough data and modeling available to derive rat AED<sub>50</sub> values.

Each AC50 value from positive assay endpoints was transformed into an AED<sub>50</sub> value using the IVIVE approach described in section 2.3.6.2. The number of positive assay endpoints by species, the minimum

and median AC50 values by species, are summarized by OP chemical in Table 14. Boxplots of the AED<sub>50</sub> values versus the BMD10/BMDL10 values, by species, are illustrated in Figure 7. Substances with zero positive human (+ human) assay endpoints in Table 14 will not have any AEDs computed or displayed for the human comparisons in Figure 7 (acephate, diazoxon, dicrotophos, ethoprop, fosthiazate, omethoate, phorate, profenofos, tebupirimfos). As stated above, malaoxon was completely negative in the DNT-NAM battery and as such is also excluded from the AED comparisons visualized in Figure 7. In order to indicate the influence of measures of cell viability, selectivity was calculated for each AC50 value used to calculate an AED, per the logic described in Section 2.3.5.1. Briefly, selectivity required a 0.3 log<sub>10</sub>-micromolar separation from the parallel cell viability AC50 value, and selectivity could not be calculated (shown as NA) for the apoptosis and cell viability assays conducted in the HCI assays using hNP1 cells.

**Table 14. DNT-NAM data availability for AED<sub>50</sub> computation.**

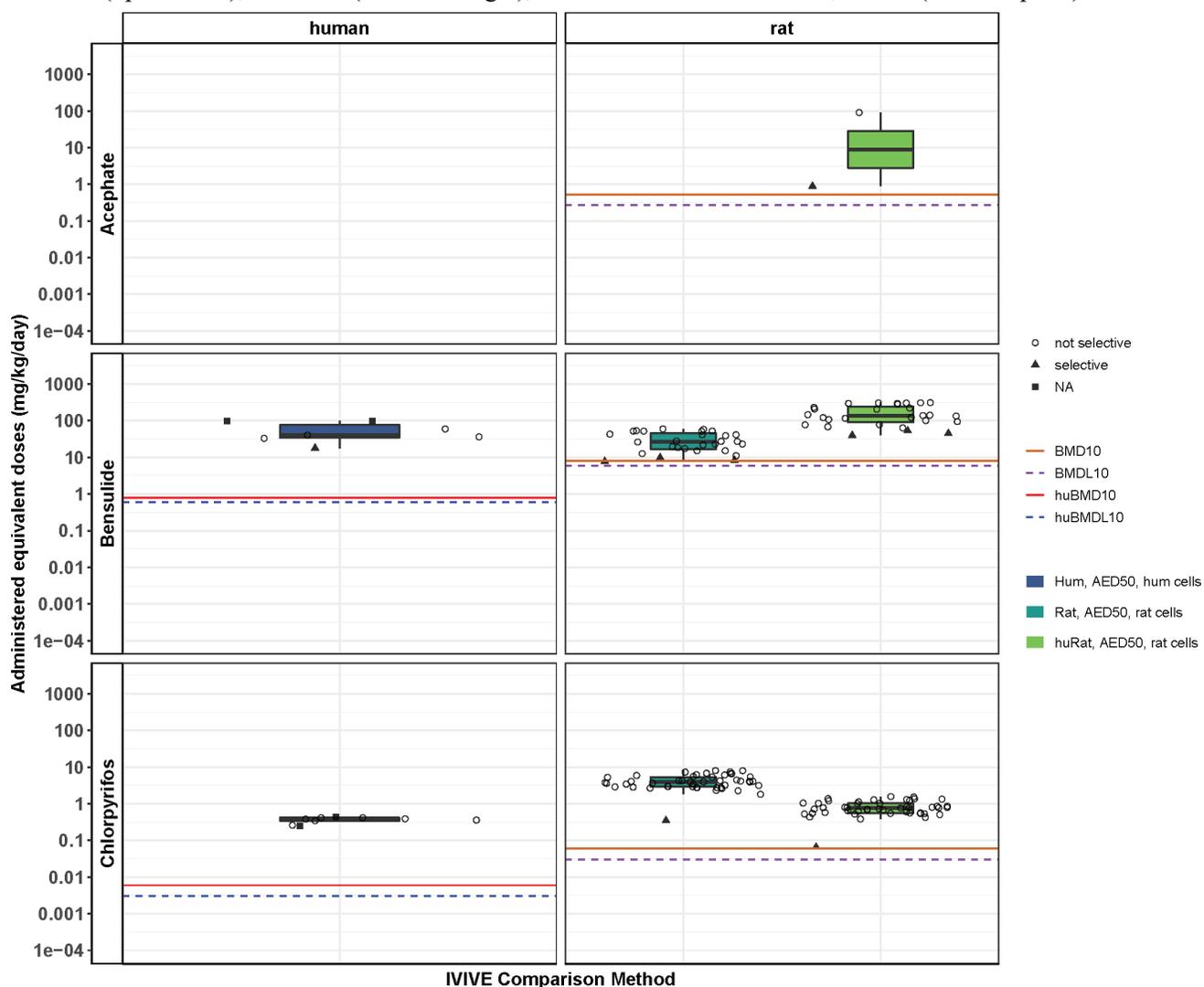
The OP chemical, DTXSID, number of positive assay endpoints for rat and human ((+) rat, (+) human), and the minimum (min) and median (med) AC50 values (μM) available for IVIVE are listed. In the case where replicate samples of a chemical were available, the data have been combined to the level of chemical. Three substances, chlorethoxyfos, naled, and Z-tetrachlorvinphos are italicized because AEDs could not be computed due to lack of reliable HTTK data measurements. A fourth substance, malaoxon, is italicized because there were no positive assay endpoint data available for comparison.

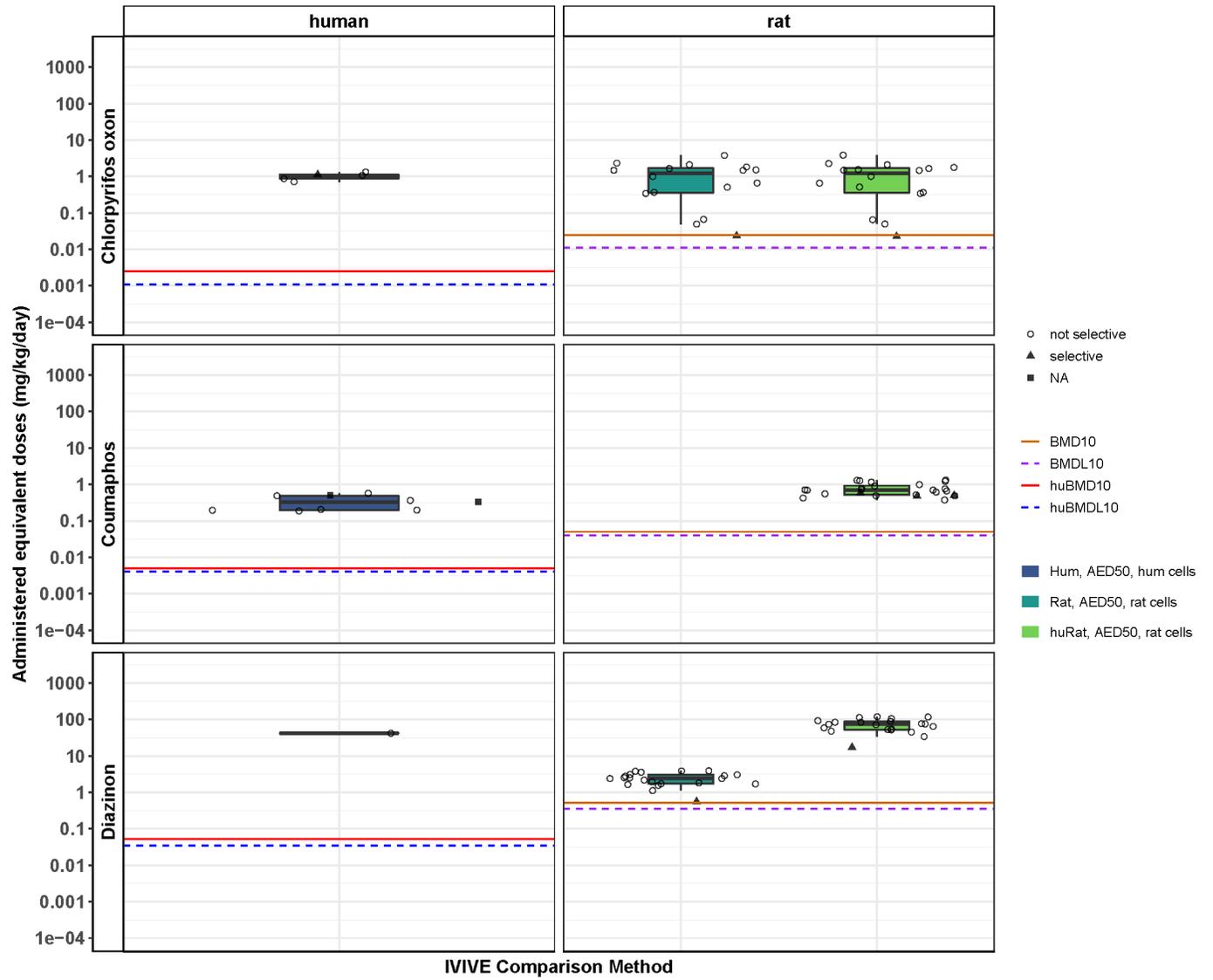
Chemical	DTXSID	(+) rat	Min rat AC50 (μM)	Med rat AC50 (μM)	(+) human	Min human AC50 (μM)	Med human AC50 (μM)
Acephate	DTXSID8023846	2	0.91	47.89	0	NA	NA
Bensulide	DTXSID9032329	28	11.69	39.31	7	11.57	25.15
<i>Chlorethoxyfos</i>	<i>DTXSID2032344</i>	28	<i>13.78</i>	<i>48.61</i>	3	<i>38.59</i>	<i>43.1</i>
Chlorpyrifos	DTXSID4020458	47	1.57	18.22	9	18.26	28.38
Chlorpyrifos oxon	DTXSID1038666	16	0.14	7.81	5	28.79	43.29
Coumaphos	DTXSID2020347	24	24.61	45.73	9	21.75	38.76
Diazinon	DTXSID9020407	21	12.07	52.06	1	53.38	53.38
Diazoxon	DTXSID5037523	1	0.88	0.88	0	NA	NA
Dichlorvos	DTXSID5020449	1	4.63	4.63	2	0.11	3.39
Dicrotophos	DTXSID9023914	1	91.72	91.72	0	NA	NA
Dimethoate	DTXSID7020479	39	11.54	39.26	1	70.17	70.17
Ethoprop	DTXSID4032611	17	3.84	25.58	0	NA	NA
Fosthiazate	DTXSID0034930	1	57.05	57.05	0	NA	NA
Malathion	DTXSID4020791	46	0.8	12.25	7	33.19	39.15
<i>Malaoxon</i>	<i>DTXSID9020790</i>	<i>0</i>	<i>NA</i>	<i>NA</i>	<i>0</i>	<i>NA</i>	<i>NA</i>
Methamidophos	DTXSID6024177	5	0.1	25.11	1	44.42	44.42
<i>Naled</i>	<i>DTXSID1024209</i>	27	<i>3.06</i>	<i>7.97</i>	4	<i>8.42</i>	<i>15.88</i>
Omethoate	DTXSID4037580	4	65.31	88.15	0	NA	NA
Phorate	DTXSID4032459	23	31.57	62.86	0	NA	NA
Phosmet	DTXSID5024261	19	8.01	33.59	5	29.12	32.98
Pirimiphos-methyl	DTXSID0024266	28	21.75	39.39	9	33.88	43.54
Profenofos	DTXSID3032464	2	1.41	3.63	0	NA	NA
Tebupirimfos	DTXSID1032482	19	44.07	81.17	0	NA	NA

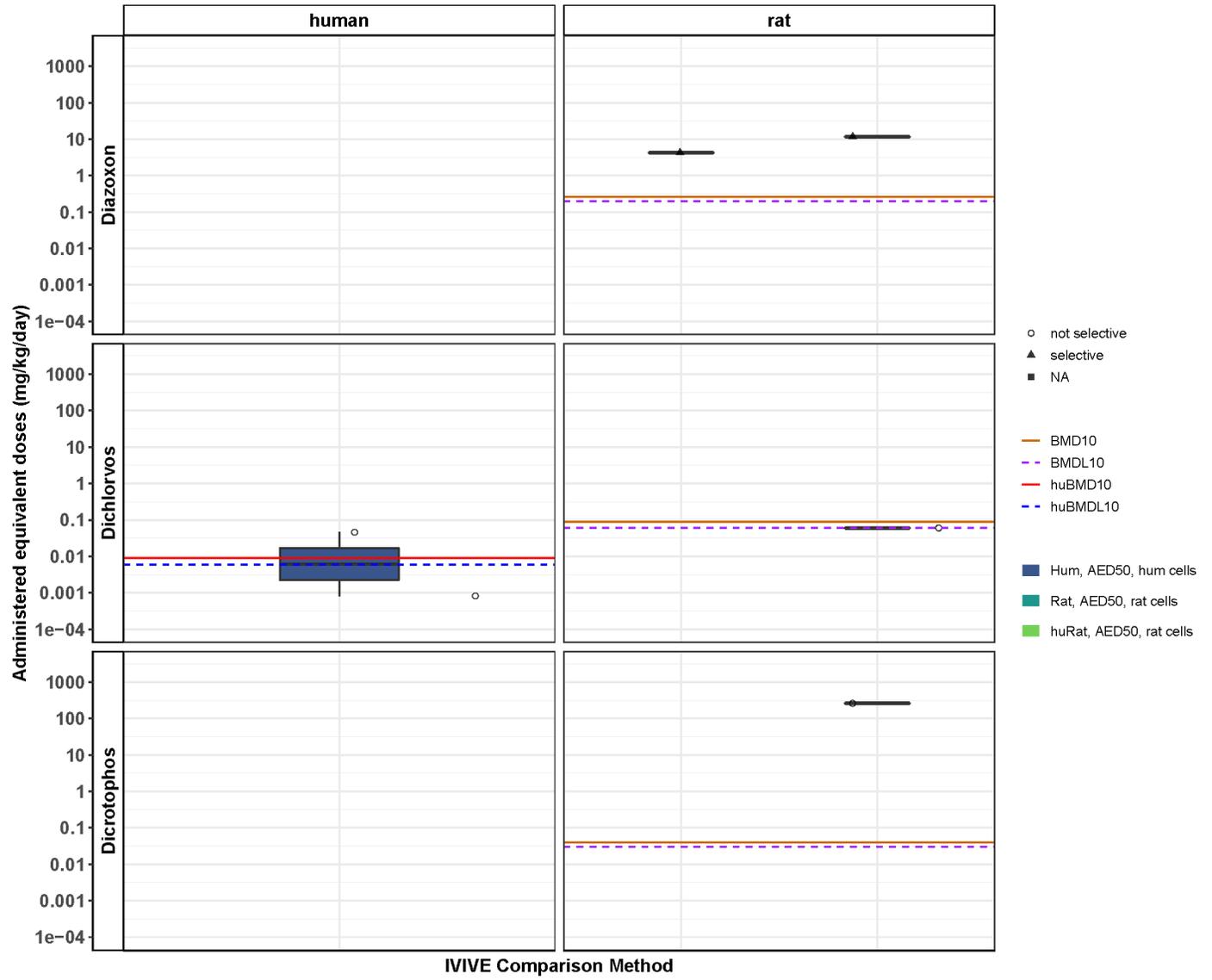
Chemical	DTXSID	(+) rat	Min rat AC50 ( $\mu\text{M}$ )	Med rat AC50 ( $\mu\text{M}$ )	(+) human	Min human AC50 ( $\mu\text{M}$ )	Med human AC50 ( $\mu\text{M}$ )
Terbufos	DTXSID2022254	28	4.09	19.04	3	0.01	35.68
Tribufos	DTXSID1024174	28	27.32	36.9	4	31.53	35.24
Trichlorfon	DTXSID0021389	35	1.91	32.39	3	14.81	57.05
Z-Tetrachlorvinphos	DTXSID1032648	15	31.52	91.4	7	3.85	6.73

