SCIENTIFIC OPINION

EFSA Journal

ADOPTED: 21 April 2021

doi: 10.2903/j.efsa.2021.6599

Development of Integrated Approaches to Testing and Assessment (IATA) case studies on developmental neurotoxicity (DNT) risk assessment

EFSA Panel on Plant Protection Products and their Residues (EFSA PPR Panel), Antonio Hernández-Jerez, Paulien Adriaanse, Annette Aldrich, Philippe Berny, Tamara Coja, Sabine Duquesne, Andreas Focks, Marina Marinovich, Maurice Millet, Olavi Pelkonen, Silvia Pieper, Aaldrik Tiktak, Christopher Topping, Anneli Widenfalk, Martin Wilks, Gerrit Wolterink, Kevin Crofton, Susanne Hougaard Bennekou, Martin Paparella and Ioanna Tzoulaki

Abstract

The EFSA Panel on Plant Protection Products and their Residues (PPR) has developed, as a self-task mandate (EFSA-Q-2019-00100), two adverse outcome pathway (AOP)-informed integrated approach to testing and assessment (IATA) case studies to answer a developmental neurotoxicity (DNT) hazard identification and characterisation problem formulation that could support the regulatory decisions for the pesticide active substances deltamethrin and flufenacet. The IATA were developed to assess the applicability of the DNT *in vitro* testing battery (IVB), designed to explore fundamental neurodevelopmental processes, in the regulatory risk assessment of pesticides. For this purpose, an evidence-based-approach methodology was applied: 1) systematic literature review and critical appraisal of all the evidence i.e. human observational studies, *in vivo* data from rodent models and new approach methodologies (NAMs, i.e. *in vitro* studies including high-throughput testing from IVB and zebrafish studies from the literature) for both case studies; 2) a quantitative uncertainty analysis of all the evidence using expert knowledge elicitation (EKE) and a probabilistic approach; 3) integration of all the evidence using the AOP conceptual framework. This stepwise approach resulted in the postulation of an evidence-based AOP network for one of the case studies. A probabilistic quantification of the weight of evidence (WoE) using Bayesian network analysis allowed the assessment and the quantification of the uncertainty in the postulated AOP. The approach taken allowed conclusions to be drawn with an acceptable level of certainty in DNT hazard identification and characterisation of deltamethrin and that flufenacet is not a developmental neurotoxicant, supporting the relevance of the mechanistic understanding. The case studies show the applicability of the DNT-IVB for hazard identification and characterisation and illustrate the usefulness of an AOP-informed IATA for regulatory decision making. The overall activity led to improved interpretation of human data by providing a plausible mechanistic link to adverse outcomes, which would support their contextualisation in the risk assessment process. This Scientific Opinion allows the PPR Panel to draft several recommendations for the implementation of the AOP-informed IATA methodology and of the DNT-IVB in the regulatory risk assessment of pesticides.

Keywords: IATA, developmental neurotoxicity, DNT, AOP, Deltamethrin, Flufenacet, *in vitro*, evidence-based, uncertainty analysis

Requestor: PPR Panel

Question number: EFSA-Q-2019-00100

Correspondence: pesticides.ppr@efsa.europa.eu

Panel members: Adriaanse Paulien, Aldrich Annette, Berny Philippe, Coja Tamara, Duquesne Sabine, Focks Andreas, Hernández-Jerez Antonio, Marinovich Marina, Millet Maurice, Pelkonen Olavi, Pieper Silvia, Tiktak Aaldrik, Topping Christopher, Widenfalk Anneli, Wilks Martin and Wolterink Gerrit.

Declarations of interest: The declarations of interest of all scientific experts active in EFSA's work are available at https://ess.efsa.europa.eu/doi/doiweb/doisearch.

Acknowledgements: The Panel wishes to thank the following for the support provided to this scientific output: Andrea Terron, Iris Mangas, Martina Panzarea, Luca Martinenghi, Laura Martino, Federica Barrucci, Elisa Aiassa and Irene Munoz Guajardo from EFSA staff, the hearing experts Ellen Fritsche and Marcel Leist and the observers Anna Price and Verena Haudek-Prinz.

Suggested citation: EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), Hernández-Jerez A, Adriaanse P, Aldrich A, Berny P, Coja T, Duquesne S, Focks A, Marinovich M, Millet M, Pelkonen O, Pieper S, Tiktak A, Topping C, Widenfalk A, Wilks M, Wolterink G, Crofton K, Hougaard Bennekou S, Paparella M and Tzoulaki I, 2021. Scientific Opinion on the development of Integrated Approaches to Testing and Assessment (IATA) case studies on developmental neurotoxicity (DNT) risk assessment. EFSA Journal 2021;19(6):6599, 63 pp. https://doi.org/10.2903/j.efsa.2021.6599

ISSN: 1831-4732

The EFSA Journal is a publication of the European Food Safety Authority, a European agency funded by the European Union.

Summary

There is a growing concern that chemicals can contribute to the increase in neurodevelopmental disorders (NDDs), as the developing nervous system may be more sensitive to certain hazardous chemicals. However, the current OECD test guidelines (TG) 426 (developmental neurotoxicity study, DNT) and TG 443 (extended one-generation reproductive toxicity study) (OECD, 2007, 2018), which were designed for the characterisation of potential chemical-induced DNT hazards, have significant limitations. A new framework is therefore needed for the assessment of chemicals on their potential to disrupt brain development.

In 2017, the European Food Safety Authority (EFSA) launched a procurement for the implementation and interpretation of an *in vitro* testing battery (IVB), consisting of DNT *in vitro* assays with high test readiness, for the assessment of DNT. The final report of this procurement (Masjosthusmann et al., 2020) represents a fundamental step in the DNT assessment paradigm as it includes DNT-IVB results for many chemicals from different domains, the assessment of the DNT-IVB performance, the contextualisation of this mechanistic data and case studies. Next, the mechanistic information provided by this IVB should be integrated for the postulation of new adverse outcome pathways (AOPs). In addition, a collaborative effort between EFSA, OECD, US EPA and Danish EPA, with engagement of academic and industry researchers, has provided a roadmap that will result in an OECD guidance document for *in vitro* DNT testing. The guidance will focus on use and interpretation of the IVB which will be incorporated in an integrated approach to testing and assessment (IATA) framework for the regulatory use of *in vitro* DNT data.

In this Scientific Opinion, the IATA was used since it represents an optimal tool for the inclusion of a combination of methods to address a defined question in a specific regulatory context. The EFSA Panel on Plant Protection Products and their Residues (PPR) has therefore conducted IATA case studies for DNT hazard identification and characterisation for pesticide active substances (therefore not covering all the steps of risk assessment), which was then expressed as an acceptable level of uncertainties, to be used later in the process of risk assessment. The main purpose of developing the IATA case studies in this Scientific Opinion was to show the applicability of an IVB for DNT in the context of the European pesticide regulations (EU) 283/2013 and 1107/2009 (European Commission, 2009, 2013). Two pesticide active substances, tested in the current DNT-IVB, were selected for the case studies: deltamethrin, a type II pyrethroid with a well characterised neurotoxic effect and a neurotoxic pesticidal mode of action, and flufenacet, a herbicide with a mode of action not related to a neurotoxic effect but for which a regulatory DNT study (i.e. OECD TG 426; OECD, 2007) was available. The AOP conceptual framework was used to organise and integrate all the evidence and contextualise the mechanistic evidence with the apical toxicity endpoints. Exposure data were used to support the dose and time concordance for the empirical support of the key event relationships (KERs) in the postulated AOP informing IATA.

A protocol was therefore developed based on EFSA Scientific Committee (SC) (2020), which included detailed information on the strategy and methods for the systematic review process used. The systematic literature review was conducted for three lines of evidence (human observational studies (HOS), experimental *in vivo* data in rodents (including data from pesticide dossiers) and *in vitro* studies (including high-throughput testing from the DNT-IVB and zebrafish studies)) for deltamethrin and flufenacet. Then the evidence was appraised for risk of bias (RoB) using a tailored version of the OHAT-NTP RoB tool.

To assure that the assessment provided reliable information for decision making, special considerations were given to the uncertainty analysis, with the identification, characterisation and evaluation of uncertainties and limitations. The uncertainty analysis was performed for each line of evidence to support conclusions on the hazard identification and characterisation using a predefined set of questions. Therefore, the systematic review, including the critical appraisal of the internal validity, of all the evidence and the uncertainty analysis were considered and accounted when integrating the evidence. When a causal association between the exposure to the substance and any of the relevant endpoints was identified, the lowest concentration/dose triggering the effect was elicited. A threshold of 66% (indicating that the probability of an event to occur is twice as possible as not) was used as the minimum subjective probability leading to the conclusion of a causal association. Only molecular initiating events (MIEs), key events (KEs) and adverse outcomes (AOs) with an estimated probability $\geq$ 66% were considered for the next step of the AOP network postulation. For the second step, the hazard characterisation (i.e. identification of the lowest concentration/dose), the uncertainty was expressed quantitatively using either ranges of possible values including the true

concentration/dose with 100% probability, or full probability distributions over plausible values as suggested in the EFSA guidance on uncertainty analysis (EFSA, 2018). The approach taken for the uncertainty analysis was via an expert knowledge elicitation (EKE).

This stepwise approach was therefore culminating in the postulation of the evidence-based AOP network for deltamethrin with a probabilistic quantitative estimation of the weight of evidence (WoE) using a Bayesian network analysis. The stressor-based approach was selected because the intention was to use the AOP conceptual framework to establish a causal relationship between the prototypical chemical deltamethrin exposure and the AO. For flufenacet, no DNT hazard was identified, hence, no AOP was postulated.

Related to HOS, the level of probability for a causal association of deltamethrin exposure with all DNT outcomes was lower than 66%, principally due to 'probably' or 'definitively high' RoB for exposure measurements in the studies. However, the available human evidence linking prenatal exposure to deltamethrin with NDDs)is supported by experimental data from animal models of some neurodevelopmental outcomes and is consistent with the biology of NDDs and with the available evidence structured in the AOP conceptual framework, which indicates that a plausible mechanistic link between exposure to deltamethrin and NDDs exists. However, the current (low) doses of exposure during pregnancy may not be enough for triggering the whole process leading to NDDs and additional co-exposure to other chemicals, environmental factors or maternal lifestyles may be needed, apart from the role of a specific genetic background.

The string of adjacent KEs from the MIE1 to KE4 (i.e. binding to voltage-gated sodium channels (VGSC); KE1: disruption of sodium channels-gate kinetic; KE2: disruption of action potential; KE3: disruption of axon terminal depolarisation; changes in neurotransmitter release; KE4: altered neuronal network function; AO: impairment behavioural function (sensory–motor reflex and learning)), represents the most robust pathway in the network mainly because of the considerable amount of biological knowledge and information available for deltamethrin. However, because of the conservatism of this pathway and the broad knowledge on pyrethroid toxicity, the PPR Panel considers that this postulated sequence of KEs would also be relevant for other pyrethroids.

Overall, the development of the AOP network identified two major uncertainties. The first one is related to the network itself as potentially important KEs may be missing. This is due to the decision to include only data and KEs that have been studied for the stressor deltamethrin. The second is due to a knowledge gap that drives uncertainty in the last two KERs: the lack of empirical data that provide correlative and/or causal relationships between disturbed neuronal network function (KE4) and/or hypomyelination (as a result of KE5) and altered behavioural function (AO). These uncertainties can be reduced by including additional KEs and KERs using the biological knowledge rather than using a stressor specific approach.

The quantitative assessment of the WoE for this putative AOP network and for the strength of the KER was performed using a Bayesian network approach. This was based on available data and expert knowledge for deltamethrin. The approach also provided a quantitative estimate for the mechanistic knowledge within the putative AOP network. According to this approach, the marginal probability for each MIE/KE/AO in any AOP string in the AOP network was always greater than 0.5 (50%), indicating that it is more likely than not that the MIE, KEs and AO would occur upon deltamethrin exposure at the triggering concentration/dose.