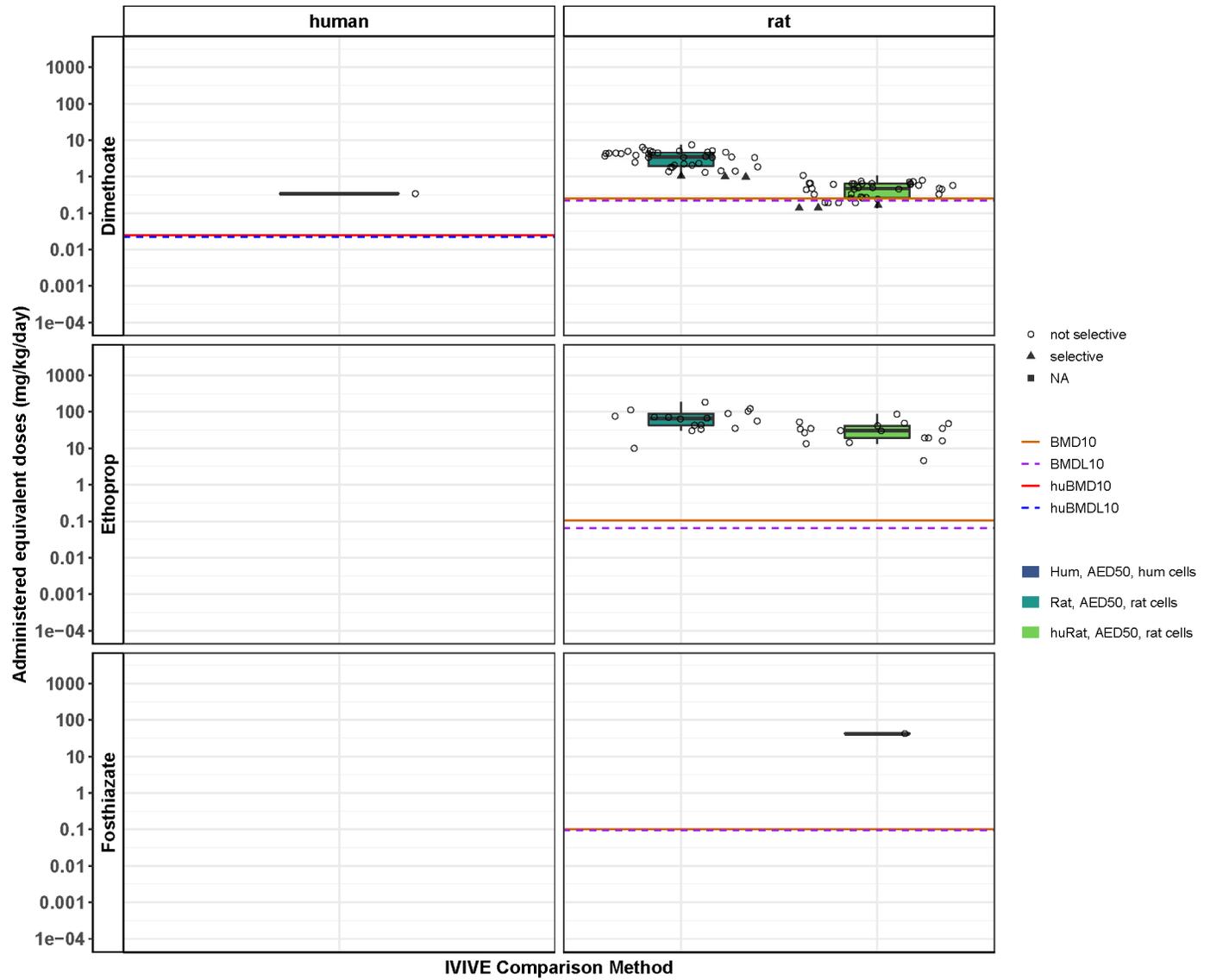
**Figure 7. Comparison of AED and BMD values by OP substance**

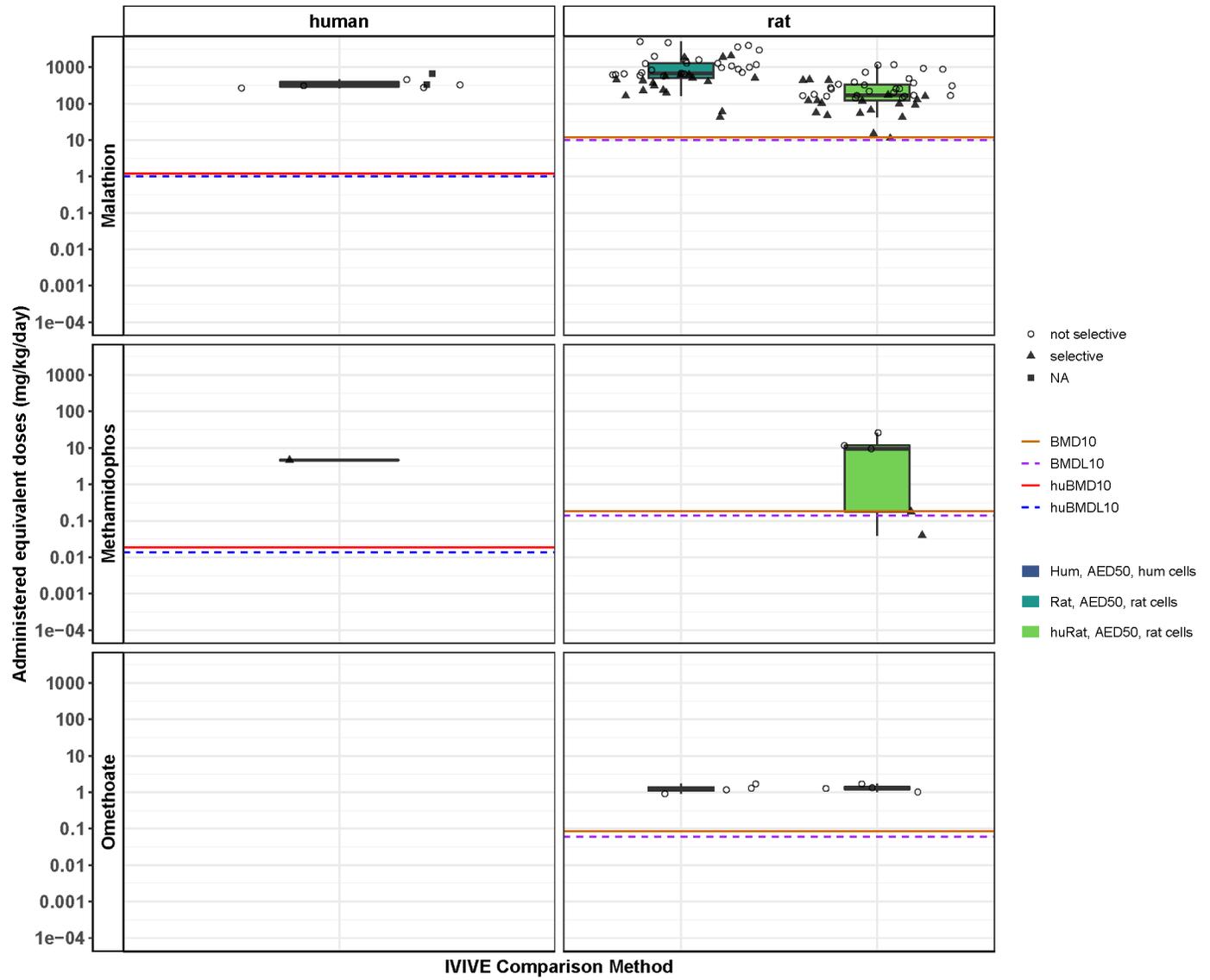
Each row contains two panels for human (left) and rat (right) IVIVE comparisons. Boxplots define the range of calculated human AED<sub>50</sub> values for human assays and rat/huRat AED<sub>50</sub> values for rat assays. BMD10 (solid orange) and BMDL10 (dashed purple) are indicated for rat by horizontal lines. These values are divided by an uncertainty factor of 10 to derive huBMD10 (solid red) and huBMDL10 (dashed blue) values represented by horizontal lines in the human comparison panels. The individual AED<sub>50</sub> values are superimposed over the boxplots (black points), with potential selectivity indicated as not selective (open circle), selective (closed triangle), and could not be calculated, or NA (closed square).

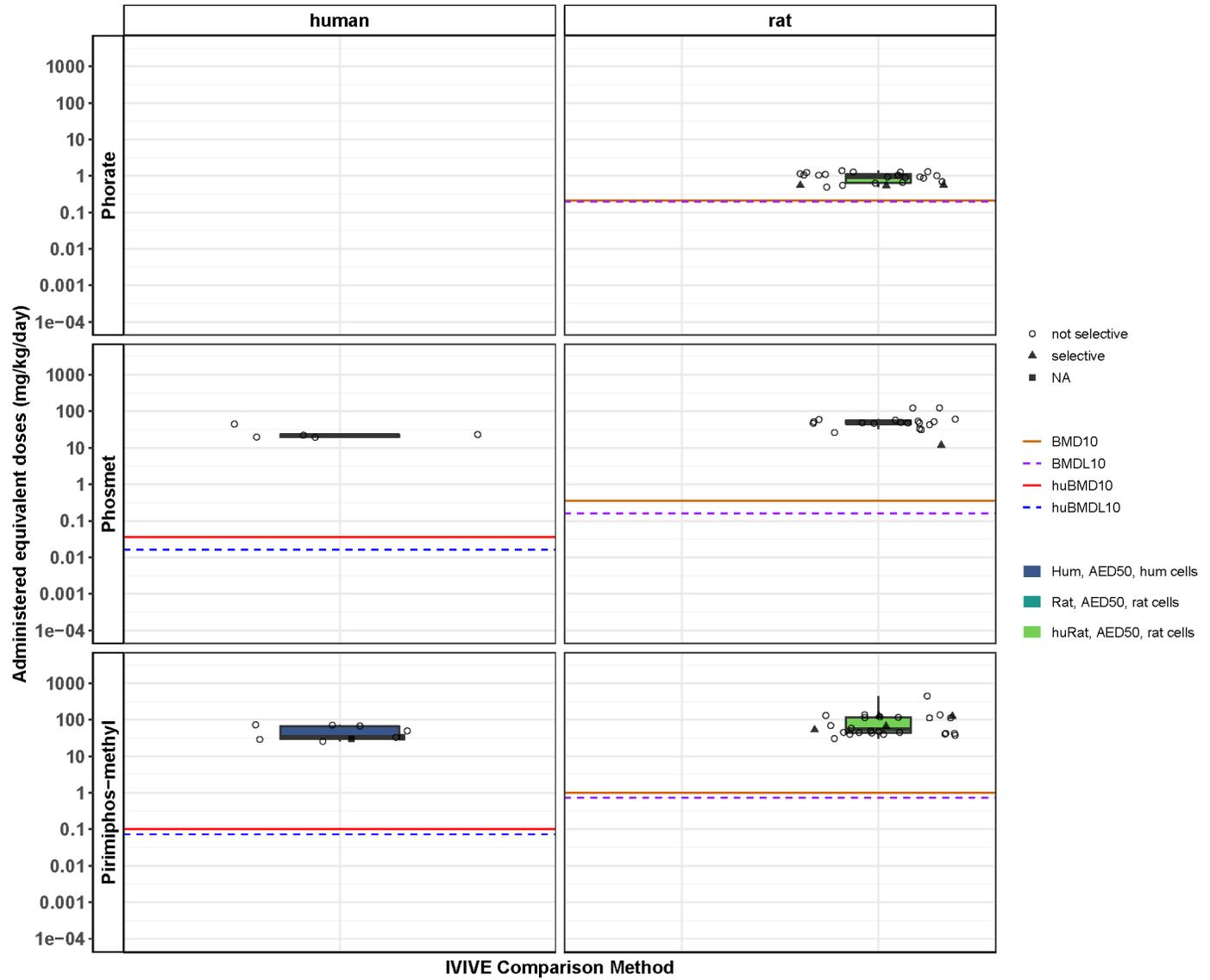


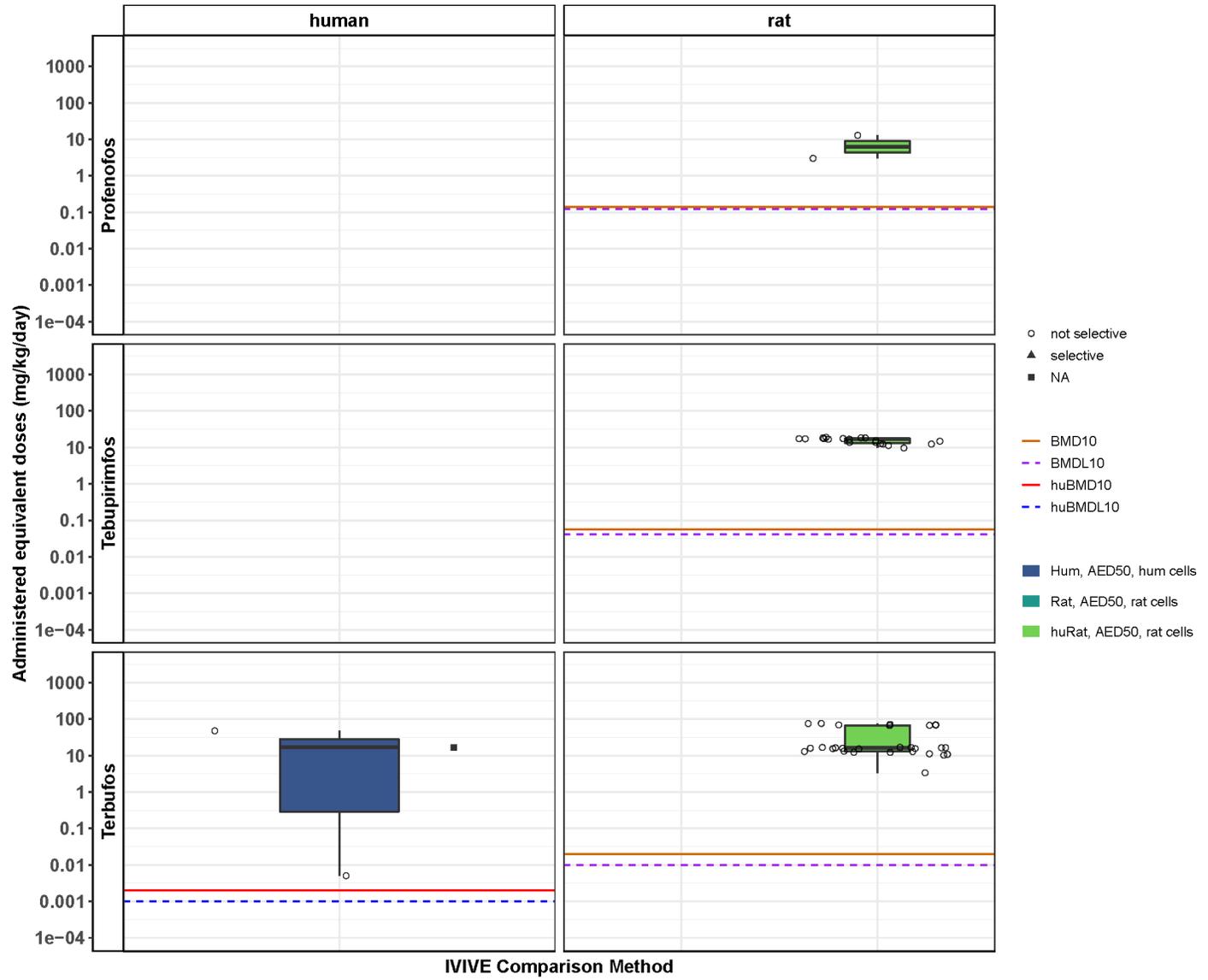


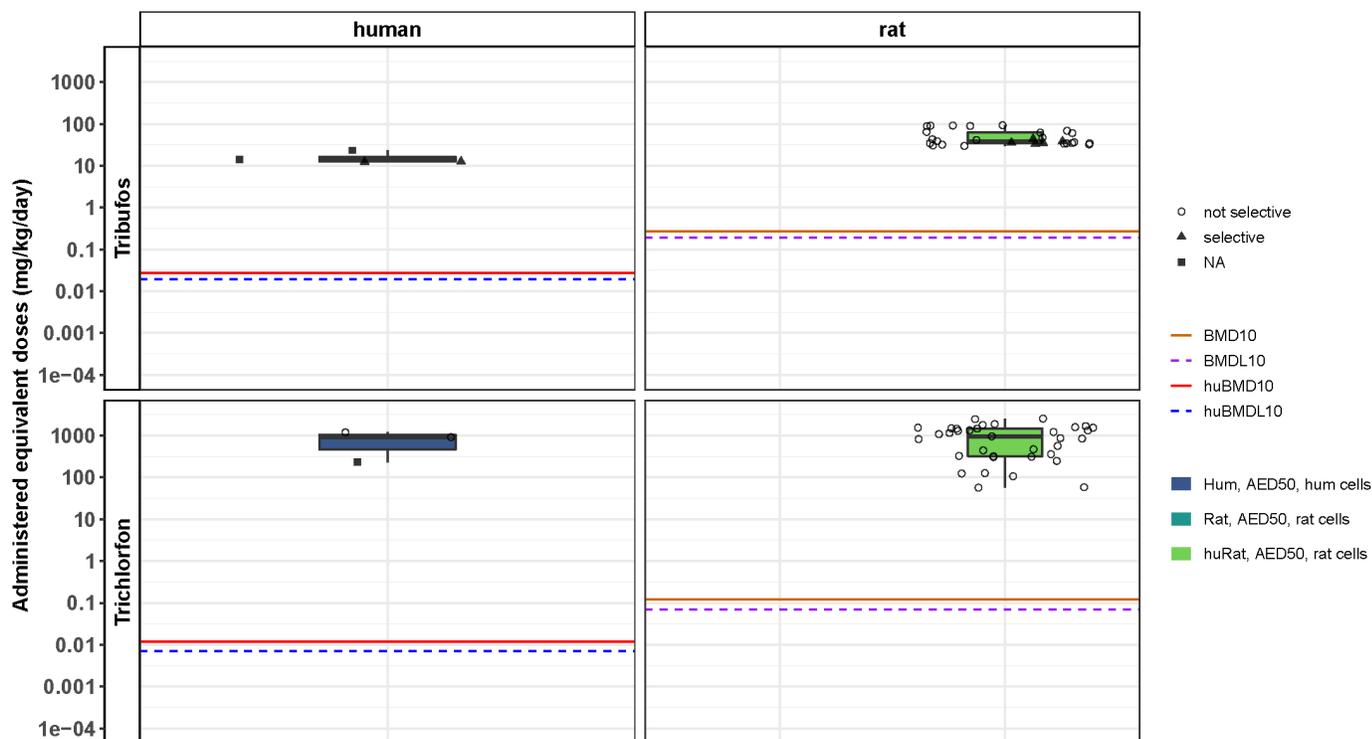












#### IVIVE Comparison Method

Some of the rat and huRat AED<sub>50</sub> values approached the *in vivo* rat BMD10 and BMDL10 thresholds identified using *in vivo* rat studies of AChE. The lower rat or huRat AED<sub>50</sub> values for acephate, bensulide, chlorpyrifos, chlorpyrifos oxon, diazinon, dimethoate, malathion, methamidophos, and phorate ranged within 1 log<sub>10</sub> order of magnitude of the BMD10 and BMDL10 values. For all of these aforementioned substances, at least a subset of the AED<sub>50</sub> values that range within 1 log<sub>10</sub> order of magnitude of the BMD10 and BMDL10 values are from selective bioactivity. The lower quartile of huRat AED<sub>50</sub> values for dimethoate and methamidophos appear similar to the BMD10 and BMDL10 values, and these AED<sub>50</sub> values appear to have included selective assay endpoints. The huRat AED<sub>50</sub> value for dichlorvos (only one positive rat assay endpoint) overlaps with the BMDL10 value, and it was not based on selective bioactivity in the DNT-NAM battery.

The human IVIVE comparison was more constrained because fewer OP chemicals had positive responses in this smaller subset of the DNT-NAM assay set. Acephate, diazoxon, dicotophos, ethoprop, fosthiazate, omethoate, phorate, profenofos, and tebupirimfos had positive rat assay data but lacked positive responses

in the human cell-based assays. For bensulide, chlorpyrifos, chlorpyrifos oxon, coumaphos, diazinon, dichlorvos, dimethoate, malathion, methamidophos, phosmet, pirimiphos-methyl, terbufos, tribufos, and trichlorfon, the human AED<sub>50</sub> values can be compared to BMD10/10 and BMDL10/10 values. With the exceptions of dichlorvos and terbufos, the full range of human AED<sub>50</sub> values are at least 2 log<sub>10</sub> orders of magnitude higher than the BMD10/10 and BMDL10/10 values. For dichlorvos, only two AED<sub>50</sub> values are available for comparison, and these values are centered around the BMD10/10 and BMDL10/10 values. Neither of these AED<sub>50</sub> values appear selective because the bioactivity was observed in assay endpoints relevant to cell viability. Similarly, for terbufos, only 3 human AED<sub>50</sub> values are available for comparison, and the lowest one of these values approaches the BMD10/10 value. This lowest AED<sub>50</sub> value for terbufos does not appear selective because it is derived from a cell viability related assay endpoint (object count in the HCI hNP1 proliferation assay endpoint). The other rat and human AED<sub>50</sub> values for terbufos appear to range approximately from 3 to 100 mg/kg/day.

The human AED<sub>50</sub> values and huRat AED<sub>50</sub> values were typically similar; both of these sets of values use the human HTTK data to inform human and rat models, respectively. Deviation between the human and huRat AED<sub>50</sub> values appears to be impacted by the smaller dataset available for human AED<sub>50</sub> derivation. Chemical-dependent differences between the rat and huRat AED<sub>50</sub> values are apparent when both are available; though for some chemicals (chlorpyrifos oxon, ethoprop, malathion, omethoate) the values are very similar, for other chemicals (bensulide, chlorpyrifos, diazinon, diazoxon, dimethoate) there may be as much as 1 log<sub>10</sub> order of magnitude separation between the median AED<sub>50</sub> values. There is no uniform direction to these differences, but these differences are expected and consistent with the impact of using human or rat HTTK data to inform a rat physiology-based model (Supplemental Appendix D).

Overall, these comparisons suggest that the doses required to achieve plasma concentrations (in the median individual in the general population) that demonstrate *in vitro* bioactivity relevant to DNT are higher than and in some cases approaching the doses that have been associated with significant changes in AChE activity in rats.

In addition to the HTTK model, the PBPK-PD models for dimethoate/omethoate and malathion were also used to estimate AEDs, assuming AC50 value from positive assay endpoints to be equivalent to steady state concentration in plasma. A PBPK-PD model was also available for malaoxon, but an AED cannot be estimated for malaoxon because all activity types examined in this Issue Paper were negative for malaoxon. The purpose of this exercise was to demonstrate the use of two different types of kinetic models in a tiered IVIVE approach, and to examine how AEDs differ when using different models. The simplest HTTK model only requires hepatic Clint to simulate steady state blood concentration, which has not been compared with data available for dimethoate, omethoate, and malathion. The customized PBPK-PD models require significant amount of time and resources to develop, but these models have been evaluated using chemical-specific blood and tissue concentration data. The dimethoate/omethoate and malathion models were developed based on the structure of a PBPK-PD model for chlorpyrifos, which had been reviewed by several SAPs (e.g., 2012, 2016<sup>10</sup>). Unfortunately, the chlorpyrifos model was written in a software (acslX) that was no longer available, so it cannot be used for estimating AEDs for chlorpyrifos and chlorpyrifos-oxon. The models for dimethoate/omethoate and malathion will be publicly available in the near future, and the review of these models is outside the scope of the current review. Since customized PBPK-PD models were only available for dimethoate, malathion, and their oxons, it is not possible to comment on particular comparisons or trends between the PBPK- and HTTK-derived AEDs for all OPs. Due to the amount of time and computing resource needed to compute the PBPK-derived AEDs, only the minimum, median, and maximum AC50 values from the DNT-NAMs were used

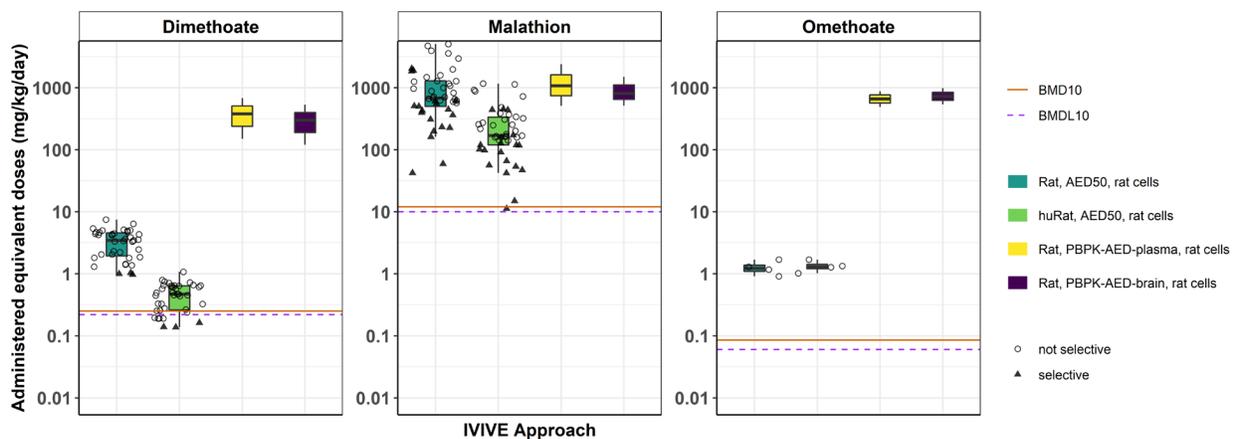
---

<sup>10</sup> <https://www.epa.gov/sap/fifra-scientific-advisory-panel-meetings>

as surrogates for steady state plasma AUC and brain AUC values. The plasma-based PBPK-AED values for the minimum, median, and maximum AC50 values were: 150, 380, 680 mg/kg/day (dimethoate); 510, 1080, 2390 mg/kg/day (malathion); and, 490, 660, 880 mg/kg/day (omethoate). Similarly, the brain-based PBPK-AED values for the minimum, median, and maximum AC50 values were: 120, 300, 530 mg/kg/day (dimethoate); 515, 810, 1500 mg/kg/day (malathion); and, 540, 730, 980 mg/kg/day (omethoate). For dimethoate and omethoate, the PBPK-AED values using plasma and brain AUC were more than two orders of magnitude greater than the HTTK-derived AEDs (Figure 8). For malathion, the PBPK-derived AED values were similar to the range of HTTK-derived AED<sub>50</sub> values for rat. In addition to predicting chemical disposition over time, these PBPK-PD models were also used to predict AChE inhibition in RBCs and in brain at steady state plasma concentrations equivalent to DNT-NAM AC50 values. For dimethoate and omethoate, both RBC and brain AChE was almost completely inhibited (> 99.9% for omethoate, >99% for dimethoate). For malathion, RBC AChE was at least 95% inhibited, and brain AChE was 35%-75% inhibited. For the three OPs examined, the HTTK-derived AEDs are more conservative estimates of the dose required to achieve blood concentrations approaching the bioactive AC50s from DNT-NAMs. Though the HTTK model employed and the PBPK-PD models all assumed 100% bioavailability, the HTTK model accounts for hepatic Clint whereas PBPK-PD models incorporate additional metabolism sites in plasma, brain, and kidneys.