The full process culminated in the IATA. The two case studies allowed for a comparative evaluation of the DNT-IVB while assessing the impact of the mechanistic understanding in two very different scenarios, supporting the validity of the approach taken. Therefore, the IATA addresses the Terms of Reference proposed by the EFSA PPR Panel.

The IATA case study for deltamethrin found apparent equivocal results between *in vivo* experimental studies from the open literature and the OECD TG 426 study, although the study design and the doses administered differed across studies. Moreover, epidemiological data showed a potential DNT concern for the pyrethroid chemical class. This led to the conclusion that the integration of mechanistic data, using DNT-IVB data relevant for a normal neurodevelopment, reduces the uncertainty for DNT hazard identification and characterisation of deltamethrin.

The IATA case study for flufenacet indicated that based on all the evidence from the literature and the DNT-IVB data, and after accounting for the uncertainties, no DNT hazard can be identified, supporting the conclusion that flufenacet is not a developmental neurotoxicant. This conclusion is in line with lack of adverse DNT outcomes in the available regulatory *in vivo* study.

Overall, the case studies included in this Scientific Opinion show the applicability of the DNT-IVB for hazard identification and characterisation and, in the context of these case studies, its specificity and

sensitivity. Furthermore, the case study illustrates the usefulness of the postulation of an AOP network and of the probabilistic quantification of the WoE with the potential of aiding regulatory decision making. The overall process and the mechanistic understanding also increased the ability of interpreting HOS by providing a plausible mechanistic link to a human-relevant AO, thus supporting the contextualisation of these studies in the future risk assessment process.

Table of contents

1. Introduction

There is a societal concern that chemicals can contribute to the increase in neurodevelopmental disorders (NDDs) and this could be explained by the fact that the developing nervous system may be more sensitive to chemical exposures. Characterisation of potential chemically induced developmental neurotoxicity (DNT) is considered for risk assessment purposes by many regulatory sectors. These potential health effects following exposure to environmental chemicals led to the development of testing guidelines by the US EPA in 1986, with refinement by OECD in 2007.

The current OECD test guideline (TG) 426 on DNT (OECD, 2007), which was founded on the US EPA OPPTS 870.6300 from 1998 (Makris et al., 2009), requires assessing the impact of prenatal and early postnatal exposure to chemicals (via dams during gestation and lactation) on the development of several neurological endpoints in the offspring. However, despite the socio-economical relevance of DNT, there has been limited use of these test guidelines. This is the result of several factors including time and resource implications, the need to limit animal use, the lack of regulatory requirements for DNT testing, and historically limited assessments of potential data variability and human predictivity (Bal-Price and Fritsche, 2018; Paparella et al., 2020). Therefore, hazard identification and actions to reduce exposure to potential DNT chemicals is a priority in chemical risk assessment and risk management. To reach this goal, a cost-efficient testing strategy based on a reliable *in vitro* testing battery should be developed. Although representing a huge challenge in risk assessment, available data and methodologies support the development of a predictive model able to respond to different regulatory-based problem formulations (Terron and Bennekou, 2018).

This situation has resulted in a consensus of scientific stakeholders from regulatory agencies, academia and industry that a new framework for assessment of chemicals with the potential to disrupt brain development is needed. Therefore, beyond the OECD tests, work has been carried out by individual research groups aimed at identifying the DNT effects of chemicals (Fritsche et al., 2015). This led to the development of several *in vitro* models for neurodevelopmental testing, in which methods using human-derived cell-based test systems represent a significant progress.

A collaborative effort between EFSA, OECD, US EPA and DK-EPA, supported by academic and industry researchers (Fritsche et al., 2017), provided a roadmap that is currently culminating in the production of an OECD guidance document (GD) for *in vitro* DNT testing. The GD will advise on use and interpretation of an *in vitro* testing battery consisting of DNT *in vitro* assays with high test readiness (Sachana et al., 2019; Bal-Price et al., 2018). The GD will strategically incorporate the *in vitro* testing battery (IVB) (Masjosthusmann et al., 2020) in an integrated approach to testing and assessment (IATA) framework for facilitation of regulatory use of *in vitro* DNT data. Therefore, targeted DNT *in vitro* testing should be guided by a problem formulation approach based on regulatory needs using the IATA framework.

Therefore, this battery of assays has a wide variety of potential uses. Because the battery is far less expensive and has much higher throughput compared with guideline DNT studies it can be used for screening and prioritisation of large chemical libraries for which DNT hazard is unknown. Similarly, it can be used to help prioritise decisions about which of a smaller set of compounds might be the best candidates for additional focused investigations. Because the battery is based on disruption of important neurodevelopmental processes, it also provides information on potential mechanisms or processes that a compound is disrupting that cannot be obtained from *in vivo* guideline studies. Such information could also be used to provide scientific rationale for more focused hypothesis-driven *in vivo* studies when necessary. When there are data from *in vivo* studies that indicate a weak or borderline equivocal effect on a neurodevelopmental endpoint, data from the battery could be helpful to interpret those data based on a weight-of-evidence approach. Finally, the battery will be useful for the generation of DNT AOPs and will have many potential uses when implemented in the context of IATAs.

In addition, this Scientific Opinion values the scientific, ethical, and legal reasons of using human-relevant non-animal methods in the regulatory process of pesticides active substances approval (EU Directive 2010/63/EU,[1] Regulation (EC) 1107/2009[2]; Regulation (EU) 283/2013[3]).

[1] Directive 2010/63/EU on the protection of animals used for scientific purposes. Article 1(a) and Article 4. Available online: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:276:0033:0079:en:PDF

[2] Regulation (EC) 1107/2009 concerning the placing of plant protection products on the market. Article 62. Available online: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:309:0001:0050:EN:PDF

[3] Regulation (EU) 283/2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009. Annex I point 5. Available online: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:093:0001:0084:EN:PDF

Several case studies will be included in the OECD GD. As a partner of this project, EFSA developed two case studies based on single chemical risk assessment using the IATA framework to answer the defined question on potential DNT effects considering the level of uncertainty associated with the decision context, i.e. chemically-induced DNT hazard characterisation.

1.1. Background and Terms of Reference as provided by the requestor

Hazards associated with DNT are of particular significance for the risk assessment of pesticides (Regulation No 1107/2009 and Regulation 283/2013, European Commission, 2009, 2013) resulting in an increased margin of safety when a concern on DNT exist. In its Scientific Opinion on the DNT potential of acetamiprid and imidacloprid (EFSA PPR Panel, 2013), the PPR Panel recommended the development of an integrated *in vitro* neurotoxicity testing strategy complementary to *in vivo* assays included into OECD TG 426 and TG 443 to screen the DNT potential of pesticides and other chemicals entering food chains. In that Scientific Opinion, the PPR Panel comprehensively discussed the issues associated with the current *in vivo* DNT testing strategy. These studies are complex and very resource demanding and raised concerns in terms of animal welfare, excessive variability in results, difficulties in the interpretation, and lack of understanding on how to link animal's endpoints to complex human diseases (Paparella et al., 2020). As a follow-up activity, the EFSA Peer Review (PREV) Unit awarded a contract to prepare a literature review on *in vitro* and alternative DNT testing methods (Fritsche et al., 2015). Subsequently, an EFSA workshop was held in November 2016 in collaboration with the OECD and US EPA (Fritsche, 2017). The outcome of the workshop was presented and discussed at the PPR Panel in November 2016 and at the SC in December 2016. The PPR Panel and the SC supported the elaboration of an action plan by EFSA with a view of developing a testing strategy on this multisector issue, which was included as a key activity in the EFSA Strategy 2020.

Therefore, the PREV Unit submitted a Standard Project Submission Form (SPSF) at the OECD for the preparation of guidance on the application and interpretation of *in vitro* DNT assays for testing and assessment, led by EFSA, US EPA and Danish EPA with the participation of interested OECD Member States. To provide an appropriate scientific background and facilitate the preparation of this guidance, EFSA launched a procurement for the conduction of experimental work with the overall goal to accelerate the development and use of *in vitro* test methods capable of cost and time-efficient testing of chemicals for their potential to disrupt the development of the nervous system. The experimental work has been completed and the report was published in October 2020 as an EFSA supporting publication (Masjosthusmann et al., 2020). The report includes the outcome of the experimental work with generation of data using relevant chemicals for validation purposes, and the design and employment of data analysis tools. The report also includes a proposal for data interpretation and use guidance, descriptions of possible application domains and case studies.

In the context of the development of the OECD guidance, EFSA proposed to develop specific case studies dealing with the risk assessment of pesticides using the IATA framework. IATAs are pragmatic, science-based approaches for chemical hazard characterisation that rely on an integrated analysis of existing information coupled with the generation of new information using testing strategies based on new approach methodologies (NAMs). IATAs follow an iterative process to answer a defined question in a specific regulatory context, considering the acceptable level of uncertainty associated with the decision context. The principles of an IATA are to include in the assessment multiple data from various information sources, including adverse outcome pathways (AOP) developed for DNT. AOP-informed IATA will be used in the DNT risk assessment in which the DNT *in vitro* testing battery will address the mechanistic understanding of the effects observed in experimental toxicology and epidemiological studies, along with the newly generated information based on a testing strategy (DNT *in vitro* testing battery, DNT-IVB), and the key element of the uncertainty analysis. For the Scientific Opinion, the PPR Panel will develop IATA case studies using a DNT risk assessment-based problem formulation with all available information on defined pesticide active substances. The PPR Panel will specifically:

1) Propose a DNT risk assessment problem formulation that could support the regulatory decisions (increased Margin of Safety -MoS- or decision on candidate for substitution) under Regulation 1107/2009 regarding the approval of active substances in the future.
2) Propose pesticide active substances for which an IATA case study should be developed. In selecting those active substances, the PPR Panel will consider: the completeness of the available database, the concern for DNT based on the qualitative toxicological profile of each active substance and the chemical class.

3) Develop an iterative AOP-informed IATA case study for each selected pesticide active substance taking into account the level of uncertainties associated with the available database.
4) Integrate the outcome of the *in vitro* testing battery for DNT in the AOP-informed IATA case study and provide a new uncertainty analysis to guide on the use and interpretation of the *in vitro* DNT testing battery.
5) Discuss the outcome of the case studies and contextualise in the conclusion of the Scientific Opinion whether the proposed testing battery is fit-for-purpose and its limitations.

Only a limited number of AOPs are currently endorsed by OECD for DNT and there are data gaps in mechanistic information. As the proposed DNT testing battery was designed to explore fundamental processes in neurodevelopment, the outcome of this battery will be critical for a new AOP postulation, eventually leading to a testing battery informed AOP followed by an AOP-informed IATA (Masjosthusmann et al., 2020).

1.2. Interpretation of the Terms of Reference

The PPR Panel made use of the ToRs to develop the following problem formulations:

1) How certain are we that substance A is a developmental neurotoxicant in humans, based on the data collected, appraised, synthesised and integrated using an operational protocol discussed and agreed by the PPR Panel and in line with the IATA framework?
2) To what extent do the results of the *in vitro* testing battery (Masjosthusmann et al., 2020) on substance A influence the level of uncertainty as assessed in point 1?
3) Same question as 1, but on substance B.
4) Same question as 2, but on substance B.

The PPR Panel considered two main aspects in the interpretation of the terms ToRs.

The first aspect considers the scientific basis on which the DNT *in vitro* testing battery (DNT-IVB) was developed. The use of DNT-IVB assays that measure disruption of key neurodevelopmental processes at cellular level assumes that the impairment of these processes will impact multiple downstream events and all the disruptive effects can be integrated in a consistent manner (Radio and Mundy, 2008).However, although this assumption is generally accepted, more evidence is needed to conclude that the DNT-IVB can replace the use of *in vivo* evidence available, including, but not limited to, the DNT OECD test guideline studies (i.e. TG426 and TG443; OECD, 2007; 2018, respectively) for derivation of all hazard-based decisions. In the context of this Scientific Opinion, evidence of adverse outcomes (AOs) generated in *in vivo* experimental studies will be considered for the hazard identification and characterisation. However, the PPR Panel agreed that evidence from human observational studies (HOS) will also be accounted for in the process and in the overall assessment to link effects that may be seen in humans and determine if a biological plausible link exists between human exposure and the AOP generated by the PPR Panel.