### Figure 8. Comparison of PBPK-PD model-derived AEDs with HTTK-derived AEDs.

For dimethoate, malathion, and omethoate, more complex PBPK-PD models were used to predict AED values (mg/kg/day) using the minimum, median, and maximum DNT-NAM AC50 values only from assays that employed primary rat cortical neurons. In the PBPK-PD model, these AC50 concentrations (µM) were assumed to reflect plasma or brain AUC values, resulting in two PBPK-PD AED predictions, respectively (Rat, PBPK-AED-plasma, rat cells [yellow]; Rat, PBPK-AED-brain, rat cells [purple]). These were compared to the HTTK-derived AEDs for rat and huRat (also presented and described above in Figure 7). Boxplots define the median and range of calculated PBPK-PD AED values for rat assays and the rat/huRat AED<sub>50</sub> values for rat assays. BMD10 (solid orange) and BMDL10 (dashed purple) are indicated for rat by horizontal lines. For the HTTK-derived AED<sub>50</sub> values, the individual AED<sub>50</sub> values are superimposed over the boxplots (black points), with potential selectivity indicated as not selective (open circle), selective (closed triangle), and could not be calculated, or NA (closed square).



### 2.3.7 Future research directions to increase utility of DNT-NAMs

As discussed in Section 2.1, EPA has shifted its testing focus from the *in vivo* DNT guideline study to more targeted testing due to several challenges associated with the study and its limited impact on human health risk assessments for pesticides. The development of a DNT-NAM battery for assessing potential DNT-related effects provides an opportunity to overcome some of the challenges with the *in vivo* DNT guideline study by evaluating underlying critical processes of neurodevelopment and incorporating human relevant information. Although the DNT-NAM battery presented in this Issue Paper covers many of the processes critical to neurodevelopment, the results from additional assays that provide information on other processes (migration, neuronal differentiation, and oligodendrocyte differentiation and maturation), or provide orthogonal/confirmatory information on processes currently covered (proliferation, neurite outgrowth), will be evaluated for their utility as they become available. As detailed in Section 2.3.2, on-going work suggests that more assays may be useful to evaluate processes not evaluated in the MEA-NFA or HCI assays, including neuronal migration (Barenys et al., 2017; Baumann et al., 2016; Nyffeler et al., 2017), some aspects of differentiation of neuronal subtypes (Barenys et al., 2017; Baumann et al., 2016); and, oligodendrocyte differentiation and maturation (Barenys et al., 2017; Baumann et al., 2016). At this time, it is unknown as to whether the additional orthogonal assays for proliferation, apoptosis, migration, neuron differentiation, oligodendrocyte differentiation and maturation, neurite outgrowth, synaptogenesis, or network formation (Sachana et al., 2019) are needed to improve prediction of DNT potential for OPs or other chemicals. Further research using more chemicals may be useful in characterizing the potential value of additional data above and beyond the DNT-NAM battery presented here. In Charge Question 3, the Agency is soliciting comment from the SAP to consider whether any additional processes, beyond those covered by the MEA NFA, HCI, and additional assays being developed via international efforts, should be considered with regard to the development of an internationally recognized DNT-NAM battery.

The DNT-NAM battery presented in this Issue Paper reflects the use of MEA NFA and HCI in an integrated way. Additional work to understand the combined value of the MEA-NFA and HCI cellular event assays, taken together, for an *in vitro* DNT activity reference chemical list (Aschner et al., 2017; Crofton et al., 2011; Mundy et al., 2015) may be informative in the future in terms of identifying strengths, weaknesses, and limitations for these assay technologies in identifying putative DNT effects *in vitro*. However, an uncertainty in this type of analysis for a DNT-NAM battery, as well as the *in vivo* DNT guideline study, is that there are so few chemicals for which human evidence of DNT is well-accepted (Grandjean and Landrigan, 2006; Grandjean and Landrigan, 2014). Evaluation of drugs with known neural activity could be used to demonstrate what assay parameters may be needed or not needed for future bioactivity screening efforts to identify *in vitro* DNT activity (as a research activity).

The impact of the cell types chosen for the MEA NFA and HCI assays is not yet fully characterized, and additional research may be needed to better understand differences between cell types in the MEA NFA and HCI assays. Primary rat cortical neurons were used in the MEA NFA and a subset of the HCI assays evaluating neuronal maturation and synaptogenesis due to a lack of network-forming (i.e., synaptically-competent) human cell models at the time these assays were developed. An advantage of using primary rat cortical cells is that the results can be compared to existing *in vivo* rodent models of DNT. It is unknown how the use of primary rat cortical neurons and human cell lines (hN2 and hNP1) in the HCI assays and rat cortical neurons in MEA NFA may produce different responses than iPSC-derived human neurons (Tukker et al., 2020). As the field of stem cell research advances and iPSC-derived human neurons become more affordable for research and screening purposes, human cell models may replace

rodent cells for the MEA NFA and other HCI assays. Though the specific pattern of effects in the MEA NFA, or the HCI assays, may differ by cell type used, the approach applied to this case study considers any effect to estimate a threshold for bioactivity in these DNT-NAMs. For the case study presented herein, any perturbation in the MEA NFA and HCI assay endpoint data using primary rat cortical cells were compared to rat BMD/BMDL values, and hN2 and hNP1 derived data from the HCI were compared to BMD/BMDL values divided by 10, regardless of the specific pattern of effects on biological processes.

A common limitation in applying *in vitro* bioactivity screening data has been the use of models that do not have the same metabolism as the *in vivo* scenario (DeGroot et al., 2018). The metabolic capacity of these DNT-NAMs is not well understood; however, in the case study presented herein, important metabolites (i.e., oxon metabolites) were known and tested as separate chemical samples for several OPs (e.g., omethoate, malaoxon, chlorpyrifos oxon); however, oxons were not tested for all the OPs that require activation (e.g., phosmet, bensulide, phorate). Brain tissue does express cytochrome P450 enzymes, and typically different isozymes than those expressed in the liver. The DNT-NAMs used here have not been thoroughly examined for expression of Phase I or II enzymes. While some limited metabolism of substances is possible, the extent to which this occurs, and the substances that are possible substrates, cannot be predicted at this time. In the future, important metabolites could be predicted and screened as separate samples in the DNT-NAMs as needed like oxons in the OP case study, or more complex retrofitting of the assays with metabolism could be considered (DeGroot et al., 2018).

### 2.3.8 Summary and conclusions on the DNT-NAMs

In Sections 2.3.1 to 2.3.4, extensive evidence and published references are provided to indicate that the MEA NFA and HCI assays allow for measurement of disruption of neural network formation and function as well as key cellular events for DNT. Assay reference chemicals (Section 2.3.4 and Supplemental Appendix A) known to perturb these processes *in vitro* provide reproducible responses in these assays on both a quantitative and qualitative basis, increasing confidence that these assays provide a reliable signal of bioactivity possibly relevant to DNT. In the MEA NFA, the positive control substances (loperamide hydrochloride, bisindolylmaleimide I, L-domoic acid, mevastatin, and sodium orthovanadate) work to inhibit (decrease) neuronal network formation. In the HCI assays, similarly, the positive controls (aphidocholin, lithium chloride, bisindolylmaleimide I, sodium orthovanadate, NSC 23766 trihydrochloride, and staurosporine) all work to inhibit key critical neurodevelopmental processes (i.e., neurite outgrowth, synaptogenesis, proliferation) in specific cell types, save staurosporine, which reduces cell viability via apoptosis (Supplemental Appendix A). Overall, these assays appear robust and reproducible for fit-for-purpose application to identify putative DNT bioactivity, with reasonable assay performance controls established. The current DNT-NAM battery covers much of the network of DNT-relevant processes suggested in Section 2.3.2 and represents a major milestone in filling data gaps for evaluation of putative DNT activity. Charge Questions 1 and 2 solicit comment from the SAP on the strengths and limitations of these assays for fit-for-purpose evaluation of putative DNT-related activity.

As described in detail in Section 2.3.5, for the case study application of DNT-NAMs to OPs, there are some key findings for this 27-chemical set. First, the range of log<sub>10</sub>-AC<sub>50</sub> values for active assay endpoints was similar for the MEA NFA and the HCI assays, and relatively limited in terms of the total concentration range (0-2 log<sub>10</sub>-μM). Though not all OPs demonstrated the same pattern of effects in the MEA NFA and HCI assays, chemicals that were active in the HCI assays tended to be active in the MEA NFA, and negative/equivocal activity in the MEA NFA appeared to correspond to limited, sparse activity in the HCI assays, giving some confidence in strong positives and strong negatives across these two

different technologies. There is only 1 OP chemical (ethoprop) that was positive in MEA NFA and demonstrated minimal activity in the HCI assays, indicating that negatives tend to be consistent between MEA NFA and HCI assays. Acephate, diazoxon, dichlorvos, dicrotophos, malaoxon, omethoate, and profenofos demonstrated limited to no activity in the MEA NFA and the HCI assays (Table 11). Considering all chemicals with greater activity in the HCI assays, only 1 OP chemical (methamidophos) was not positive in the MEA NFA, indicating that if activity is observed in the HCI assays, it is likely that the OP chemical will also be active in the MEA NFA.

Currently there are some limitations in understanding whether cytotoxicity in neural cell types is relevant for putative DNT-related bioactivity. As such, efforts have been made to distinguish the “selectivity” of OPs in the MEA NFA and HCI assays, i.e., bioactivity in these assays at concentrations at least 0.3 log<sub>10</sub>-micromolar less than the concentrations at which cytotoxicity was observed, in Sections 2.3.5 and 2.3.6. In comparing AED<sub>50</sub> values from the DNT-NAMs to BMD/BMDL values, it is evident that a subset of the AED<sub>50</sub> values were based on “selective” activity in the DNT-NAMs, and that these “selective” AED<sub>50</sub> values tend to fall on the lower end of the distribution of AEDs, having been derived from lower AC50 values.

The AED<sub>50</sub> values from DNT-NAM bioactivity presented for comparison, importantly, used the median individual in the general population for the HTTK-based IVIVE approach; accounting for interindividual variability (i.e., first order hepatic clearance, plasma protein binding, liver physiology, and glomerular filtration rate) and/or using a more sensitive individual or subpopulation would result in different AED values (indeed, lower values would result by using an estimate of a more toxicokinetically-sensitive individual). Within the case study application presented here, the median individual from the general population was chosen because susceptible subpopulations and population-based differences are handled in separate parts of the risk assessment. Overall, the AED<sub>50</sub> to BMD10/BMDL10 comparisons suggest that the doses required to achieve plasma concentrations (in the median individual in the general population) that demonstrate *in vitro* bioactivity relevant to DNT are higher than and in some cases approach the doses that have been associated with significant changes in AChE activity in rats. Oxon metabolites were tested as separate chemical samples for several OPs; however, oxons were not tested for all the OPs that require activation and lack of these data will need to be considered in the context of this case study. In Charge Question 4, the Agency is soliciting comments on the strengths and limitations of the IVIVE approach taken in this proof-of-concept application for evaluation of putative DNT bioactivity for OPs.

The analysis presented in this Issue Paper demonstrates several conclusions:

- the MEA NFA and HCI assay suite recapitulates key cellular events and processes relevant to DNT, as demonstrated through the use of appropriate assay performance controls;
- the DNT-NAMs presented here represent a major milestone for *in vitro* fit-for-purpose identification of putative DNT-related hazard, though additional methods may be available in the future;
- the MEA NFA and HCI assay suite demonstrates reproducibility in terms of positive responses and potency of these responses;
- the 27 OP chemicals in this set are differentially active in the MEA NFA and HCI assay suite; and,

- application of IVIVE approaches for the *in vitro* bioactivity observed in these DNT-NAMs results in AED<sub>50</sub> values that are greater than or in some cases approximate the doses that inhibit AChE *in vivo*.

As discussed in Section 2.1, EPA has shifted its testing focus from the *in vivo* DNT guideline study to more targeted testing due to several challenges associated with the study and its limited impact on human health risk assessments for pesticides. The development of a DNT-NAM battery for assessing potential DNT-related effects provides an opportunity to overcome some of the challenges with the *in vivo* DNT guideline study by evaluating underlying critical processes of neurodevelopment and incorporating human relevant information. The MEA NFA and HCI assays provide an opportunity to evaluate the majority of the critical processes of neurodevelopment (as presented in Table 3 and Figure 2) and represent a significant advancement toward developing a DNT-NAM battery for fit-for-purpose DNT evaluation. Additional assays are currently under development by EFSA-funded researchers that evaluate processes not covered by the MEA NFA and HCI assays (i.e., neuronal migration and oligodendrocyte differentiation and maturation). The relative contribution of these assays for a NAM battery will be considered once data are available. Additionally, any OP data from these other assays will be considered in combination with the results of the MEA NFA and HCI assays as part of an overall weight of evidence evaluation of the DNT potential for individual OPs.

### 3.0 Development of DDEFs for Interspecies and Intraspecies Extrapolation

#### 3.1 Background

For risk assessment, default uncertainty factors are commonly applied to extrapolate toxicity data derived from animal models to humans (interspecies or UF<sub>A</sub>) and to account for human variability (intraspecies or UF<sub>H</sub>). For OP human health risk assessments, except chlorpyrifos, the default 10X interspecies and 10X intraspecies extrapolation factors have been applied. For chlorpyrifos, the PBPK-PD model is used to derive PODs. Since the model accounts for the pharmacokinetic and pharmacodynamic differences between animals and humans, the default interspecies factor was reduced to 1X for chlorpyrifos human health risk assessments. Similar PBPK-PD models have been developed for dimethoate and malathion and their oxons; these models are under review by EPA for their potential use in future risk assessments.

The EPA's 2014 *Guidance for Applying Quantitative Data to Develop DDEF for Interspecies and Intraspecies Extrapolation*<sup>11</sup> provides guidance on the process for identifying reliable data that are useful for quantifying interspecies and intraspecies differences to serve as the basis for empirically deriving DDEFs. When using DDEFs, interspecies and intraspecies extrapolation factors are divided into two components representing toxicokinetic (TK) variability and toxicodynamic (TD) variability. Therefore, four DDEFs can be calculated given sufficient information. Two extrapolation factors for interspecies extrapolation are: 1) extrapolation factor covering interspecies toxicokinetics (EF<sub>AK</sub>) to account for TK variability and 2) extrapolation factor for interspecies toxicodynamics (EF<sub>AD</sub>) to account for TD variability. Similarly, the two extrapolation factors for intraspecies extrapolation are: 1) extrapolation factor covering intraspecies toxicokinetics (EF<sub>HK</sub>) to account for TK variability; and 2) extrapolation factor for intraspecies toxicodynamics (EF<sub>HD</sub>) to account for TD variability. The composite factor (CF) is calculated after the appropriate DDEF values for interspecies and intraspecies differences in TK and TD have been derived as shown in Equation 3.

<sup>11</sup> <https://www.epa.gov/risk/guidance-applying-quantitative-data-develop-data-derived-extrapolation-factors-interspecies-and>

$$CF = EF_{AK} \times EF_{AD} \times EF_{HK} \times EF_{HD} \quad \text{Equation 3}$$

The CF calculation is analogous to calculating composite uncertainty factors when using the 10X defaults for  $UF_A$  and  $UF_H$ . If data are only available to develop a DDEF for one component of extrapolation or another, the remaining extrapolation is done by an appropriate default procedure.

As described in the EPA's DDEF guidance, information on MOA, even when a complete understanding of mechanism is not available, is important for DDEF derivation since DDEFs are considered in the context of toxicity endpoints most relevant for risk assessment purposes. As discussed in Section 2.2, EPA has determined that the OPs share a common mechanism of toxicity based on their shared ability to inhibit AChE.

In 2016, three OP pesticide registrants (AMVAC, FMC, and Gowan) worked in conjunction with their consultant (Exponent) and Dr. Janice Chambers from Mississippi State University to develop an experimental plan to determine if differences exist in AChE inhibition between rats and humans and estimate intra-human variability and these differences. Representatives of OPP and ORD considered the plan and supported continued development of these efforts to generate data to inform the TD variability in the DDEF development. Since 2016, there have been numerous meetings with EPA to discuss the experimental plan, data review, and statistical analyses. EPA provided feedback on experimental protocols to determine the AChE inhibition kinetic constants of OP compounds to inform interspecies and intraspecies pharmacodynamic differences. Three study reports were submitted testing 17 OP pesticides and oxons:

- MRID 50773501: malaoxon (the active metabolite of malathion) and omethoate (the active metabolite of dimethoate)
- MRID 50773502: DDVP, naled, dicotophos, tribufos, phorate oxon sulfone (phorate metabolite), phorate oxon sulfoxide (phorate metabolite), ethoprop, methamidophos, fenamiphos, terbufos oxon sulfone (terbufos metabolite), terbufos oxon sulfoxide (terbufos metabolite), chlorethoxyfos oxon (chlorethoxyfos metabolite), and tebupirimphos oxon (tebupirimphos metabolite)
- MRID 50773503: bensulide oxon (the active metabolite of bensulide) and phosmet oxon (the active metabolite of phosmet).

A separate report was submitted on using these data to calculate pharmacodynamic DDEFs (MRID 50773504). This report includes an evaluation by Exponent of the criteria outlined in the EPA's DDEF guidance for using *in vitro* data for estimating DDEFs<sup>12</sup>, which includes:

- Was the toxicologically active form of the agent studied?
- How directly was the measured response linked to the adverse effect?
- Are the biological samples used in the assays derived from equivalent organs, tissues, cell types, age, stage of development, and sex of the animals/humans in which the target organ toxicity was identified?

---

<sup>12</sup> These criteria can be found in Section 2.4 on page 19 of the EPA's 2014 *Guidance for Applying Quantitative Data to Develop DDEFs for Interspecies and Intraspecies Extrapolation*.

- What is the range of variability (e.g., diverse human populations and life stages) that the biological materials cover?
- If the effect occurs or can be measured in several tissues, is the studied tissue or tissue preparation an appropriate surrogate? Or, in situations where the effect is not localized, is the effect consistent across tissues?
- Does the design of the study allow for statistically valid comparisons based on such factors as replicate and sample size?
- Was chemical uptake considered when the chemical was applied to the samples so as to give comparable intracellular concentrations across tissues?
- Were similar tissues or samples evaluated across species?
- Do the concentrations in the *in vitro* studies allow for comparison with *in vivo* conditions?

EPA believes all of the criteria have been appropriately considered and met for use of *in vitro* data for estimating pharmacodynamic DDEFs for OP compounds based on the evaluation provided by Exponent; however, EPA has noted concerns with the sample size for intraspecies analyses. EPA provided feedback on the statistical analyses performed in this report, as well as supplemental statistical analyses (MRID 51182301) that were subsequently submitted to address EPA's comments and concerns. Although statistical analyses and results are presented in earlier reports from Exponent, the SAP should focus on the statistical analyses and results presented in the supplemental analysis (MRID 51182301) for their evaluation and comment. Lastly, EPA has reviewed a supplemental whitepaper submitted by Exponent (MRID 50773504) that provides a summary of existing knowledge regarding AChE in rats and humans, including amino acid sequence alignments and 3D structures.

## 3.2 Methods

### 3.2.1 Laboratory Experiments

The methods used to measure kinetic constants (bimolecular rate constant  $k_i$ , dissociation constant  $K_I$ , and phosphorylation constant  $k_p$ ) for AChE inhibition are described in detail in MRIDs 50773501-50773503. Briefly, a continuous spectrophotometric assay was used to determine AChE activities for 17 OPs, as well as paraoxon as the positive control. The assays were performed using "erythrocyte ghost" preparations (i.e., erythrocyte cell membranes separated from hemoglobin and other cytoplasmic constituents) that were obtained from either human or rat erythrocytes as the source of AChE. For humans, the kinetic constants were determined for 18 individual samples (9 adults, 5 juveniles, and 4 cord blood samples). Human AChE utilized in the experiments was derived from blood samples from individual healthy humans of both sexes (adults age 16-60, and juveniles age 10-13), as well as cord blood samples. Blood samples from multiple race and ethnic groups were included in the study. For rats, the kinetic constants were determined for three individual pooled samples from adult rats (Sprague Dawley [CrI:CD(SD)BR]). Each male sample (n=3) was prepared from the pooled blood of five male rats. Similarly, each female sample (n=3) was prepared from the pooled blood of five female rats.

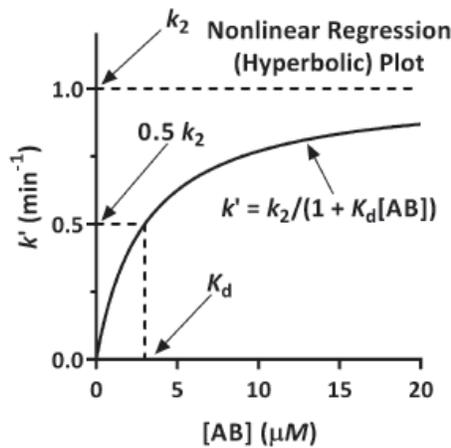
The absorbance in each well was measured and the velocity of each reaction was calculated by determining the slope of the line from a plot of absorbance (proportional to product formed) as a function of time. The inhibition reaction was terminated by addition of the substrate acetylthiocholine. For each species and each inhibitor, all regressions and calculations were performed using either Microsoft Excel 2010 or SigmaPlot version 14 to obtain the AChE velocity remaining and apparent rate of AChE

phosphorylation ( $k_{app}$ ) for determination of  $k_i$ ,  $k_p$ , and  $K_I$ . The key result from the experiments was  $k_i$ , which is directly proportional to the rate of *in vivo* inhibition for a given internal dose. Since the inhibition of AChE from an internal dose is a pharmacodynamic process, it can be directly used in the estimation of DDEFs.

### 3.2.2 Statistical Approach

For each of the 16 OPs, Exponent used SAS PROC NL MIXED to fit nonlinear mixed models to  $k_{app}$  (the apparent rate of AChE phosphorylation in mol/L/min) as a function of the inhibitor concentration (*Conc*). These mixed models were specified to have a hyperbolic functional form (referred to as the classic Michaelis-Menten equation for describing the kinetics of enzyme-substrate reactions), where the parameter estimates correspond to the phosphorylation rate constant ( $A$ ,  $min^{-1}$ ) and disassociation rate constant ( $B$ ,  $mol/L$ ):

$$k_{app} = \frac{A * Conc}{B + Conc}$$



In turn,  $k_i$  was estimated as the ratio of the phosphorylation rate constant and the dissociation rate constant ( $A/B$ ) parameters from nonlinear mixed model.

$$k_i = \text{Bimolecular Rate Constant (L/mol/min)} = k_p / K_I$$

As part of the nonlinear mixed-effects model approach where  $k_{app}$  was nonlinearly regressed on inhibition concentration (*Conc*), each sample<sup>13</sup> was treated as a unique experimental subject, such that each sample was allowed to have a different hyperbolic curve (i.e., each sample had a different random value of  $A$  and different random value of  $B$ ). Additionally, in the nonlinear mixed-effects models, each group (e.g., species, sex, age, or ethnicity groups) is allowed to have different average  $A$  value and average  $B$  value (i.e., each group has a fixed effect  $A_m$  value and fixed effect  $B_m$  value). These fixed effect values are the average of all random  $A$  values or the average of all random  $B$  values within each group. Thus, each group is allowed to have an “average” hyperbolic curve.

<sup>13</sup> The term “sample” refers to blood samples either collected from an individual human subject or pooled together from a group of rats.

$$k_{app.mit} = \frac{(A_m + \gamma_{mi}) \times Conc_t}{(B_m + \varphi_{mi}) + Conc_t} + \varepsilon_{mit}$$

Where  $k_{app.mit}$  is the  $k_{app}$  value of the sample  $i^{th}$  in the group  $m$  at the inhibition concentration  $Conc_t$ ,  $A_m$  and  $B_m$  are the fixed effects  $A$  and  $B$  of group  $m$ ,  $\gamma_{mi}$  and  $\varphi_{mi}$  are the random effects associated with  $A_m$  and  $B_m$  of the sample  $i^{th}$  in the group  $m$ ,  $\gamma_{mi}$  and  $\varphi_{mi}$  follow normal distributions with mean = 0 and

$$\text{variance-covariance matrix} = \begin{bmatrix} G_{Am} & G_{ABm} \\ G_{ABm} & G_{Bm} \end{bmatrix}, \text{ and}$$

$\varepsilon_{mit}$  is the error of the measurement of the sample  $i^{th}$  in the group  $m$  at the inhibition concentration  $Conc_t$ , with  $\varepsilon_{mit} \sim N(0, \sigma^2)$ .

The above equation can be rewritten and implemented in SAS as:

$$k_{app.mit} = \frac{\{(A_1 + \gamma_{1i}) \times I[m = 1] + (A_2 + \gamma_{2i}) \times I[m = 2]\} \times Conc_t}{\{(B_1 + \varphi_{1i}) \times I[m = 1] + (B_2 + \varphi_{2i}) \times I[m = 2]\} + Conc_t} + \varepsilon_{mit}$$

where,  $A_1$  and  $B_1$  are the fixed effects  $A$  and  $B$  of group  $m=1$ ,  $A_2$  and  $B_2$  are the fixed effects  $A$  and  $B$  of group  $m=2$ ,

$$I[m=1] = \begin{cases} 1 & \text{if } m = 1 \\ 0 & \text{if } m = 2 \end{cases}, \text{ and}$$

$$I[m=2] = \begin{cases} 1 & \text{if } m = 2 \\ 0 & \text{if } m = 1 \end{cases}$$

and the variance-covariance matrix of random effects  $\gamma_{1i}$ ,  $\varphi_{1i}$ ,  $\gamma_{2i}$  and  $\varphi_{2i}$  is

$$\text{variance-covariance matrix} = \begin{bmatrix} G_{A1} & G_{AB1} & 0 & 0 \\ G_{AB1} & G_{B1} & 0 & 0 \\ 0 & 0 & G_{A2} & G_{AB2} \\ 0 & 0 & G_{AB2} & G_{B2} \end{bmatrix}$$

As shown in the variance-covariance matrix above, the random effects are independent between the groups  $m=1$  and  $m=2$ .

For each chemical, the ratio of  $k_i$  values between two groups (e.g.,  $m=1$  for human and  $m=2$  for rat) was estimated using the estimated values of the parameters in the model as  $\frac{A_1/B_1}{A_2/B_2}$ , and the 95% confidence

interval (CI) of the estimated ratio was calculated based on the standard error that was approximately computed using the delta method<sup>14</sup> (all of these computations were done with the ESTIMATE statement in SAS PROC NL MIXED). When the ratios of the  $k_i$  values are from the human and rat groups, the chemical-specific ratios and their corresponding 95% CIs are used to evaluate the interspecies differences in AChE phosphorylation between two species. When the ratios of the  $k_i$  values are from the various human subpopulation groups (i.e., males and females; adults and infants; or Caucasian and non-

<sup>14</sup> Billingsley, P. (1986). *Probability and Measure*. 2nd ed. New York: John Wiley & Sons.