The second aspect is dealing with the IATA methodology. IATA is a framework that allows for the integration of all available data for use in chemical regulatory assessments (OECD, 2016c). IATAs are:

> 'pragmatic, science-based approaches for chemical hazard characterisation that rely on an integrated analysis of existing information, with optional use of the adverse outcome pathway (AOP) framework, coupled with the generation of new information if necessary. IATAs follow an iterative approach to answer a defined question in a specific regulatory context, considering the acceptable level of uncertainty associated with the decision context' (OECD, 2016c).

In line with the ToRs, the PPR Panel opted for the development of an AOP-informed IATA as a selected tool to conduct a DNT hazard characterisation for substance A and B which will be expressed as an acceptable level of uncertainties in the context of IATA, to be used later in the process of risk assessment. Therefore, the Panel interpreted the ToRs as dealing with a DNT hazard identification and characterisation, not covering all the steps of risk assessment. As the outcome of the DNT-IVB was used from the beginning of the process, the impact of the DNT-IVB is discussed as part of the outcome of the IATA.

2. Data and methodologies

2.1. Data

Data related to three lines of evidence (*in vitro* and *in vivo* experimental data, and human observational data) have been collected and appraised:

1) I DNT-IVB experimental work has been completed, reported and published in October 2020 as an EFSA supporting publication (Masjosthusmann et al., 2020). The report includes the outcome of the experimental work and a proposal for data interpretation. The experimental work entails testing of 119 chemicals, including the pesticide active substances deltamethrin and flufenacet.
2) *In vivo* experimental studies for regulatory purpose (OECD TG 426) are available for both deltamethrin and flufenacet.
3) Additional data were deemed necessary to address the ToR, specifically to develop an AOP-informed IATA for each of the selected pesticide active substance. A systematic literature review was therefore conducted for this purpose for the human observational data, *in vivo* data in rodents and *in vitro* data (including high-throughput testing and zebrafish studies).

2.2. Methodologies

2.2.1. Chemical selection for the case studies

The first consideration in selecting the chemicals to be used in the case studies was the main scope of the Scientific Opinion, i.e. the impact of the DNT-IVB in the process of hazard characterisation with an acceptable level of uncertainties. To address this question, availability of *in vivo* experimental studies conducted with pesticide active substances was considered necessary. The second requisite was based on the desire to test the DNT-IVB for a known neurotoxic substance and for a non-neurotoxic substance. Therefore, two chemicals were selected: deltamethrin, a type II pyrethroid with a well characterised neurotoxic effect in mammals and a neurotoxic pesticidal mode of action and where some evidence of DNT exists; and flufenacet, an herbicide with a mode of action not related to a neurotoxic effect, where no evidence of DNT exists. For both substances an OECD TG 426 study was available. In addition, both active substances were tested in the current DNT-IVB.

2.2.2. Development of the IATA case studies

The IATA case studies are fully reported in ***Annexes N*** and ***O*** for deltamethrin and flufenacet, respectively.

Within an IATA, data from various information sources are evaluated and integrated to draw conclusions on the hazard and/or risk of chemicals (OECD GD 329; OECD, 2020). Important features of IATA are the need to explicitly set out the problem formulation and context of use, since these will determine the acceptable level of uncertainty, the choice of methods (building blocks), the approach to evidence integration (OECD, 2016a,b) and the application of the WoE in an iterative process until a conclusion is reached.

In this Scientific Opinion, the IATA was used as a tool to address the ToR of the mandate since it represents an optimal tool for the inclusion of a combination of methods, including NAMs, to address a defined question in a specific regulatory context. Therefore, the IATA was applied first to address the hazard characterisation of the two selected substances, and then to assess the impact of the DNT-IVB in addressing the regulatory question. To understand the relationship between what is tested in the DNT-IVB and the apical toxicity endpoint being predicted, the use of the AOP framework was deemed necessary. The AOP is therefore the key element to inform the IATA and is needed to characterise the biological and toxicological relevance of the novel methods used for predicting the adverse effect. The same framework was used for the integration of the additional evidence retrieved from the systematic review which will benefit from an integrated approach.

Figure 1 shows the workflow applied in the IATA developed to address the ToR and includes the components of the IATA which is explained in detail in Annexes *N* and *O* for deltamethrin and flufenacet, respectively. Data were systematically collected and appraised for *in vitro*, *in vivo* and human lines of evidence. The biological knowledge and the empirical evidence for selected endpoints were used to postulate the KEs to be included in the AOP. For deltamethrin, the AOP-informed IATA was therefore the approach selected to draw conclusions based on a quantitative WoE. For flufenacet, the same approach was taken; however, no hazard was identified up to Step 4 and the workflow stopped at the Step 4 and no AOP was postulated.

To assure that the assessment provides reliable information for decision making, special considerations have been given in these IATA case studies to the process for evidence retrieval and appraisal and to the uncertainty analysis, with the identification, characterisation and evaluation of uncertainties and limitations. This was carried out first, by planning the methods and ensuring that the

questions and methods were defined in a protocol for the assessment (EFSA Scientific Committee, 2020); second, by conducting a systematic data collection and appraisal using Critical Appraisal Tools; and third, by conducting an uncertainty analysis of each part, which was based on predefined domains and guiding questions and led to a transparent reporting of the uncertainty expressed quantitatively using probability. This stepwise approach has culminated in the development of an evidence-based, stressor-based, AOP with a probabilistic quantitative estimation of the WoE using a Bayesian network (BN) approach. Importantly, exposure considerations were used to contextualise dose and time concordance for the empirical support of the KERs in the postulated AOP informing IATA. This approach represents a step forward in terms of the methodological approach in delivering a sound, transparent and accessible scientific advice in support to decision making.

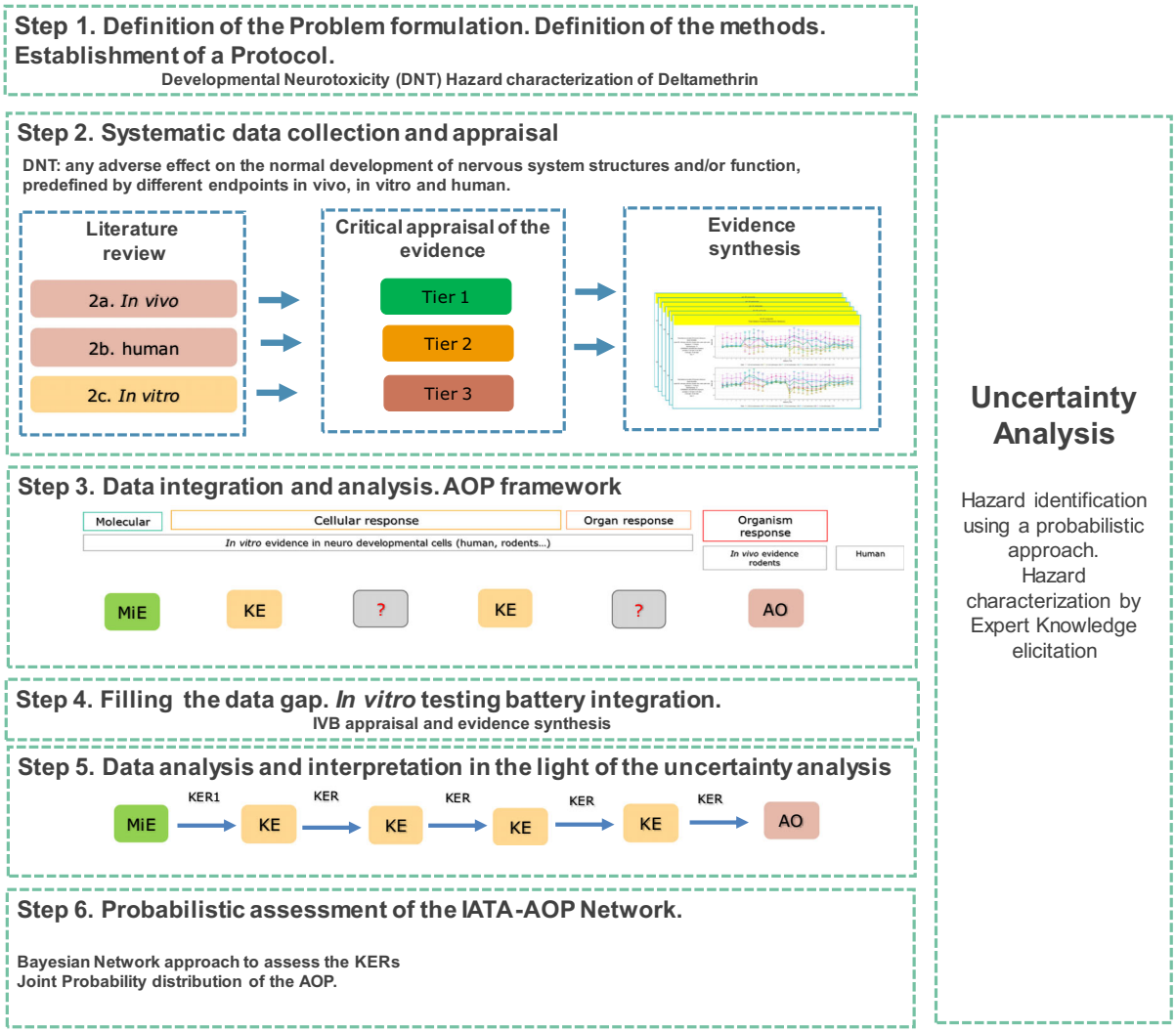


Figure 1: Workflow used in the development of the IATA

2.2.3. Evidence-based approach

Protocol

A predefined protocol for the systematic review process was developed based on EFSA (2020) and is reported in Annex A. The protocol includes in detail the strategy and methods for the systematic review process, specifically:

- translation of the mandate into sub-questions.
- predefinition of DNT endpoints.
- search string to retrieve the studies for *in vitro* (including high throughput), *in vivo* and human evidence and full list of literature databases to be used.

- studies eligibility criteria and screening for relevance.
- data model and procedures for the data extraction.
- Critical appraisal tools (CATs) and procedures for assessing the risk of bias (RoB), including the rationale used for each line of evidence for DNT.

The protocol however only partially includes the methods for evidence synthesis, integration and uncertainty analysis, since the novelty of the approach made it was difficult to anticipate the most appropriate methodology for addressing quantitatively the AOP framework.

Systematic literature review

The literature searches were conducted using three electronic bibliographic databases (PubMed, Web of Science, Toxnet) and three resources indexing PhD theses (DART, EBSCO and PQDT). The exact time period considered is included in Table 8 of Annex A Protocol and it was of more than 15 years and until July 2020, and then updated on 23 of November 2020 (for deltamethrin) and on 7 December 2020 (for flufenacet) by an information specialist. Search strings are described in the protocol (Annex A). Terms for the exposure were combined with relevant terms for DNT outcomes (human and *in vivo* studies) or methods (*in vitro* studies) and a specific search string was designed to identify studies applying high-throughput methods to evaluate potential DNT without terms of exposure (Annex A). The DNT outcomes were predefined by a series of toxicological *in vivo* and *in vitro* endpoints, and NDDs (based on The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; APS, 2013) for HOS and categorised in endpoint categories translated into keywords for the searches (see Annex A).

Critical appraisal of the evidence

RoB for the eligible *in vivo* studies in rodents (experimental toxicological studies) and humans (observational studies) was appraised using tailored versions of the OHAT-NTP RoB tool (NTP, 2015). For the *in vitro* studies, a modified tool developed by OHAT-NTP for a specific project (NTP, 2016) was adapted. CATs were defined upfront and are described in the protocol (Annex A). Table 1 shows the different questions and domains appraised for *in vivo*, *in vitro* and human lines of evidence and the three key questions selected for this specific assessment.

Endpoints measured in the studies were classified as being of low (Tier 1), moderate (Tier 2) or high (Tier 3) RoB. These tiers were derived weighing the appraisal from the individual RoB domains where some were identified as key for the overall appraisal. RoB was appraised by endpoint in those studies in which the diverse endpoints measured used different methodology and therefore may have different RoB.