Caucasian), the chemical-specific ratios and their corresponding 95% CIs are used to evaluate the intraspecies differences in AChE phosphorylation between the human subpopulations.

For each data analysis, there were two final models: one assumed that both groups (e.g., human and rat species) had the *same* variance-covariance matrix for random effects  $\gamma_{mi}$  and  $\varphi_{mi}$  (model 2 in SAS code) and one assumed that both groups (e.g., human and rat) had *different* variance covariance matrices (model 3 in SAS code)<sup>15</sup>. Exponent selected the results of the final model that had the lower Akaike's Information Criterion (AIC) value. However, if the model with the lower AIC value had a SAS warning statement, but the alternative final model did not have a warning statement, Exponent selected the alternative model. In some cases, both models 2 and 3 had SAS warning statements (i.e., model fit issue), indicating the results of the selected final model should be interpreted with caution.

Exponent also conducted regression diagnostics for the selected final models of all data analyses. Based on a visual evaluation of diagnostic plots, some of the selected final models had severe outliers or a severe imbalance in the distribution of residuals in the *residuals vs. predicted values* plots (i.e., model fit is questionable). The results of the estimated ratios from these analyses should similarly be interpreted with caution.

For each chemical, phosphorylation rate constant ( $A$ ) and disassociation rate constant ( $B$ ) values of the human blood samples were calculated using the sample random value coefficients obtained from the selected final models for the human vs. rat analyses. The bimolecular rate constant of inhibition ( $k_i$ ) was then calculated as the ratio of these rate constants ( $A/B$ ). Exponent then provided normal Q-Q plots of the  $k_i$  values and the  $\ln(k_i)$  values to support rationale for assuming the  $k_i$  values were lognormally distributed for all chemicals. Coverage ratios (i.e., 90%-tile/GM, 95%-tile/GM, 97.5%-tile/GM, and 99%-tile/GM) were then calculated based on the assumption that  $k_i$  values are lognormally distributed. These chemical-specific coverage ratios were calculated only for the final selected models with the human and rat groups and were used to evaluate the variability AChE phosphorylation within humans (i.e., intraspecies variability of AChE phosphorylation).

### 3.3 Results

The sections below provide the summary results for the interspecies data (rat to human) and the intraspecies (within human) analyses. Values are reported for 16 of the 17 OPs tested because tribufos did not produce any AChE inhibition even at a final concentration of 1 mM; therefore,  $k_i$  values are not reported for tribufos.

#### 3.3.1 Interspecies data analyses

- There were 9 (of 16) chemicals for which we can interpret the results of estimated  $k_i$  ratios human/rat without any concern for severe outliers or model fit issues in the selected final models (Table 15). For all these 9 chemicals, both the 95% upper confidence limit (UCL) and lower confidence limit (LCL) for the ratios between the human estimated  $k_i$  values and the rat estimated  $k_i$  values are less than 3-fold from the point estimate (i.e.,  $UCL < 3$  and  $LCL > 0.33$ ).
- As shown in Table 15, there were 6 (of 16) chemicals for which the estimated ratios of human/rat should be interpreted with caution due to severe outlier/model fit concern. Except for naled with

---

<sup>15</sup>The same variance-covariance matrix model (model 2) assumes that the variations of random effects and the correlation between the random effects of the samples are the same or similar between human samples and animal samples. A different variance-covariance matrix model (model 3) assume the variation of random effects and the correlation between random effects of samples are different between human samples and rat samples.

estimated  $k_i$  ratio human/rat = 0.5 (95% CI = 0.35-0.66) and omethoate with estimated  $k_i$  ratio human/rat = 0.26 (95% CI = 0.09-0.43), the other 4 chemicals had similar  $k_i$  values between human and rat, and both the 95% UCL and LCL for the ratios between the human estimated  $k_i$  values and the rat estimated  $k_i$  values differ less than 2-fold (i.e., UCL < 2 and LCL > 0.5).

- There was 1 (of 16) chemical for which there was a SAS warning statement (model fit issue) in the selected final model (Table 15). The result of estimated  $k_i$  ratio human/rat from this model should be interpreted with caution.

**Table 15. Interspecies biomolecular rate constant ( $k_i$ ) ratios and associated 95% confidence intervals for 16 OPs**

Chemical	Estimated $k_i$ ratio human/rat	95% CI		Severe outliers/ model fit concern?	Warning statement/ model fit issue
		Lower	Upper		
Bensulide oxon	0.62*	0.42	0.81	No	No
Chlorethoxyfos	0.48*	0.39	0.57	No	No
DDVP	1.14	0.98	1.29	No	No
Ethoprop	2.06*	1.45	2.67	No	No
Fenamiphos	0.89	0.63	1.15	No	No
Methamidophos	0.86	0.64	1.09	No	No
Phorate oxon sulfoxide	0.84*	0.72	0.96	No	No
Phosmet oxon	1.07	0.83	1.3	No	No
Terbufos oxon sulfone	1.2	0.89	1.52	No	No
Terbufos oxon sulfoxide	1.35	1.2	1.5	No	Yes
Dicrotophos	0.76*	0.62	0.9	Yes	No
Malaoxon	0.96	0.53	1.39	Yes	No
Naled	0.50*	0.35	0.66	Yes	No
Omethoate	0.26*	0.09	0.43	Yes	No
Phorate oxon sulfone	0.88	0.5	1.25	Yes	No
Tebuprimphos oxon	0.44*	0.21	0.67	Yes	No

\*Statistically significantly different from 1 ( $p < 0.05$ )

### 3.3.2 Intraspecies data analyses

Three stratified intraspecies data analyses were separately conducted for each of 16 chemicals. These analyses are presented on pages 23-30 of MRID 51182301. However, there is low confidence in the results of these analyses, in part, due to the small sample sizes following stratification. EPA is concerned that results based on a limited number of samples would not be representative of the subpopulations being evaluated. There were 14 adult samples and 4 infant samples in the analysis of adult vs. infant (number of samples per group would be very small if the sample was further split into additional age groups). There were 8 male samples and 10 female samples in the analysis of male vs. female. There were 13 Caucasian samples and 5 non-Caucasian samples<sup>16</sup> in the analysis of Caucasian vs. non-Caucasian. Additionally,

<sup>16</sup> The Exponent analysis and nomenclature included African American (n=3) and Hispanic subjects (n=2) in the “non-Caucasian” category for this analysis.

several of the models for these analyses resulted in model outliers or had model convergence issues. Further exploratory analysis would be needed to address these model fit issues, which may in part be due to the limited number of samples.

Exponent estimated the coverage ratios of  $k_i$  distributions (Table 16), where the  $k_i$  values were estimated from the nonlinear mixed effects models using all available human samples. The human fixed effects and sample random of *Phosphorylation Constant*  $k_p$  (SAS coding as *HA* and *RHA*, respectively) and *Dissociation Constant*  $K_i$  (SAS coding as *HB* and *RHB*, respectively) obtained from the selected final models of human vs. rat analyses were used to calculate the  $k_i$  values ( $k_i = (HA + RHA)/(HB + RHB)$ ). The coverage ratios of the  $k_i$  distributions, that is, the ratio between various upper percentiles values and the geometric mean or GM (e.g., 90%-tile/GM, 95%-tile/GM, 97.5%-tile/GM, and 99%-tile/GM) were calculated based on the assumption that the  $k_i$  values were distributed lognormally. Among 9 chemicals with no issue or concern of model fit, all the  $k_i$  ratios of the 95%-tile/GM were less than 3. Of those 7 chemicals with model fit issues or model fit concerns in the analysis of human vs. rat, 2 chemicals (malaoxon and naled) had estimated  $k_i$  ratios of the 95%-tile/GM > 3.

Exponent also presented the results of an alternative coverage ratio analysis in MRID 51182301, which was based on fitting separate nonlinear models to each of the samples. For each chemical, the estimated nonlinear model for each human subject was considered completely independent and the model parameters were estimated completely independent of one another. However, EPA prefers the other mixed effect model approach over this alternative analysis because it better reflects the structure of the observed data and leverages the information contained in the entire set of chemical-specific data.

<b>Table 16. Intraspecies coverage ratios of biomolecular rate constant (<math>k_i</math>) distributions for 16 OPs assuming lognormal distribution</b>							
<b>Model fit issue/ model fit concern</b>	<b>Chemical</b>	<b>GM</b>	<b>GSD</b>	<b>90%-tile/ GM</b>	<b>95%-tile/ GM</b>	<b>97.5%-tile/ GM</b>	<b>99%-tile/ GM</b>
No	Bensulide oxon	3.7E+02	1.30	1.40	1.54	1.67	1.84
	Chlorethoxyfos	2.5E+07	1.17	1.23	1.30	1.37	1.45
	DDVP	5.1E+04	1.18	1.24	1.31	1.38	1.47
	Ethoprop	1.5E+03	1.18	1.24	1.32	1.39	1.48
	Fenamiphos	1.5E+02	1.23	1.31	1.41	1.50	1.62
	Methamidophos	1.1E+03	1.23	1.31	1.41	1.51	1.63
	Phorate oxon sulfoxide	5.2E+04	1.12	1.15	1.20	1.24	1.29
	Phosmet oxon	1.0E+05	1.21	1.27	1.36	1.45	1.55
	Terbufos oxon sulfone	3.8E+05	1.56	1.77	2.08	2.39	2.82
Model fit concern	Dicrotophos	5.3E+03	1.22	1.29	1.38	1.47	1.58
	Malaoxon	1.1E+05	2.49	3.22	4.48	5.97	8.33
	Naled	3.1E+06	2.16	2.69	3.56	4.54	6.02
	Omethoate	6.6E+02	1.71	1.99	2.41	2.86	3.48
	Phorate oxon sulfone	7.7E+04	1.39	1.53	1.72	1.91	2.16
	Tebupirimphos oxon	6.6E+05	1.73	2.02	2.47	2.94	3.59
Model fit issue	Terbufos oxon sulfoxide	1.1E+04	1.32	1.43	1.58	1.72	1.91

GM = geometric mean; GSD = geometric standard deviation

Exponent also conducted another data analysis to evaluate the variation of  $k_i$  between human samples (Table 17). The  $k_i$  data of 3 chemicals (naled, omethoate, and phosmet oxon) from 3 subsamples from each of 4 human samples were available. Note that these  $k_i$  values were obtained from the nonlinear regression analyses that Exponent separately performed on each  $k_i$  subsample data. To achieve the normality assumption, log-transformation was applied to the  $k_i$  data, then a mixed-effects model (SAS PROC MIXED) with human sample ID as random effect (i.e., experimental subject in the model) was used to analyze the log-transformed  $k_i$  data of each chemical. The intraclass correlation (ICC), between subject variance (this is actually the estimate of the variance of human  $\log(k_i)$ ), and within subject variance (this is actually the variance of  $\log(\text{experimental error})$ ) were obtained from the model. Using the estimated variance of human  $\log(k_i)$  obtained from the mixed-effect model, the coverage ratios of the  $k_i$  distributions<sup>17</sup> were calculated for each chemical and are presented in the following table.

As described in MRID 51182301, the results from this analysis are not consistent across the 3 chemicals. For naled, there was relatively large within subject variability with 84% of the total variability due to differences observed between the 3 replicate analyses of the human blood samples. Whereas, for phosmet oxon, only 3% of the total variability was due to differences in the 3 replicate analyses of the blood sample; 97% of the observed variability was due to the differences between human subjects.

<b>Chemical</b>	<b>GM</b>	<b>GSD<sub>total</sub></b>	<b>ICC</b>	<b>GSD<sup>a</sup><sub>between human</sub></b>	<b>90%-tile/ GM</b>	<b>95%-tile/ GM</b>	<b>97.5%-tile/ GM</b>	<b>99%-tile/ GM</b>
Naled	2.7E+06	2.1	0.84	1.86	2.4	3.0	3.7	4.8
Omethoate	6.9E+02	2.2	0.20	1.17	1.6	1.8	2.0	2.3
Phosmet Oxon	9.8E+04	1.3	0.027	1.01	1.1	1.1	1.1	1.1

<sup>a</sup> GSD<sub>between human</sub> is the geometric standard deviation of between human sample variation, calculated based on the formula  $ICC = \text{between human sample variance} / \text{Total Variance} = \text{between human sample variance} / (\text{between human sample variance} + \text{variance of experimental errors})$ . Specifically,  $GSD_{\text{between human}} = \exp(\sqrt{[\ln(GSD_{\text{total}}) * \ln(GSD_{\text{total}})] * ICC})$ .

### 3.4. Proposed Approach to the Human Health Risk Assessment

#### 3.4.1 Interspecies Pharmacodynamic DDEFs

The interspecies pharmacodynamic DDEFs ranged from 0.26 to 2.06. All DDEFs, except one value (ethoprop), were approximately 1 or less. There were no concerns for warning statements, severe outliers, and/or model fit issues for 9 of the chemicals tested. These included bensulide oxon, chlorethoxyfos, DDVP, ethoprop, fenamiphos, methamidophos, phorate oxon sulfoxide, phosmet oxon, and terbufos oxon sulfone. EPA’s preliminary proposal is to use the pharmacodynamic DDEFs presented in Table 15 for these 9 chemicals in lieu of the 3X default interspecies pharmacodynamic uncertainty factor for human health risk assessment, pending review by the SAP. Although some values are less than 1, the EPA’s DDEF guidance states that interspecies pharmacodynamic values “can be less than 1 if the data show humans are

<sup>17</sup> The ratio between various upper percentiles values and the geometric mean or GM (e.g., 90%-tile/GM, 95%-tile/GM, 97.5%-tile/GM, and 99%-tile/GM)

less sensitive than test species” but the decision of what value to use is considered policy and beyond the scope of the SAP.

For the remaining chemicals, there are concerns with using the DDEFs due to warning statements, potential outliers, and/or model fit issues. The OPP is soliciting feedback and advice from the SAP on the warning statements and model fit issues as part of Charge Question 6.