Table 1: Critical appraisal tool questions for the RoB analysis and key questions selected for each line of evidence (see Annex A, Protocol for further details)

Selection Bias	**Human**	**In vivo**	**In vitro**
1. Was administered dose or exposure level adequately randomized?	–	Key Q	Yes
2. Was allocation to study groups adequately concealed?	–	Yes	–
3. Did selection of study participants result in appropriate comparison groups	Yes	–	Yes
Confounding Bias			
4. Did the study design or analysis account for important confounding and modifying variables?	Key Q	–	–
Performance Bias			
5. Were experimental conditions identical across study groups?	–	Yes	Yes
6. Were the research personnel blinded to the study group during the study?	–	Yes	Yes
Attrition/Exclusion Bias			
7. Were outcome data complete without attrition or exclusion from analysis?	Yes	Yes	Yes
Detection Bias			
8. Can we be confident in the exposure characterisation?	Key Q	Key Q	Key Q
9. Can we be confident in the outcome assessment?	Key Q	Key Q	Key Q

Selective Reporting Bias			
10. Were all measured outcomes reported?	Yes	Yes	Yes
Other Sources of Bias			
11. Were there other potential threats to internal validity?	Statistics	Systematic tox	Cytoxic
12. Were there other potential threats to internal validity?			N replic

2.2.4. Uncertainty analysis and expert knowledge elicitation

The uncertainty analysis and the expert knowledge elicitation (EKE) methodologies are reported in Annex B (statistical report). According to EFSA recommendations, uncertainty should be addressed as part of the scientific assessment questions using quantitative approaches whenever possible (EFSA and EBTC, 2018; EFSA Scientific Committee, 2018). An uncertainty analysis was therefore performed for each line of evidence and hierarchical level (specific endpoints and categories) to support conclusions on the hazard identification and characterisation questions. The final purpose was to screen the evidence and identify molecular initiating events (MIEs), key events (KEs) and adverse outcomes (AOs) to be included in the putative AOP network.

All data were mapped in specific endpoints and endpoint categories (obtained from the systematic review, the DNT-IVB and the TG 426 *in vivo* study) and this was used to postulate an AOP network. In the AOP, evidence was classified as MIEs, KEs or AOs based on their categorisation as molecular, cellular, organ, organism or population responses (see Figure 10). However, this was only possible for deltamethrin as all data retrieved or obtained for flufenacet were negative and mapping was not therefore possible.

The uncertainty analysis was conducted in two steps (Steps 1 and 2) to identify and characterise the DNT hazard in the body of evidence by providing answers to specific assessment questions and expressing the uncertainty in a probabilistic way; therefore, probability is used here to express the uncertainty of the working group experts on these questions.

In Step 1, the hazard identification uncertainty analysis was conducted by addressing the questions shown in **Table 2**. MIEs, KEs and AOs for which the probability of a causal association was judged by the working group experts consensus to be 66% or above were included in an initial list that was then used in Step 2 and to postulate the resulting AOP network (see Section 2.2.5). The threshold of 66% was considered sufficiently conservative since it means that it is judged at least twice as probable that deltamethrin is activating the MIEs, KEs and AOs rather than not activating. Considering the intrinsic limitations of the HOS, the working group decided not to address the hazard characterisation due to the overall high RoB affecting the body of evidence. However, for this line of evidence the certainty on the hazard identification question was assessed more precisely using an approximate probability scale instead of a bounded probability (see Table 2). Indeed, in this case, scientific consensus means that all the working group experts agreed with the overall conclusion. The 66% identified the lower bound (minimum) of the probability attached to the conclusion, i.e. that a causal association exists between exposure to deltamethrin and an adverse outcome. Therefore, it is reflecting the uncertainty identified in the evidence when concluding on the causal association and is not related with the percentage of working group experts agreeing on this conclusion. Indeed, all working group experts agreed on the causal association with some uncertainty and they believed that it is at least as double as probable that the causal association occurs than not occurs.

Table 2: Assessment questions for the uncertainty analysis for hazard identification, answered by each working group expert individually, discussed and all working group experts agreed with the overall conclusion (see Annex B, Statistical Report for further details)

Line of evidence	Question 1. Hazard identification	Expression of the uncertainty (probability)
***In vitro* experimental studies**	Does exposure to deltamethrin/flufenacet trigger the specific endpoint/KE as measured in acute and developmental protocol (wash-out yes/no) (assuming a monotonic concentration–response relationship) in *in vitro* studies (EFSA DNT-IVB and open literature) carried out in human and/or rat and/or mouse neuro cells in development?	**Bounded probability** No: Prob $<$ 0.66 Yes: Prob $\geq$ 0.66

Line of evidence	Question 1. Hazard identification	Expression of the uncertainty (probability)
***In vivo* experimental studies**	Does exposure to deltamethrin/flufenacet affect this specific endpoint/endpoint category/adverse outcome in a dose–response relationship in **experimental animal studies** exposed during pregnancy and/or postnatal until weaning (maximum up to 21 days postnatal for rats and mice)?	**Bounded probability** No: Prob < 0.66 Yes: Prob ≥ 0.66
Human observational studies	What is the probability that an association between human individual exposure to deltamethrin/flufenacet in uterus (mothers might have been exposed via dietary and non-dietary sources) and the specific endpoint/adverse outcome occurs?	Approximate probability(*): [0–10) % [10–33) % [33–50) % [50–66) % [66–100] % (*) A round or squared parenthesis indicates that the extreme is excluded or included, respectively

To address the assessment questions and the related level of certainty, the working group experts were asked to assess the available body of evidence and a predefined set of factors/domains that could represent either justification for inconsistencies in the evidence or sources of uncertainty because of their inadequateness for answering the questions. The predefined factors/domains were tailored by line of evidence and their assessment facilitated by guiding questions. The list of factors considered included (but was not limited to): (a) for *in vivo* studies (including zebrafish): species, strain, sexes, exposure duration and stage, maternal and or systemic toxicity, RoB, route of administration, imprecision; (b) for *in vitro* studies: exposure conditions, test system used, RoB, effect measurements and analysis, imprecision; (c) for human studies: children's age at the time of outcome assessment, timing of maternal urine collection for biomonitoring, exposure characterisation (type of metabolites measured and expression of exposure and limit of detection (LOD)), methods for outcome assessment, RoB, confounding factors, statistical analysis (Annex E for deltamethrin and Annex F for flufenacet).

All potential sources of uncertainty identified were tabulated in Annexes K and L for deltamethrin and flufenacet evidence, respectively. The working group experts were asked to independently assess each of these factors starting with the lower hierarchical levels (specific endpoints) and progress through the higher levels (endpoint categories and MIEs, KEs or AOs). The working group experts were instructed to provide synthetic answers (Yes/No/Not Relevant) accompanied by a narrative explanation and rationale for the answer. The results of these independent assessments were then discussed in a meeting and the results of the discussion were summarised in the uncertainty analysis tables. The working group experts were asked to address question 1 in Table 2 and judgement was made on whether exposure to deltamethrin is causing that specific endpoint, endpoint category and/or MIE, KE or AO. Although a similar process was applied for flufenacet, no AO, MIE or KEs were identified as triggered by flufenacet. This was also true following testing of flufenacet in the DNT-IVB. Therefore, no AOP was postulated allowing movement to a conclusive step.

Step 2. The hazard characterisation and the related uncertainty analysis was conducted by providing answers to the questions shown in Table 3. This was performed for the MIE/KE/AO that were judged as being triggered by deltamethrin in Step 1 with a probability (level of certainty) higher than 66%.

In the cases for which data were obtained from different studies, the uncertainty affecting the body of evidence used for identifying the lowest concentration/dose triggering the KE/AO were derived using a semi-formal EKE (EFSA, 2014; EFSA and EBTC, 2018; EFSA Scientific Committee, 2018). First, the working group experts were requested to provide an individual estimate of a range (in cases where the probability distribution was agreed upfront as uniform) or a full probability distribution. When the uncertainty at individual level was elicited in the form probability distribution, the Roulette method was used to elicit their knowledge (EFSA, 2014' O'Hagan, 2019). Then a consensus range or probability distribution was achieved based on discussion among the experts (see Annex B for further details).

Table 3: Assessment questions for the uncertainty analysis for hazard characterisation, answered by each working group expert individually, and then the probability distribution was integrated, discussed and all working group experts agreed collectively with the overall conclusion (see Annex B, Statistical Report for further details)

Line of evidence	Question 2. Hazard characterisation	Expression of the uncertainty (expert knowledge elicitation)
***In vitro* experimental studies**	What is the lowest concentration at which the exposure to deltamethrin triggers the MIE/KE (assuming a monotonic concentration–response relationship)?	Range (uniform distribution) or full probability distribution (using the Roulette method) for the lowest concentration/dose
***In vivo* experimental studies**	What is the lowest dose at which deltamethrin causes the AO in a dose–response relationship in rodents exposed during pregnancy and/or postnatal until weaning?	

2.2.5. Putative adverse outcome pathway development and weight of evidence (WoE)

The putative AOP network is fully described in Annex M.

An AOP is a logical sequence of KEs triggered by chemical exposure and occurring at different levels of biological organisation, i.e. molecular, cellular, organ, whole organism or population. These KEs are causally linked to the AO under consideration and they are measurable. The AOP is anchored at one end by a MIE, which represents the direct interaction of a chemical with a molecular biological target, and at the other end by an AO at the organism or population level. In this Scientific Opinion, the AO is anchored to DNT endpoints observed in toxicological evaluations conducted in *in vivo* experimental models.

The link between an upstream KE and a downstream KE in an AOP is called the key event relationship (KER). The KERs include the available evidence supporting the causal relationship between a pair of adjacent KEs. KERs contain mechanistic information of biological processes, i.e. biological plausibility, that are involved and connect the upstream KE to the downstream KE. The underlying evidence supporting the KERs is therefore based on the biological plausibility and empirical support of the KER and on the essentiality of the KE. These considerations are used as part of the WoE for the relationship between KEs and for the overall assessment of the AOP.

To fulfil the IATA needs, evidence was therefore structured using the AOP framework.

An AO can also be based on reported endpoints from HOS. However, in this Scientific Opinion, only *in vivo* experimental evidence was included in the AO while evidence from HOS was considered as supporting information in the conclusive remarks of the IATA. This decision was taken because (1) all HOS were appraised as Tier 3 in the RoB analysis (i.e. high RoB); (2) according to the EKE the probability for an association between deltamethrin exposure *in utero* and the group of endpoints measured in humans was less than 10% for most endpoints and less than 33% for one endpoint in one study; and (3) the limited translational value of animal models, which adds uncertainty to the comparison of the neurodevelopmental outcome measured in HOS with the AO observed in animal studies. Therefore, the human evidence was used as additional information in the deltamethrin IATA case study (see Section 3.6 for a detailed review on this topic).

The AOP conceptual framework was therefore applied to integrate information obtained from various lines of evidence by means of a systematic literature review, and the experimental outcome from the DNT-IVB to provide a structured contextualisation of the MIEs and KEs leading to the AO.

Because of the lack of predefined AOPs, and as a follow-up of the uncertainty analysis and EKE, an AOP was postulated with the inclusion of MIEs, KEs and AO (see Section 2.1.2 in Annex M).

Afterwards, the KERs were assessed for their biological plausibility, empirical support and essentiality of the KE. The three criteria were summarised using a Bayesian network (BN) approach (see Annexes B and M). The latter is a probabilistic graphical model that allows quantification of the uncertainty in the KERs via conditional probabilities (see Table 4 for definitions) that express the subjective belief that a MIE/KE/AO would occur conditionally to the status of the upstream KEs. Resorting to the BN allows: (1) to quantify the global dependency structure among MIEs, KEs and AOs described in the AOP network (joint probability, see Table 4 for definitions); (2) to assess the probability that each MIE, KE and AO would occur when exposure to the stressor occurs (marginal probability, see Table 4 for definitions); and (3) to perform scenario analyses assessing the impact of

new evidence, such as an individual MIE/KE occurring/not occurring with certainty, on the probability of the other KEs/AO in the network to occur.

Table 4: Probabilities associated to the Bayesian network assessment of the AOP (see Annexes B and M Section 1.4, for further details)

Probability	Definition	Elicitation
Conditional	The probability of each of the possible statuses of a downstream event given each possible status (or combination of statuses) of the linked upstream event(s) in the network (i.e. the conditioning events).	Provided as expert judgement based on the biological plausibility, essentiality and empirical evidence assessment of the KERs
Marginal	The probabilities associated to each possible state of a KE/variable (e.g. activated/not activated, occurrence/not occurrence) irrespective of the state of all the others.	Calculated based on the conditional probabilities (see Annex B)
Joint	The probability of each of the possible combinations of the status of the KEs in the network.	Calculated based on the conditional probabilities (see Annex B)

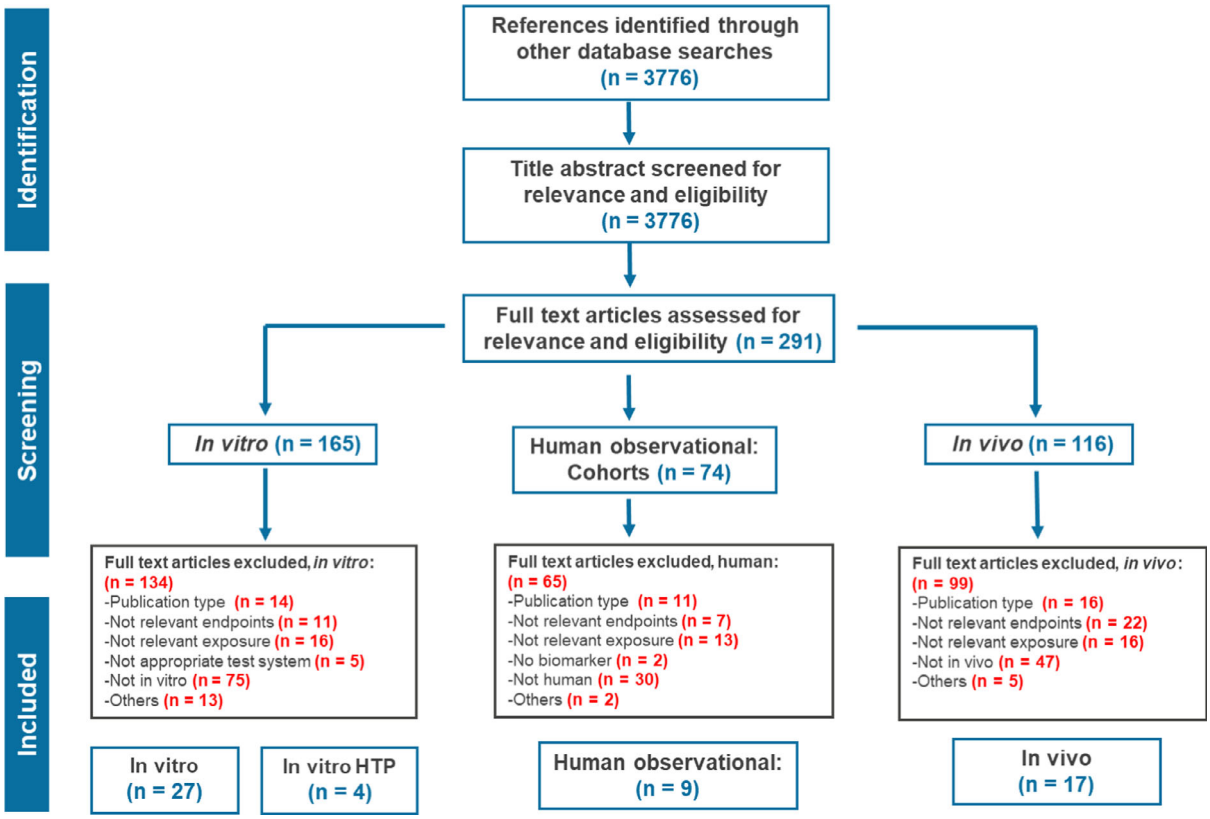
The outcome of the conditional probabilities was therefore used to quantify the strength of the relationship between KEs and the associated uncertainty. The marginal probability distribution was used to predict the most probable status (activation/not activation, occurrence/not occurrence) of the KEs/AO when exposed to the substance. The outcome of the joint probabilities computation was used to quantify the certainty that all events occur concurrently (see Annexes B and M).

The incorporation of statistical or probabilistic relationships into the AOP was used to develop a probabilistic quantitative WoE for the overall assessment of the AOP.

3. Assessment

3.1. Deltamethrin systematic review outcome

For deltamethrin, two independent reviewers screened the literature identified through the searches; 3,776 unique references were identified after removing duplicates (see PRISMA flow chart, **Figure 2**). The evidence was clustered as *in vivo* (containing *in vivo* experimental studies), *in vitro* (containing *in vitro* mechanistic studies and behavioural studies conducted in zebrafish up to 120 h post-fertilisation) or human (containing HOS) during the title and abstract screening. The title and abstract screening left 291 relevant articles that underwent a full-text review, of those, 165 were classified as *in vitro*, 74 as human evidence and 116 as *in vivo*. For *in vivo*, 99 publications were excluded and 17 were included. For human, 65 publications were excluded and 9 were included. For *in vitro* (including zebrafish studies), 134 publications were excluded and 31 were included. The main reasons for final exclusion of non-relevant literature were not relevant endpoints, not relevant exposure, not relevant system (*in vitro*, *in vivo*, human studies) or publication type. Full list of references included and excluded and reasons for exclusion are described in Annex C.



*: HTP: high throughput.

Figure 2: PRISMA flow chart of the systematic literature search process for deltamethrin, including the screening for relevance

3.2. Flufenacet systematic review results

For flufenacet, two independent reviewers screened the literature identified through the searches; 137 unique references were identified after removing duplicates (see PRISMA flow chart, **Figure** 3). The evidence was clustered as *in vivo* (containing *in vivo* experimental studies), *in vitro* (containing *in vitro* mechanistic studies and behavioural studies conducted in zebrafish up to 120 h fertilisation) or human (containing HOS) in the title and abstract screening. The title and abstract screening left five relevant articles of which three were classified as *in vitro*, 1) as human evidence; and 2) as *in vivo* (of those one was also classified as *in vitro*). After undergoing a full-text review, only one *in vivo* study (the OECD 426) and three *in vitro* studies were included. No relevant human publications were retrieved (see full list of references included and excluded and reasons for exclusion, Annex D).

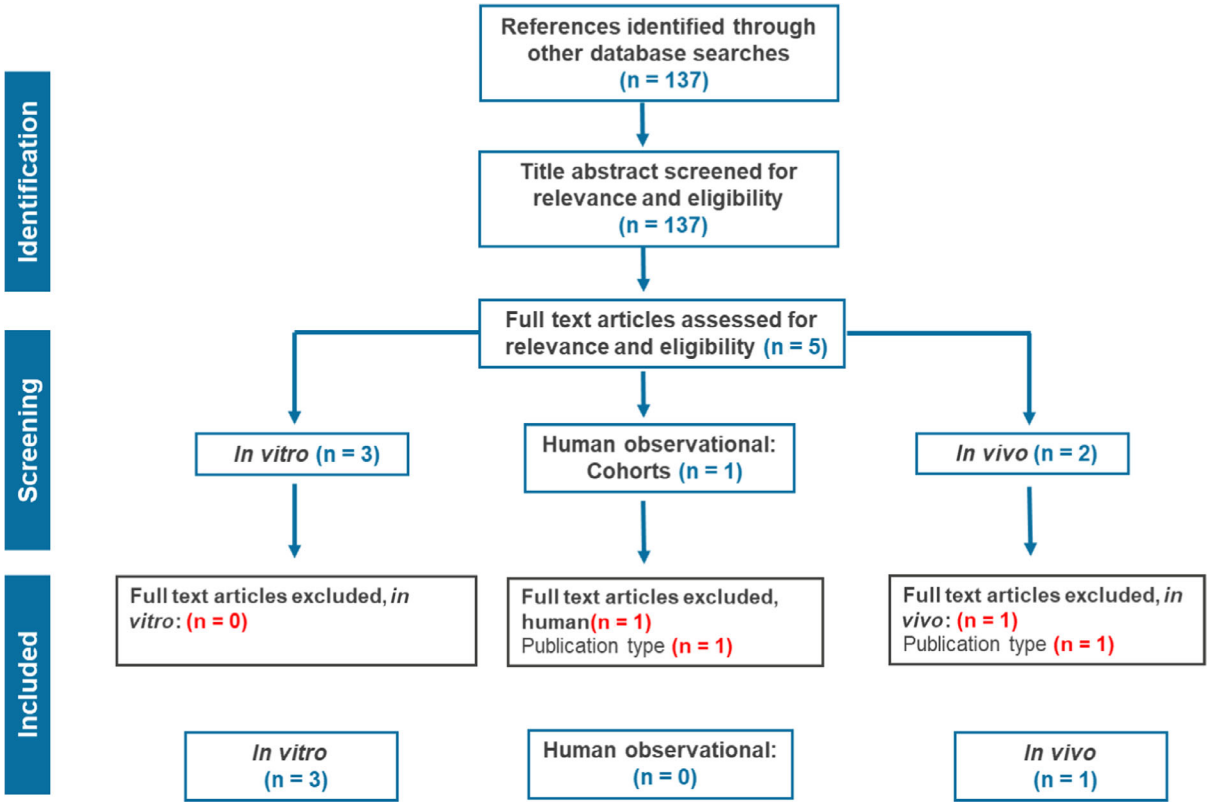
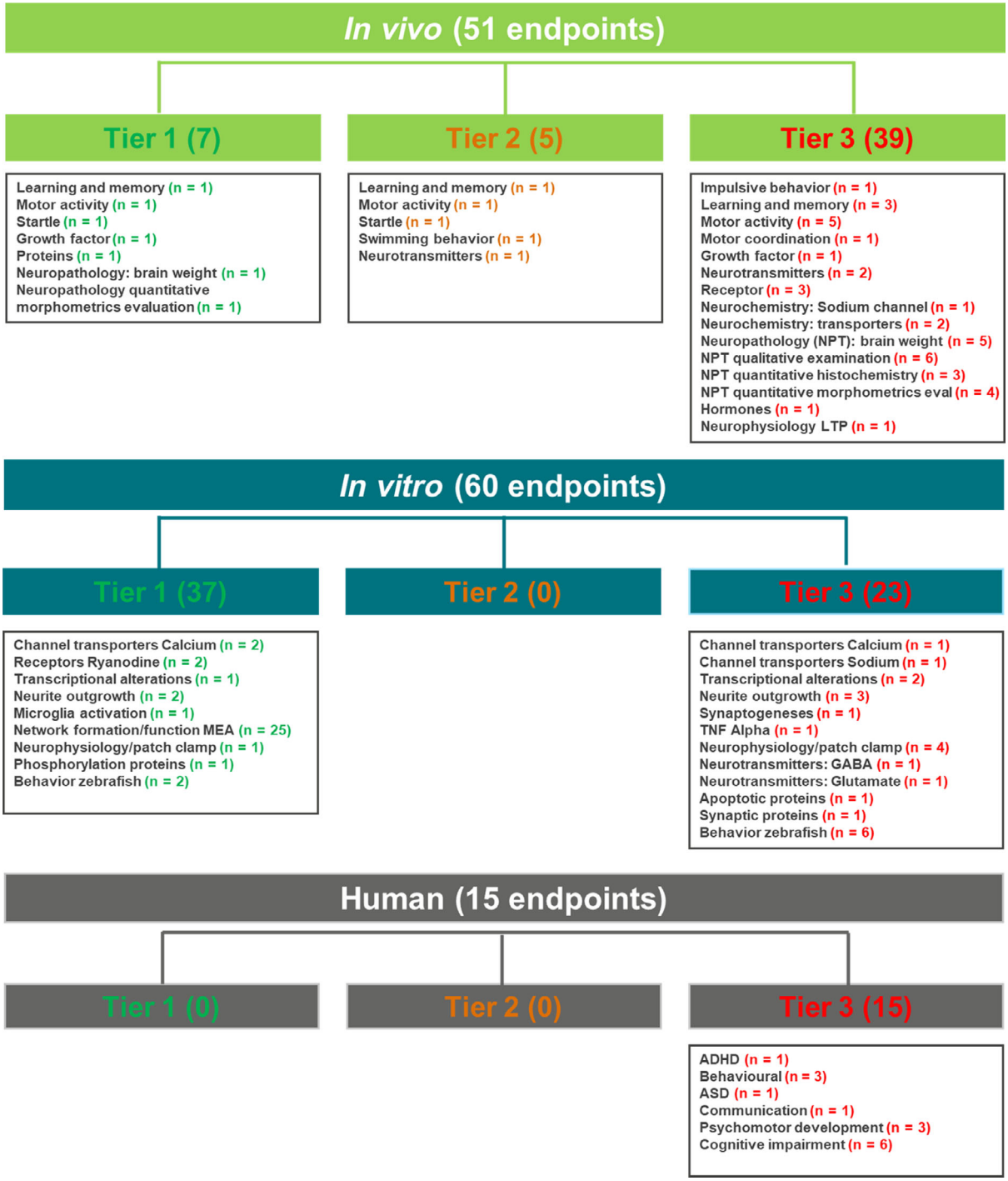