### 3.4.2 Intraspecies Pharmacodynamic DDEFs

Due to small sample sizes, the Agency does not have confidence in the stratified analyses (adult/infant, male/female, and Caucasian/non-Caucasian). There is more confidence in the coverage ratios (the ratio between various upper percentiles values and the geometric mean) that were calculated using all of the available human samples; however, it is still questionable whether the sample size (n=18) is sufficient and representative for these analyses. The same 9 chemicals listed in Section 3.4.1 for interspecies extrapolation with no concerns for warning statements, severe outliers, and/or model fit issues were also found to have no concerns for intraspecies pharmacodynamic DDEFs. Therefore, EPA’s preliminary proposal is to use the intraspecies pharmacodynamic DDEFs generated for bensulide oxon, chlorethoxyfos, DDVP, ethoprop, fenamiphos, methamidophos, phorate oxon sulfoxide, phosmet oxon, and terbufos oxon sulfone (Table 16) in lieu of the 3X default intraspecies pharmacodynamic uncertainty factor, pending review by the SAP. The percentile used is a policy decision that will be made by the OPP when applying these data for human health risk assessments and beyond the scope of the SAP.

For the remaining chemicals, there are concerns with using the DDEFs due to warning statements, potential outliers, and/or model fit issues. The OPP is soliciting feedback and advice from the SAP on the warning statements and model fit issues as part of Charge Question 6. Additionally, in Charge Question 8, EPA is seeking advice from the SAP on the utility of the analyses performed for a subset of OPs to evaluate experimental and intrinsic variability and whether additional data should be generated for other OPs.

## 4.0 Summary & Next Steps

The EPA’s OPP is actively engaged in numerous activities related to developing and implementing NAMs, which reduce reliance on laboratory animal studies and have the ability to provide human relevant information that may be challenging to obtain from *in vivo* studies. This Issue Paper presented two approaches related to these efforts. The first approach presented work completed by ORD thus far to develop a battery of NAMs for evaluating DNT, using OPs as a case study. The MEA NFA and HCI assays were found to be robust and reproducible for fit-for-purpose application to identify putative DNT-related bioactivity, with reasonable assay performance controls established. Differential activity was observed across the 27 OPs tested in the MEA NFA and HCI assay suite. These assays evaluate the majority of the critical processes of neurodevelopment (as presented in Table 3 and Figure 2) and represent a significant advancement toward developing a NAM battery for fit-for-purpose DNT evaluation.

IVIVE approaches with HTTK models were utilized to approximate NAM-derived AEDs, which were greater than or in some cases approximating BMD/BMDL values based on AChE inhibition. Predictions from PBPK-PD models were found to be higher than HTTK predictions; however, these comparisons

were limited to the 3 chemicals where PBPK-PD models could be applied. The IVIVE approach used is driven by the chemical-specific data available. Due to the lack of rat clearance data for many of the OPs tested, the use of human data in the rat model (i.e., huRat) was evaluated. Predictions were not found to be uniformly higher or lower when using human rather than rat data. Therefore, for chemicals that are lacking rat clearance data and have huRat AEDs approximating the BMD/BMDL values, it is unknown whether the predictions are over- or underestimates. Pending the recommendations from the SAP and the overall weight of evidence evaluation of DNT potential based on the currently available data, the OPP may determine that the rat *in vitro* clearance data for particular OPs are needed to improve AED vs. BMD/BMDL comparisons and may reach out to registrants in the future to generate these data.

Additional assays are currently under development by EFSA-funded researchers that evaluate processes not covered by the MEA NFA and HCI assays (i.e., migration and differentiation). The relative contribution of these assays for a fit-for-purpose NAM battery will be considered once data are available. Additionally, any OP data from these assays will be considered in combination with the results of the MEA NFA and HCI assays as part of an overall weight of evidence evaluation of the DNT potential for individual OPs.

The second approach presented *in vitro* data generated by academia on behalf of pesticide registrants to calculate pharmacodynamic DDEF values for 16 OPs in accordance with the EPA's 2014 guidance on DDEFs. For both interspecies and intraspecies DDEFs, there were no statistical issues identified for 9 chemicals. EPA's preliminary proposal is to use the DDEFs calculated for these 9 chemicals (bensulide oxon, chlorethoxyfos, DDVP, ethoprop, fenamiphos, methamidophos, phorate oxon sulfoxide, phosmet oxon, and terbufos oxon sulfone) in lieu of the default pharmacodynamic uncertainty factors, pending review by the SAP. For the remaining chemicals, SAP recommendations regarding warning statements and model fit issues will be considered before determining the potential use of the calculated DDEFs. Ultimately, the selection and application of DDEFs for each chemical is a policy decision that will be made by EPA.

## 5.0 References

- Alloisio, S., et al., 2016. Microelectrode array (MEA) platform as a sensitive tool to detect and evaluate *Ostreopsis cf. ovata* toxicity. *Harmful Algae*. 55, 230-237.
- Alloisio, S., et al., 2015. Multiparametric characterisation of neuronal network activity for in vitro agrochemical neurotoxicity assessment. *Neurotoxicology*. 48, 152-65.
- Aschner, M., et al., 2017. Reference compounds for alternative test methods to indicate developmental neurotoxicity (DNT) potential of chemicals: example lists and criteria for their selection and use. *ALTEX*. 34, 49-74.
- Atchison, W. D., 1988. Effects of neurotoxicants on synaptic transmission: lessons learned from electrophysiological studies. *Neurotoxicol Teratol*. 10, 393-416.
- Bal-Price, A., et al., 2018a. Recommendation on test readiness criteria for new approach methods in toxicology: Exemplified for developmental neurotoxicity. *ALTEX*. 35, 306-352.
- Bal-Price, A., et al., 2018b. Strategies to improve the regulatory assessment of developmental neurotoxicity (DNT) using in vitro methods. *Toxicol Appl Pharmacol*. 354, 7-18.
- Ball, K. R., et al., 2017. A multivariate extension of mutual information for growing neural networks. *Neural Netw*. 95, 29-43.
- Barenys, M., et al., 2017. Epigallocatechin gallate (EGCG) inhibits adhesion and migration of neural progenitor cells in vitro. *Arch Toxicol*. 91, 827-837.
- Barone, S., Jr., et al., 2000. Vulnerable processes of nervous system development: a review of markers and methods. *Neurotoxicology*. 21, 15-36.
- Baskar, M. K., Murthy, P. B., 2018. Acute in vitro neurotoxicity of some pyrethroids using microelectrode arrays. *Toxicol In Vitro*. 47, 165-177.
- Baumann, J., et al., 2016. Comparative human and rat neurospheres reveal species differences in chemical effects on neurodevelopmental key events. *Arch Toxicol*. 90, 1415-27.
- Bell, S. M., et al., 2018. In vitro to in vivo extrapolation for high throughput prioritization and decision making. *Toxicol In Vitro*. 47, 213-227.
- Bertrand, R., et al., 1994. Induction of a common pathway of apoptosis by staurosporine. *Exp Cell Res*. 211, 314-21.
- Bisio, M., et al., 2014. Emergence of bursting activity in connected neuronal sub-populations. *PLoS One*. 9, e107400.
- Bondy, S. C., Campbell, A., 2005. Developmental neurotoxicology. *J Neurosci Res*. 81, 605-12.
- Bradley, J. A., et al., 2018. In Vitro Screening for Seizure Liability Using Microelectrode Array Technology. *Toxicol Sci*. 163, 240-253.
- Bradley, J. A., Strock, C. J., 2019. Screening for Neurotoxicity with Microelectrode Array. *Curr Protoc Toxicol*. 79, e67.
- Brown, J. P., et al., 2016. Editor's Highlight: Evaluation of a Microelectrode Array-Based Assay for Neural Network Ontogeny Using Training Set Chemicals. *Toxicol Sci*. 154, 126-139.
- Carr, R. L., et al., 2017. Decreased anxiety in juvenile rats following exposure to low levels of chlorpyrifos during development. *Neurotoxicology*. 59, 183-190.
- Carr, R. L., et al., 2014. Low level chlorpyrifos exposure increases anandamide accumulation in juvenile rat brain in the absence of brain cholinesterase inhibition. *Neurotoxicology*. 43, 82-89.
- Charlesworth, P., et al., 2015. Quantitative differences in developmental profiles of spontaneous activity in cortical and hippocampal cultures. *Neural Dev*. 10, 1.
- Cheng, C., et al., 2017. Highly Expandable Human iPS Cell-Derived Neural Progenitor Cells (NPC) and Neurons for Central Nervous System Disease Modeling and High-Throughput Screening. *Curr Protoc Hum Genet*. 92, 21 8 1-21 8 21.

- Cheng, C. H., Kuchta, R. D., 1993. DNA polymerase epsilon: aphidicolin inhibition and the relationship between polymerase and exonuclease activity. *Biochemistry*. 32, 8568-74.
- Chiappalone, M., et al., 2006. Dissociated cortical networks show spontaneously correlated activity patterns during in vitro development. *Brain Res*. 1093, 41-53.
- Coecke, S., et al., 2013. Toxicokinetics as a key to the integrated toxicity risk assessment based primarily on non-animal approaches. *Toxicol In Vitro*. 27, 1570-7.
- Colombi, I., et al., 2013. Effects of antiepileptic drugs on hippocampal neurons coupled to micro-electrode arrays. *Front Neuroeng*. 6, 10.
- Cotterill, E., et al., 2016. Characterization of Early Cortical Neural Network Development in Multiwell Microelectrode Array Plates. *J Biomol Screen*. 21, 510-9.
- Crofton, K. M., et al., 2011. Developmental neurotoxicity testing: recommendations for developing alternative methods for the screening and prioritization of chemicals. *ALTEX*. 28, 9-15.
- Culbreth, M. E., et al., 2012. Comparison of chemical-induced changes in proliferation and apoptosis in human and mouse neuroprogenitor cells. *Neurotoxicology*. 33, 1499-1510.
- de Esch, C., et al., 2012. Zebrafish as potential model for developmental neurotoxicity testing: a mini review. *Neurotoxicol Teratol*. 34, 545-53.
- de Groot, M. W., et al., 2014. Characterization of calcium responses and electrical activity in differentiating mouse neural progenitor cells in vitro. *Toxicol Sci*. 137, 428-35.
- de Souza Anselmo, C., et al., 2018. Zebrafish (*Danio rerio*): A valuable tool for predicting the metabolism of xenobiotics in humans? *Comp Biochem Physiol C Toxicol Pharmacol*. 212, 34-46.
- DeGroot, D. E., et al., 2018. mRNA transfection retrofits cell-based assays with xenobiotic metabolism. *J Pharmacol Toxicol Methods*. 92, 77-94.
- Delp, J., et al., 2018. A high-throughput approach to identify specific neurotoxicants/ developmental toxicants in human neuronal cell function assays. *ALTEX*. 35, 235-253.
- Dingemans, M. M., et al., 2016. Chronic 14-day exposure to insecticides or methylmercury modulates neuronal activity in primary rat cortical cultures. *Neurotoxicology*. 57, 194-202.
- Druwe, I., et al., 2016. Comparison of Human Induced Pluripotent Stem Cell-Derived Neurons and Rat Primary Cortical Neurons as In Vitro Models of Neurite Outgrowth. *Applied In Vitro Toxicology*. 2, 26-26.
- Druwe, I., et al., 2015. Sensitivity of neuroprogenitor cells to chemical-induced apoptosis using a multiplexed assay suitable for high-throughput screening. *Toxicology*. 333, 14-24.
- Duarte, D. J., et al., 2017. In vitro neurotoxic hazard characterization of different tricresyl phosphate (TCP) isomers and mixtures. *Neurotoxicology*. 59, 222-230.
- Einarson, A., et al., 2000. Prospective, controlled, multicentre study of loperamide in pregnancy. *Can J Gastroenterol*. 14, 185-7.
- Feng, G., Kaplowitz, N., 2002. Mechanism of staurosporine-induced apoptosis in murine hepatocytes. *Am J Physiol Gastrointest Liver Physiol*. 282, G825-34.
- Filer, D. L., et al., 2017. tcpl: the ToxCast pipeline for high-throughput screening data. *Bioinformatics*. 33, 618-620.
- Frank, C. L., et al., 2017. From the Cover: Developmental Neurotoxicants Disrupt Activity in Cortical Networks on Microelectrode Arrays: Results of Screening 86 Compounds During Neural Network Formation. *Toxicol Sci*. 160, 121-135.
- Frega, M., et al., 2012. Cortical cultures coupled to micro-electrode arrays: a novel approach to perform in vitro excitotoxicity testing. *Neurotoxicol Teratol*. 34, 116-27.
- Frega, M., et al., 2017. Rapid Neuronal Differentiation of Induced Pluripotent Stem Cells for Measuring Network Activity on Micro-electrode Arrays. *J Vis Exp*.

- Fritsche, E., et al., 2018. Consensus statement on the need for innovation, transition and implementation of developmental neurotoxicity (DNT) testing for regulatory purposes. *Toxicol Appl Pharmacol.* 354, 3-6.
- Gao, Y., et al., 2004. Rational design and characterization of a Rac GTPase-specific small molecule inhibitor. *Proc Natl Acad Sci U S A.* 101, 7618-23.
- Goldstone, J. V., et al., 2010. Identification and developmental expression of the full complement of Cytochrome P450 genes in Zebrafish. *BMC Genomics.* 11, 643.
- Gong, X. W., et al., 2012. Loperamide, an antidiarrhea drug, has antitumor activity by inducing cell apoptosis. *Pharmacol Res.* 65, 372-8.
- Gordon, J. A., 1991. Use of vanadate as protein-phosphotyrosine phosphatase inhibitor. *Methods Enzymol.* 201, 477-82.
- Gorshkov, K., et al., 2018. Astrocytes as targets for drug discovery. *Drug Discov Today.* 23, 673-680.
- Goulian, M., et al., 1990. Two forms of DNA polymerase delta from mouse cells. Purification and properties. *J Biol Chem.* 265, 16402-11.
- Gramowski-Voss, A., et al., 2015. Enhancement of Cortical Network Activity in vitro and Promotion of GABAergic Neurogenesis by Stimulation with an Electromagnetic Field with a 150 MHz Carrier Wave Pulsed with an Alternating 10 and 16 Hz Modulation. *Front Neurol.* 6, 158.
- Grandjean, P., Landrigan, P. J., 2006. Developmental neurotoxicity of industrial chemicals. *Lancet.* 368, 2167-78.
- Grandjean, P., Landrigan, P. J., 2014. Neurobehavioural effects of developmental toxicity. *Lancet Neurol.* 13, 330-8.
- Guo, S., 2009. Using zebrafish to assess the impact of drugs on neural development and function. *Expert Opin Drug Discov.* 4, 715-726.
- Hampson, D. R., Manalo, J. L., 1998. The activation of glutamate receptors by kainic acid and domoic acid. *Nat Toxins.* 6, 153-8.
- Harrill, J. A., et al., 2015. Ontogeny of biochemical, morphological and functional parameters of synaptogenesis in primary cultures of rat hippocampal and cortical neurons. *Mol Brain.* 8, 10.
- Harrill, J. A., et al., 2018. Testing for developmental neurotoxicity using a battery of in vitro assays for key cellular events in neurodevelopment. *Toxicol Appl Pharmacol.* 354, 24-39.
- Harrill, J. A., et al., 2010. Quantitative assessment of neurite outgrowth in human embryonic stem cell-derived hN2 cells using automated high-content image analysis. *Neurotoxicology.* 31, 277-90.
- Harrill, J. A., et al., 2011a. Comparative sensitivity of human and rat neural cultures to chemical-induced inhibition of neurite outgrowth. *Toxicol Appl Pharmacol.* 256, 268-80.
- Harrill, J. A., et al., 2013. Use of high content image analyses to detect chemical-mediated effects on neurite sub-populations in primary rat cortical neurons. *Neurotoxicology.* 34, 61-73.
- Harrill, J. A., et al., 2011b. Use of high content image analysis to detect chemical-induced changes in synaptogenesis in vitro. *Toxicol In Vitro.* 25, 368-87.
- Hers, I., et al., 1999. The protein kinase C inhibitors bisindolylmaleimide I (GF 109203x) and IX (Ro 31-8220) are potent inhibitors of glycogen synthase kinase-3 activity. *FEBS Lett.* 460, 433-6.
- Hogberg, H. T., et al., 2011. Application of micro-electrode arrays (MEAs) as an emerging technology for developmental neurotoxicity: evaluation of domoic acid-induced effects in primary cultures of rat cortical neurons. *Neurotoxicology.* 32, 158-68.
- Hondebrink, L., et al., 2016. Neurotoxicity screening of (illicit) drugs using novel methods for analysis of microelectrode array (MEA) recordings. *Neurotoxicology.* 55, 1-9.
- Huang, T., et al., 2016. Microelectrode Array-evaluation of Neurotoxic Effects of Magnesium as an Implantable Biomaterial. *J Mater Sci Technol.* 32, 89-96.
- Jenkins, F., Barr, D., Chen, J., Cronin, M., Delcos, K., Ehrich, M., Georgopoulos, P., Hayton, W., Jett, D., Mcdonald, P., Mcmanaman, J., Murphy, C., Potter, T.L., Schlenk, D., SAP Minutes No. 2014-03

- for FIFRA meeting held July 29-31, 2014. A set of scientific issues being considered by the Environmental Protection Agency regarding new high throughput methods to estimate chemical exposure. . Environmental Protection Agency. , 2014.
- Johnstone, A. F., et al., 2010. Microelectrode arrays: a physiologically based neurotoxicity testing platform for the 21st century. *Neurotoxicology*. 31, 331-50.
- Karaman, M. W., et al., 2008. A quantitative analysis of kinase inhibitor selectivity. *Nat Biotechnol*. 26, 127-32.
- Kasteel, E. E., Westerink, R. H., 2017. Comparison of the acute inhibitory effects of Tetrodotoxin (TTX) in rat and human neuronal networks for risk assessment purposes. *Toxicol Lett*. 270, 12-16.
- Kepiro, M., et al., 2018. High Content, Phenotypic Assays and Screens for Compounds Modulating Cellular Processes in Primary Neurons. *Methods Enzymol*. 610, 219-250.
- Khazipov, R., Luhmann, H. J., 2006. Early patterns of electrical activity in the developing cerebral cortex of humans and rodents. *Trends Neurosci*. 29, 414-418.
- Kirshenboim, N., et al., 2004. Lithium-mediated phosphorylation of glycogen synthase kinase-3beta involves PI3 kinase-dependent activation of protein kinase C-alpha. *J Mol Neurosci*. 24, 237-45.
- Kohno, R., et al., 2006. Sphere formation of ocular epithelial cells in the ciliary body is a reprogramming system for neural differentiation. *Brain Res*. 1093, 54-70.
- Kokel, D., et al., 2010. Rapid behavior-based identification of neuroactive small molecules in the zebrafish. *Nat Chem Biol*. 6, 231-237.
- Kondo, Y., et al., 2017. Temporal relation between neural activity and neurite pruning on a numerical model and a microchannel device with micro electrode array. *Biochem Biophys Res Commun*. 486, 539-544.
- Kosnik, M. B., et al., 2020. Concentration-response evaluation of ToxCast compounds for multivariate activity patterns of neural network function. *Arch Toxicol*. 94, 469-484.
- Kota, K. P., et al., 2012. High content image based analysis identifies cell cycle inhibitors as regulators of Ebola virus infection. *Viruses*. 4, 1865-77.
- Kreir, M., et al., 2018. Do in vitro assays in rat primary neurons predict drug-induced seizure liability in humans? *Toxicol Appl Pharmacol*. 346, 45-57.
- Krokan, H., et al., 1981. Aphidicolin inhibits DNA synthesis by DNA polymerase alpha and isolated nuclei by a similar mechanism. *Nucleic Acids Res*. 9, 4709-19.
- Krug, A. K., et al., 2013. Evaluation of a human neurite growth assay as specific screen for developmental neurotoxicants. *Arch Toxicol*. 87, 2215-31.
- Lantz, S. R., et al., 2014. Glufosinate binds N-methyl-D-aspartate receptors and increases neuronal network activity in vitro. *Neurotoxicology*. 45, 38-47.
- Lein, P., et al., 2007. Meeting report: alternatives for developmental neurotoxicity testing. *Environ Health Perspect*. 115, 764-8.
- Lenk, K., et al., 2016. Simulation of developing human neuronal cell networks. *Biomed Eng Online*. 15, 105.
- McConnell, E. R., et al., 2012. Evaluation of multi-well microelectrode arrays for neurotoxicity screening using a chemical training set. *Neurotoxicology*. 33, 1048-57.
- Meyer, D. A., et al., 2008. Pyrethroid modulation of spontaneous neuronal excitability and neurotransmission in hippocampal neurons in culture. *Neurotoxicology*. 29, 213-25.
- Mochizuki, Y., et al., 2016. Similarity in Neuronal Firing Regimes across Mammalian Species. *J Neurosci*. 36, 5736-47.
- Mohana Krishnan, B., Prakhya, B. M., 2016. In vitro evaluation of pyrethroid-mediated changes on neuronal burst parameters using microelectrode arrays. *Neurotoxicology*. 57, 270-281.
- Mundy, W. R., et al., 2015. Expanding the test set: Chemicals with potential to disrupt mammalian brain development. *Neurotoxicol Teratol*. 52, 25-35.

- Mundy, W. R., et al., 2010. Neuronal models for evaluation of proliferation in vitro using high content screening. *Toxicology*. 270, 121-30.
- Nishimura, Y., et al., 2015. Zebrafish as a systems toxicology model for developmental neurotoxicity testing. *Congenit Anom (Kyoto)*. 55, 1-16.
- Nyffeler, J., et al., 2017. Combination of multiple neural crest migration assays to identify environmental toxicants from a proof-of-concept chemical library. *Arch Toxicol*. 91, 3613-3632.
- O'Brien, W. T., et al., 2011. Glycogen synthase kinase-3 is essential for  $\beta$ -arrestin-2 complex formation and lithium-sensitive behaviors in mice. *The Journal of Clinical Investigation*. 121, 3756-3762.
- OECD, Test No. 426: Developmental Neurotoxicity Study, OECD Guidelines for the Testing of Chemicals. In: 4, S., (Ed.). OECD Publishing, Paris, 2007.
- Pancrazio, J. J., et al., 2014. Botulinum toxin suppression of CNS network activity in vitro. *J Toxicol*. 2014, 732913.
- Pearce, R. G., et al., 2017a. Evaluation and calibration of high-throughput predictions of chemical distribution to tissues. *J Pharmacokinet Pharmacodyn*. 44, 549-565.
- Pearce, R. G., et al., 2017b. htk: R Package for High-Throughput Toxicokinetics. *J Stat Softw*. 79, 1-26.
- Peterson, R. T., et al., 2008. Use of non-mammalian alternative models for neurotoxicological study. *Neurotoxicology*. 29, 546-55.
- Pugazhenthii, S., et al., 1996. Inhibition of a Src homology 2 domain containing protein tyrosine phosphatase by vanadate in the primary culture of hepatocytes. *Arch Biochem Biophys*. 335, 273-82.
- Reif, D. M., et al., 2016. High-throughput characterization of chemical-associated embryonic behavioral changes predicts teratogenic outcomes. *Arch Toxicol*. 90, 1459-70.
- Robinette, B. L., et al., 2011. In vitro assessment of developmental neurotoxicity: use of microelectrode arrays to measure functional changes in neuronal network ontogeny. *Front Neuroeng*. 4, 1.
- Rodier, P. M., 1995. Developing brain as a target of toxicity. *Environ Health Perspect*. 103 Suppl 6, 73-6.
- Rolls, A., et al., 2007. Toll-like receptors modulate adult hippocampal neurogenesis. *Nat Cell Biol*. 9, 1081-8.
- Rotroff, D. M., et al., 2010. Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol Sci*. 117, 348-58.
- Ruegg, U. T., Burgess, G. M., 1989. Staurosporine, K-252 and UCN-01: potent but nonspecific inhibitors of protein kinases. *Trends Pharmacol Sci*. 10, 218-20.
- Ryves, W. J., Harwood, A. J., 2001. Lithium inhibits glycogen synthase kinase-3 by competition for magnesium. *Biochem Biophys Res Commun*. 280, 720-5.
- Sachana, M., et al., 2019. International Regulatory and Scientific Effort for Improved Developmental Neurotoxicity Testing. *Toxicol Sci*. 167, 45-57.
- Salinas, E., Sejnowski, T. J., 2001. Correlated neuronal activity and the flow of neural information. *Nat Rev Neurosci*. 2, 539-50.
- Scelfo, B., et al., 2012. Application of multielectrode array (MEA) chips for the evaluation of mixtures neurotoxicity. *Toxicology*. 299, 172-83.
- Schroeder, O. H., et al., Spike train data analysis of substance-specific network activity: application to functional screening in preclinical development. 6th international meeting on substrate-integrated microelectrodes, July 8–11, 2008, Reutlingen, Germany, 2008.
- Shafer, T. J., 2019. Application of Microelectrode Array Approaches to Neurotoxicity Testing and Screening. *Adv Neurobiol*. 22, 275-297.
- Shafer, T. J., et al., 2019. Evaluation of Chemical Effects on Network Formation in Cortical Neurons Grown on Microelectrode Arrays. *Toxicol Sci*. 169, 436-455.
- Shafer, T. J., et al., 2008. Complete inhibition of spontaneous activity in neuronal networks in vitro by deltamethrin and permethrin. *Neurotoxicology*. 29, 203-12.

- Sharma, P., et al., 2012. High-throughput screening in primary neurons. *Methods Enzymol.* 506, 331-60.
- Stambolic, V., et al., 1996. Lithium inhibits glycogen synthase kinase-3 activity and mimics wingless signalling in intact cells. *Curr Biol.* 6, 1664-8.
- Strickland, J. D., et al., 2018. Screening the ToxCast phase II libraries for alterations in network function using cortical neurons grown on multi-well microelectrode array (mwMEA) plates. *Arch Toxicol.* 92, 487-500.
- Toullec, D., et al., 1991. The bisindolylmaleimide GF 109203X is a potent and selective inhibitor of protein kinase C. *J Biol Chem.* 266, 15771-81.
- Tukker, A. M., et al., 2020. Towards animal-free neurotoxicity screening: Applicability of hiPSC-derived neuronal models for in vitro seizure liability assessment. *ALTEX.* 37, 121-135.
- Uhlhaas, P. J., et al., 2009. Neural synchrony in cortical networks: history, concept and current status. *Front Integr Neurosci.* 3, 17.
- Uhlhaas, P. J., Singer, W., 2006. Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron.* 52, 155-68.
- Upton, R. N., 2007. Cerebral uptake of drugs in humans. *Clin Exp Pharmacol Physiol.* 34, 695-701.
- USEPA, Health Effects Test Guidelines OCSPP 870.6300 Developmental Neurotoxicity Study. 1998a.
- USEPA, U.S. Environmental Protection Agency Guidelines for neurotoxicity risk assessment. EPA/630/R-95/001F. USEPA, Washington, DC, 1998b.
- USEPA, US Environmental Protection Agency Policy on a Common Mechanism of Action: The Organophosphate Pesticides. Vol. 64. Federal Register, Washington, D.C., 1999, pp. 5795-5799.
- Vandenbossche, J., et al., 2010. Loperamide and P-glycoprotein inhibition: assessment of the clinical relevance. *J Pharm Pharmacol.* 62, 401-12.
- Walton, N. M., et al., 2006. Derivation and large-scale expansion of multipotent astroglial neural progenitors from adult human brain. *Development.* 133, 3671-81.
- Wambaugh, J. F., et al., 2018. Evaluating In Vitro-In Vivo Extrapolation of Toxicokinetics. *Toxicol Sci.* 163, 152-169.
- Wambaugh, J. F., et al., 2015. Toxicokinetic Triage for Environmental Chemicals. *Toxicol Sci.* 147, 55-67.
- Wetmore, B. A., 2015. Quantitative in vitro-to-in vivo extrapolation in a high-throughput environment. *Toxicology.* 332, 94-101.
- Wetmore, B. A., et al., 2012. Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol Sci.* 125, 157-74.
- Wright, L. S., et al., 2006. Human progenitor cells isolated from the developing cortex undergo decreased neurogenesis and eventual senescence following expansion in vitro. *Exp Cell Res.* 312, 2107-20.
- Wu, L. S., Li, J., 2018. High-Content Imaging Phenotypic Screen for Neurogenesis Using Primary Neural Progenitor Cells. *Methods Mol Biol.* 1787, 101-113.
- Zhang, G., et al., 2017. A New Statistical Approach to Characterize Chemical-Elicited Behavioral Effects in High-Throughput Studies Using Zebrafish. *PLoS One.* 12, e0169408.