Figure 3: PRISMA flow chart systematic of the literature search process for flufenacet, including the screening for relevance

3.3. Critical appraisal results

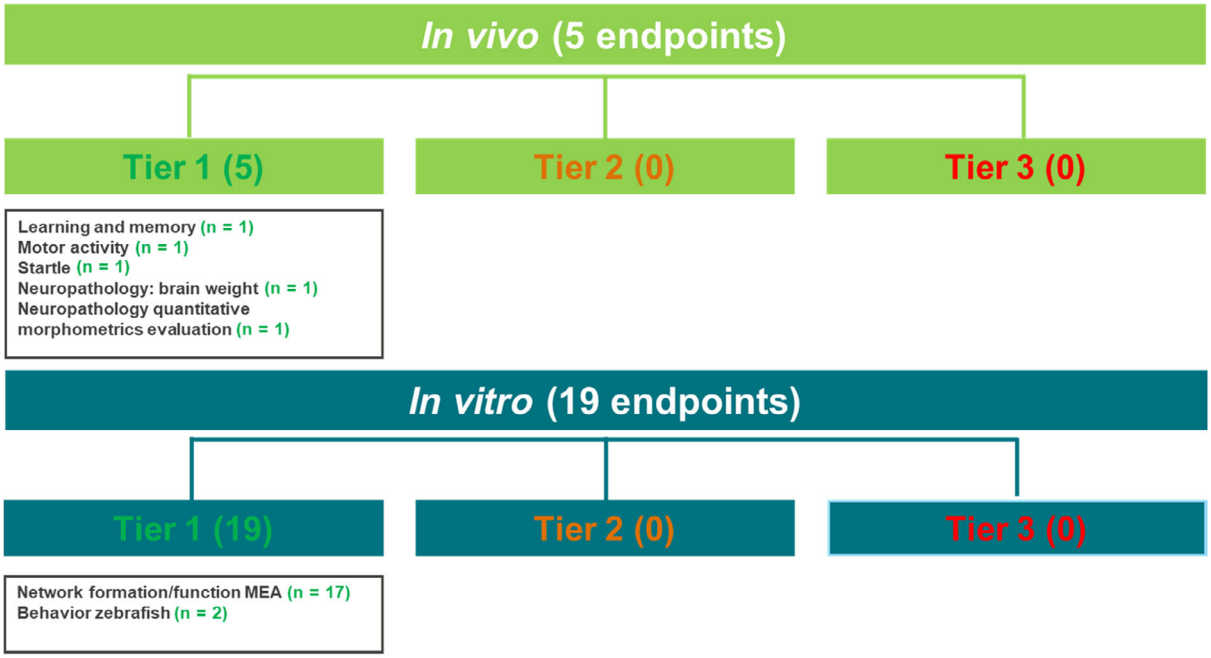
For deltamethrin, the outcome of the RoB appraisal is presented in Annex E for *in vivo*, *in vitro*, human and zebrafish lines of evidence.

For flufenacet, the outcome of the RoB appraisal and is presented in Annex F for *in vivo*, *in vitro* and zebrafish lines of evidence, respectively. Figures 4 and 5 present an overview of the results of the appraisal for deltamethrin and flufenacet, respectively. In Figures 4 and 5, endpoints categories are reported for each Tier (Tier 1 (low RoB), Tier 2 (moderate RoB), Tier 3 (high RoB)) and for each endpoint category the number of endpoints measured (n).



The total number of specific endpoints appraised in Tier 1 (low RoB), 2 (moderate RoB) or 3 (high RoB) is shown (see Section 2.2.3).

Figure 4: Summary of the appraisal of the endpoint categories of the studies selected for deltamethrin



The total number of specific endpoints appraised in Tier 1 (low RoB), 2 (moderate RoB) or 3 (high RoB) is shown (see Section 2.2.3).

Figure 5: Summary of the appraisal of the endpoint categories of the studies selected for flufenacet

The specific endpoints retained for the synthesis have been summarised and displayed graphically; the outcome of the data extraction and analysis by line of evidence and endpoint is presented in the Annex H containing all the graphs for *in vivo*, *in vitro* and zebrafish for deltamethrin, and the human evidence table in Annex I. For flufenacet graphs for *in vivo*, *in vitro* and zebrafish are presented in Annex J.

3.4. DNT *in vitro* testing battery (DNT-IVB)

In 2017, EFSA launched a procurement for the implementation and interpretation of IVB for the assessment of DNT. This procurement was awarded by a consortium between IUF – Leibniz Research Institute for Environmental Medicine and the University of Konstanz (Germany).

The overall goal of this procurement was to accelerate the development and use of *in vitro* test methods, conducted in human-relevant cell systems, capable of cost- and time-efficient testing of chemicals to assess their potential to disrupt the development of the nervous system, i.e. prediction of neurodevelopmental hazards to human health. To reach this goal, the contract was aimed to provide an appropriate scientific background for a DNT testing battery and facilitate the preparation of a guidance for alternative DNT testing by producing wet laboratory data and delivering data interpretation and user guidance. Here, the focus was on the application of a DNT-IVB (see assays in **Figure 6**) covering neurodevelopmental processes essential for brain development in a temporal context (**Figure 7**). The external scientific report produced was therefore used by the PPR Panel as critical information for the development of the two IATA case studies. In addition, the outcome of the report, and the report itself, will be used as experimental evidence available for the development of the OECD guidance on use and interpretation of the DNT *in vitro* test.

The final version of the report is available on the EFSA supporting publications (Masjosthusmann et al., 2020). The report includes not only the experimental work conducted on many chemicals from different domains but also the assessment of the DNT-IVB performance, a contextualisation of the DNT-IVB vs known DNT mechanisms of action, a proposal template for the description of the assays, classification models for compound classification and case studies. The analysis of the battery performance is also a critical step in the development of the OECD guidances. The assays included in the DNT-IVB are currently at high readiness levels when considering the scoring results (Bal-Price et al., 2018) and the performance standards are essential for the development of a performance-based test guideline. This is particularly important for DNT in which the number of known DNT-positive or -negative

compounds in human is limited and there is uncertainty on the outcome of the experimental *in vivo* DNT studies (Masjosthusmann et al., 2020; Paparella et al., 2020). The sensitivity and specificity analysis of the DNT-IVB was therefore conducted in the external report by relating categorised predefined compounds of the 119 chemicals tested to the real battery outcomes. This was performed through multiple steps. First, it was studied for the negative compounds preclassified in the substance set using the list of chemicals suggested in Aschner et al. (2017) and Masjosthusmann et al. (2020). This led to a specificity of 88% for the basic battery (without the neuronal network formation assay) and of 100% when the neuronal network formation assay was included. The next step predefined category was the human positive compounds. The nine selected human positive compounds (Masjosthusmann et al., 2020) were all testing positive in the DNT-IVB with the most sensitive endpoint scattered across the battery assays, meaning that each compound produced at least one specific hit. Besides, the 'negative' and the 'human positive', the analysis included the predefined compounds that show findings in *in vivo* experimental DNT studies in the 119 compounds (n = 20) set identified in Aschner et al. (2017) and Mundy et al. (2015). Although, some uncertainties are recognised in the analysis against the *in vivo* experimental studies, this resulted in an overall sensitivity of the test method battery (true positive/true positive + false negative), also including the neuronal network formation assay, of 90.5% and an overall specificity (true negative/true negative+false positive) of 100% (Masjosthusmann et al., 2020).

This report is therefore representing a fundamental step in the DNT assessment paradigm.

The report does not include the assays that are part of the DNT-IVB but that were conducted by the US-EPA as part of the current collaboration. The results of these assays have been reported in the public literature (Frank et al., 2017; Harrill et al., 2018; Shafer et al., 2019).

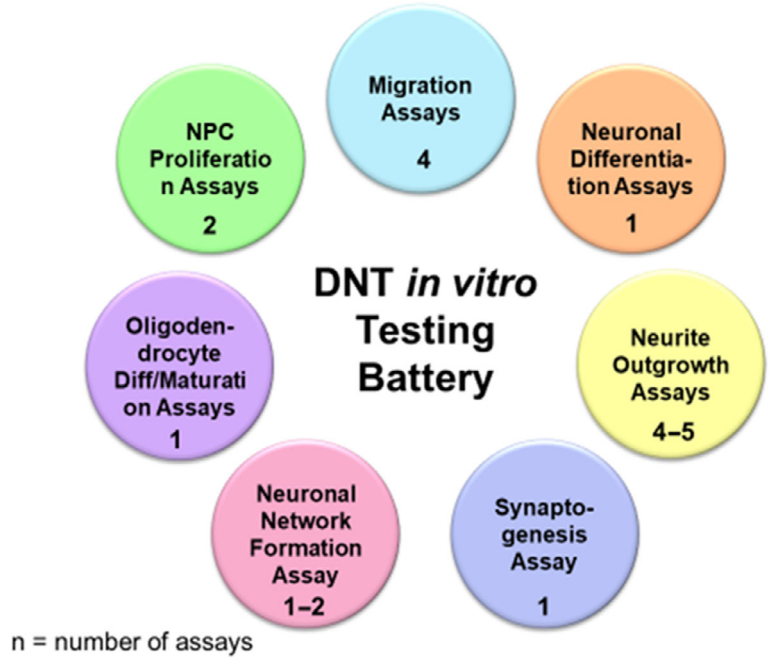
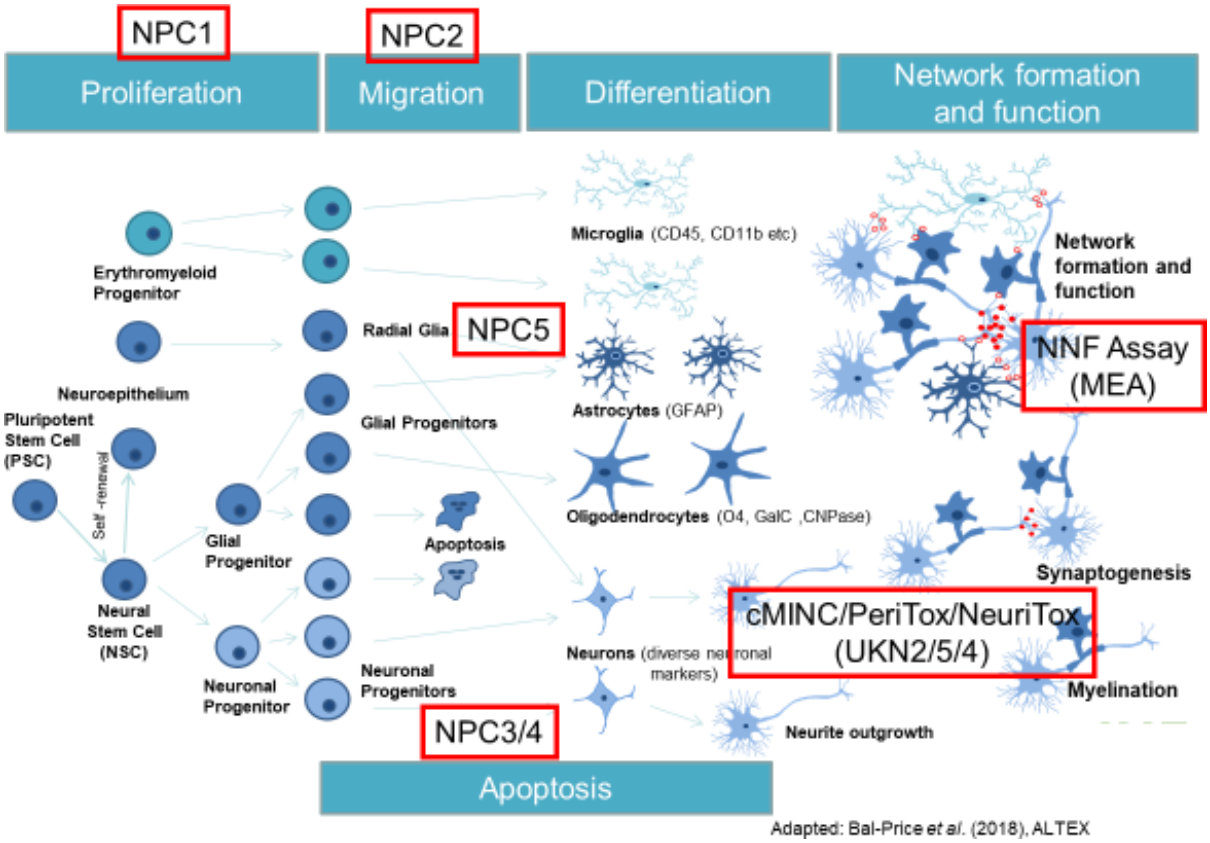


Figure 6: Test methods included in the DNT-IVB



NPC1: Primary human neuronal progenitor CELls (hNPC) proliferation assay.
NPC2: pRIMary hNPC migration assay.
NPC 3–4: pRIMary hNPC neuronal differentiation assay (NPC3) and Neuronal morphology (neurite length and area) of young neurons differentiatED From hNPC.
NPC5: Oligodendrocyte differentiATIOn.
cMINC: Neural Crest Cell Migration Assay (UKN2 Test).
PeriTox: Neurite Outgrowth of Peripheral Nervous System Neurons Test (UKN5 Test).
NeyriTox: Neurite Outgrowth of Central Nervous System Test (UKN4 Test).
NNF Assay (MEA): Neuronal network formation, microelectrode array in pre-differentiated human inducible pluripotent stem cells (hiPSC) or in rat neonatal derived primary cells.

Figure 7: Key neurodevelopmental processes and associated test methods and systems

Details of the test methods and test systems used to assess deltamethrin and flufenacet in the DNT-IVB and the results, including the concentration–response curve are given in Section 11.2 of the External Scientific Report (Masjosthusmann et al., 2020). The IVB was appraised with the CAT used for the *in vitro* evidence (see Section 2.2.3) and classified as Tier 1. The outcome of the RoB appraisal is presented in Annex G.

A summary of the results of the IVB for deltamethrin and flufenacet is given in Table 5.

Table 5: Summary of the results for the DNT-IVB for deltamethrin and flufenacet

	Deltamethrin	Flufenacet
BMC_{25} migration (UKN2)	**18.4^{s}**	> 100*
BMC_{25} neurite area (UKN5)	112.8^{ns}	> 100*
BMC_{10} migration radial cell (NPC2)	16.3^{ns}	> 20**
BMC_{30} neurite length (NPC4)	14.9^{ns}	> 20**
BMC_{30} neurite area (NPC4)	15.9^{ns}	> 20**
BMC_{30} oligodendrocyte differentiation (NPC5)	**0.6^{s}**	17.8^{ns}
BMC_{50} rat neuronal network formation (rNNF)	**0.5^{s}**	> 20
BMC_{50} human neuronal network formation (hNNF)	**4.1^{s}**	> 20
BMC_{30} neurite maturation (rat cortical neurons 2)	9.8^{ns}	–
BMC_{30} synaptogenesis (rat cortical neurons 2)	8.6^{ns}	–

s: specific hits (the BMC for the effect is separated from the BMC for cytotoxicity/cell viability).
ns: non-specific hits (the BMC for the effect is not separated from the BMC for cytotoxicity/cell viability).
Specific hits according to respective compound are marked in bold.
Units for the numerical values given in the table correspond to μM.
*: No benchmark response (BMR) at concentration higher than 100 μmol/L in the absence of cytotoxicity. The BMR is a value of effect size and is defined as an effect size that is higher than the general variability of the measured endpoint and is therefore determined based on the variability of the respective endpoint. For UKN 2 and 5 the BMR is set at 20% (BMR20). If the parameter is not affected at the BMR20 the compound is classified as a 'no hit'.
**: No benchmark response (BMR) at concentration higher than 20 μM in the absence of cytotoxicity. For NPC 2, 4 and 5, the BMR is defined as at least 1.5× the standard deviation (SD) (between experiment variation) as BMR10 for migration distance radial glia 72 h and 120 h, cytotoxicity 72 h and 120 h and BMR30 for all other endpoints.

In addition, UKN 2–5 and NPC 2–5, the table also includes data that was not generated as part of the EFSA report (Masjosthusmann et al., 2020):

- Benchmark concentration (BMC) values for neurite maturation and synaptogenesis (rat cortical neurons 2) were derived from Harrill et al. (2018). Flufenacet was not tested in these assays.
- Data for deltamethrin assessed in the rat neuronal network formation (rNNF) assay was taken from Frank et al. (2017) and flufenacet data from Shafer et al. (2019).
- hNNF values originate from unpublished data generated at IUF.

3.5. Outcome of the uncertainty analysis and of the expert knowledge elicitation

The outcome of the uncertainty analysis and of the EKE is reported in the Annexes K and L for question on hazard identification of deltamethrin and flufenacet respectively, in the Annex B for questions on hazard characterisation for deltamethrin (only for KE/AO requiring specification of a full probability distribution) .

In line with the methodology for assessing hazard identification (question 1), a judgement was made based on the identified unexplained inconsistencies and uncertainties on whether exposure to deltamethrin/flufenacet are causing the specific endpoint, endpoint category or the MIE/KE/AO. Uncertainty analysis tables are presented in Annexes K and L (for deltamethrin and flufenacet) including the answer to question 1 and the rationale agreed by the working group experts. Only when a causal association was identified, the lowest concentration/dose triggering the KE specific endpoint, endpoint category or adverse outcome was elicited. The threshold of 66% (twice as possible as not) was used as the minimum subjective probability (minimum level of certainty) leading to the conclusion of a causal association. This was considered to adequately reflect the difficulties in expressing the level of certainty and sufficiently precise when considering the purpose of this step of the process (screening potential KEs and AOs for the AOP). Only MIEs, KEs and AOs with an estimated probability of at least 66% were considered for the next step of the AOP network postulation (see Section 3.7). Table 7 in Annex N summarises the MIEs, KEs and AO kept for the AOP postulation description and its uncertainty analysis, including the EKE results. For assessing the hazard characterisation, question 2, the uncertainty affecting the body of evidence was expressed quantitatively using either ranges of probabilities (under the assumption that the uncertainty was limited and a uniform distribution could be assumed) or full probability distributions as suggested in the EFSA guidance on uncertainty analysis (EFSA and EBTC, 2018; EFSA Scientific Committee, 2018).

The estimates were derived using a semi-formal EKE. For the deltamethrin KE Altered Neuronal Network Function (assessed using a microelectrode array, MEA) and for the AO impairment behavioural function (assessed in the *in vivo* studies), a full probability distribution was elicited using the Roulette method (EFSA, 2014; O'Hagan, 2019). The individual uncertainty distributions of 12 experts for the KE and 7 experts for the AO were first summarised in a single distribution by mathematical averaging. Then the experts had the possibility to revise their individual judgements based on the collegial discussion and the change in average distribution until a consensus probability distribution was achieved. **Figures 8** and **9** represents the consensus probability and detailed methodology of the EKE conducted for the KE Altered Neuronal Network Function and the AO.

For flufenacet, such a detailed analysis was not carried out because in the question 1 of the assessment none of the KE/AO was assessed as being affected with a probability higher than 66%. Indeed, it was necessary only for deltamethrin to proceed further by contextualising the outcome of HOS and postulation of an AOP network. Figures 8 and 9 summarise how the uncertainty analysis and the EKE, with the inclusion of considerations on the biological plausibility, culminated in the postulation of the evidence-based AOP network (Figure 10 in Section 3.7). The uncertainty analysis is therefore a key building block in the IATA iterative process leading to further necessary steps, i.e. assessment of the HOS and development of a quantitative WoE as part of the overall assessment of the postulated AOP network, using deltamethrin as a model chemical. It was acting as a final conclusive step for the finalisation of the IATA for flufenacet.

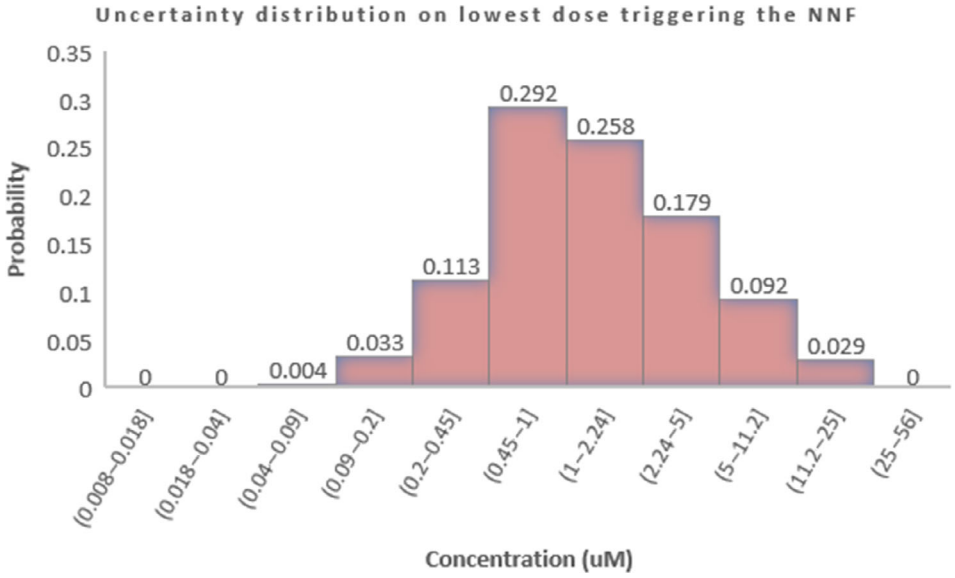
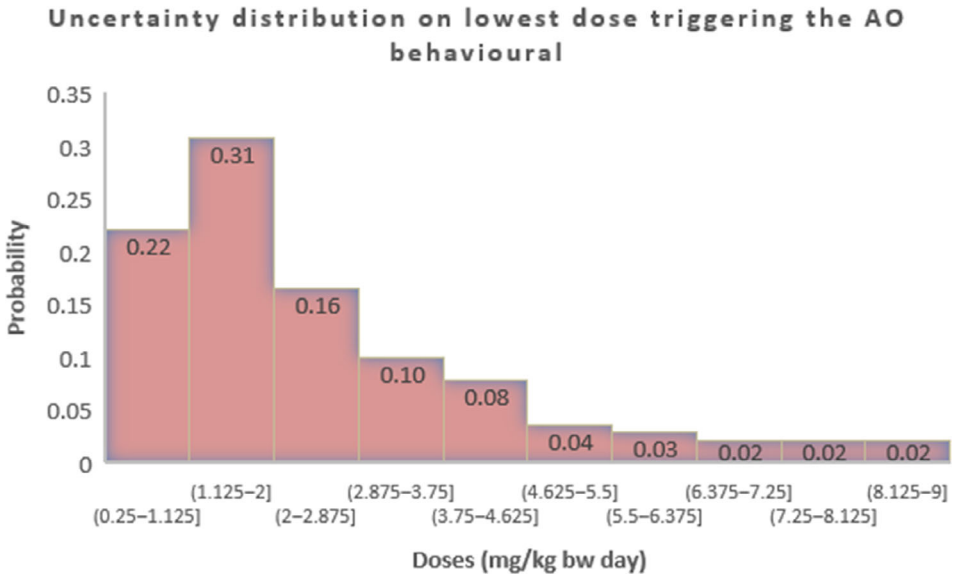


Figure 8: Uncertainty distribution on the lowest concentration triggering the KE Altered Neuronal Network Function in deltamethrin AOP, elicited through EKE. A round or squared parenthesis indicates that the extreme is excluded or included, respectively

Twelve experts participated to the elicitation. A logarithmic scale was considered by the experts more appropriate to split the overall credibility range into subsets. The bounds of the bins are in the range of –3; +2.5 in a logarithmic scale of base 5 with step increase of 0.5 (however, the figure shows the original scale). The outcome of the elicitation, based on the NNF (MEA) studies, provides a credibility range of 0.04–5 μmol/L that is expected to include the true lowest concentration with around 88% probability. The judgement stemmed from the consideration that lower bounds of concentration for individual studies with different test systems and methods mostly overlap in the range of 0.45–5 μmol/L covering both the distribution for single administration (acute protocol) and multiple administration over developmental period of the network (developmental protocol) (see Annex B for details in the EKE process and Annex M for detailed results of the discussion).



Doses are related to external doses.

Figure 9: Uncertainty distribution on the lowest dose triggering the AO (behavioural changes) in deltamethrin AOP. A round or squared parenthesis indicates that the extreme is excluded or included, respectively

The experts discussed the distribution by considering the expected corresponding brain concentration in the pups. The experts agreed that the range of external doses between 0.25 and 2 mg/kg body weight (bw) per day (based on the PK study are likely to yield the highest concentration in the brain) had a 53% probability to include the true lowest dose that would trigger the AO. This distribution mainly reflects the occurrence of the AO in what is expected to be the most sensitive population, i.e. the pups, and the AO at the lowest dose is considered based on effect on sensory motor and cognitive behaviours. Overall, the credible range 0.25–7.25 mg/kg bw per day is expected to include the true lowest dose with 96% probability (see Annex B for details in the EKE process and Annex M for detailed results of the discussion).

3.6. Human observational studies

3.6.1. Background

Neurodevelopmental disorders encompass a broad and etiologically heterogeneous group of conditions characterised by alterations in the development of the central nervous system (CNS) that manifest generally occurring very early in life due to the rapid development of brain circuitry during this period (Nakai et al., 2018). The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; APS, 2013) provided a new classification for NDDs which includes the following six categories: intellectual disabilities, communication disorders, autism spectrum disorder (ASD), attention-deficit hyperactivity disorder (ADHD), specific learning disorders, motor disorders and other NDDs (see Annex A).

NDDs affect 10–15% of all births, and subclinical decrements in brain function are thought to be even more common compared with the disorders themselves (Grandjean and Landrigan, 2014). Children with NDDs can experience difficulties with language and speech, attention, memory, learning, behaviour, motor skills or impairment of other neurological functions. Many affected children may have more than one of these conditions. Symptoms and behaviours associated with NDDs often change or evolve as the child grows older (ACE, 2019). The multiple clinical signs of NDDs may be related to a common underlying multifactorial aetiology and pathophysiology (Cheroni et al., 2020).

3.6.2. Risk factors for neurodevelopmental disorders

Multiple risk factors have been associated with NDDs, including genetic, environmental, infectious and traumatic, among others, which in general interact with each other rather than acting alone. The

co-occurrence of distinct NDD categories in the same children indicates shared underlying biological/cellular mechanisms between them (Cardoso et al., 2019).

While genetic susceptibility accounts for about 30–40% of NDDs, environmental exposures also play a role in some cases probably by interacting with genetic factors (Grandjean and Landrigan, 2014). Whole-exome and whole-genome sequencing of patients has found evidence for high genetic heterogeneity within, between and across different NDDs, having documented a considerable overlap of genes involved in more than one NDD. The current mutational spectrum of NDDs includes hundreds of genes related to neurodevelopmental pathways such as those associated with chromatin remodelling, synaptic function and transcriptional regulation (Cardoso et al., 2019). For ASD, although the heritability is high and multiple genes have been implicated as causal, most of these genes have been identified in *de novo* cases (Rubinstein et al., 2018). In sum, NDDs are highly polygenic, with pleiotropic risk alleles and a complex background of gene–environment interactions driving their pathophysiology (Del Pino et al., 2018).

The overall increase in NDDs reported in recent years has been suspected to be linked to increased exposure to environmental factors, which may influence neuronal development and activity, as well as to changes in diagnostic patterns of these diseases in populations. The perturbation produced by external agents during the prenatal and perinatal periods has been widely reported for NDDs in general and for ASD in particular. The environmental factors that have been consistently associated with NDDs include paternal age, gestational age, low birth weight, birth defects and perinatal hypoxia and respiratory stress, lower socioeconomic status and prenatal or childhood exposure to certain environmental chemicals (ACE, 2019; Carlsson et al., 2020; Cheroni et al., 2020). Serious infection and exposure to psychoactive or addictive drugs during early development are well-known environmental factors relevant to NDDs (Hsueh, 2019). However, this evidence comes mainly from observational studies in certain geographical areas of the world and is susceptible to familial and other sources of confounding and therefore causality with many of the examined factors is difficult to establish.

Toxicants from different chemical classes have been identified as potential risk factors for several NDDs. While lead, methylmercury and PCBs are the developmental neurotoxicants best known, other chemicals (and drugs) have also been studied for their potential impairment of children neurodevelopment, including exposure to organophosphate pesticides (OPs), polybrominated diphenyl ether flame retardants (PBDEs), phthalates, bisphenol A (BPA), polycyclic aromatic hydrocarbons (PAHs), arsenic and perchlorate (ACE, 2019). However, the association of chemical substances on the manifestation of NDDs is difficult to establish due to the long-term co-exposures to a complex mix of substances and the inherent limitations of human studies to provide evidence for causality (Cheroni et al., 2020).

3.6.3. Epidemiological studies on prenatal exposure to deltamethrin and neurodevelopmental outcomes in children

Relevance and limitations of HOS

Human data, where they exist, are considered to be an essential component of chemical risk assessment. However, regulators have set clear limitations on the type of human studies from which data can be used, and for what purpose. For pesticides, Regulation (EC) 1107/2009 stipulates that:

> 'the assessment of an active substance or a plant protection product should not be based on tests or studies involving the deliberate administration of the active substance or plant protection product to humans with the purpose of determining a human 'no observed effect level' of an active substance. Similarly, toxicological studies carried out on humans should not be used to lower the safety margins for active substances or plant protection products'.

In practice, this limits the use of human data to observational studies such as, surveillance schemes of exposed workers or the general population, and epidemiological investigations such as cohort, case-control or cross-sectional studies. For the present case study, i.e. consideration of the evidence for deltamethrin being a developmental neurotoxicant using an IATA approach, only birth cohorts were considered as suitable observational studies to be included in the evidence base.

The main benefit of using observational studies in human risk assessment is that they are concerned with an adverse outcome of interest in a population of interest, therefore avoiding the need for extrapolation from experimental studies (Nachman et al., 2011). They are a reflection of 'real life' in terms of including heterogeneous populations who have been exposed to relevant environmental

concentrations of a stressor by relevant routes of exposure. The exposure is cumulative and there is the right temporal relationship between exposure and effect, including relevant co-exposures to other chemical and non-chemical stressors (Christensen et al., 2015).

However, the very advantages of observational over experimental studies also provide major challenges constraining their more widespread use in human risk assessment. With very few exceptions, observational studies deal with AOs of multifactorial origin, and the contribution of a single agent to the outcome is difficult to assess, let alone to quantify. Hence, the most important limitations of observational studies relate to: (1) the presence of confounders and effect-modifying factors, and (2) their ability to provide a realistic assessment of exposure.

While confounders are factors that are associated with both the exposure and the outcome, effect modifiers are factors associated with the outcome but not the exposure. Both have the potential to distort or even invalidate findings from observational studies, unless they are appropriately taken into account in the statistical analysis or during the design of the study. The distinction between confounders and effect modifiers is not always easy, but socioeconomic status or certain lifestyle factors are typical examples of confounders whereas co-exposure to other stressors that change the magnitude of the effect (e.g. co-exposure to chemicals other than deltamethrin causing the same outcome) could be considered as effect modifiers.

Exposure assessment can take many forms, from a simple questionnaire to sophisticated analysis of chemicals, metabolites or effect biomarkers in human samples or event real time monitoring. All are subject to error and it has been argued that misclassification bias as a result of incorrect exposure assessment in epidemiology studies is likely to underestimate the true effects (Blair et al., 2007). Environmental exposures vary greatly over time and in intensity and constructing meaningful exposure-response relationships, especially when dealing with potentially small windows of vulnerability to certain effects, are exceedingly difficult.

An important aspect of the assessment of study quality is the appraisal of the 'internal validity' of the study, i.e. whether the evidence presented in the study supports a causal relationship between the exposure and the effect. It must be emphasised that observational studies are not designed to provide direct evidence of causality. Instead, the observed associations must be critically examined on the likelihood of a cause-effect relationship. This can be facilitated using criteria such as the ones originally proposed by Bradford Hill (Hill, 1965) but adapted to the needs of data integration from multiple sources (Fedak et al., 2015).

According to the selection criteria specified in the protocol for the systematic review, nine HOS were identified for deltamethrin (see main features in Table 6). An additional study was further considered in the updated literature search conducted in November 2020 (Barkoski et al., 2021). This study assessed the relationships between prenatal pyrethroid exposure, using maternal urinary levels of 3-PBA, and ASD or non-typical development at 3 years. Although the study found a moderately elevated relative risk ratio for ASD as compared with typical development, this finding was not statistically significant and the confidence interval was wide. Furthermore, no evidence for differences in 3-PBA levels was observed when comparing non-typical development with typical development.

Table 6: Main features of the nine human observational studies identified in the systematic literature search

Study (Ref. id. n)	Primary exposure	Co-exposure	Confounders	Outcome	Instrument	Children's age
Ref. 129 Viel et al. (2017); (n = 287)	3-PBA, 4-F-3-PBA *cis*-DCCA *trans*-DCCA *cis*-DBCA	OPs (ΣDAPs) Pb (home dust)	Socio-demographic Lifestyle Environmental factors Clinical information (pregnancy and birth)	Behaviour, other (prosocial behaviour, internalising disorders, externalising disorders)	Strengths and Difficulties Questionnaire (SDQ)	6 years
Ref. 433 Xue et al. (2013); (n = 497)	3-PBA *cis*-Cl2CA *trans*-Cl2CA		Socio-demographic Environmental factors Clinical information (pregnancy and children)	Cognitive impairment	Development Screen Test (DST) scale: -mental development index (MI) -developmental quotient (DQ)	1 year
Ref. 480 Fluegge et al. (2016); (n = 140)	3-PBA *trans*-DCCA	OPs metabolites (TCPy, IMPy)	Socio-demographic Lifestyle Clinical information (pregnancy and infant)	Cognitive impairment (MDI) Impaired psychomotor development (PDI)	Bayley Scales of Infant Development-II (BSID-II): MDI and PDI	3 months
Ref. 876 Dalsager et al. (2019); (n = 948)	3-PBA *trans*-DCCA	OPs (TCPy)	Socio-demographic Lifestyle Clinical information (pregnancy and infant)	Attention-deficit/hyperactivity disorder (ADHD)	Child Behaviour Check List for ages 1.5–5 years (CBCL: 1½-5)	1.7–4.1 years
Ref. 1117 Eskenazi et al. (2018); (n = 432)	3-PBA *cis*-DBCA *cis*-DCCA *trans*-DCCA	*o,p'*-DDT *p,p'*-DDT *o,p'*-DDE *p,p'*-DDE Pb (children blood)	Socio-demographic Lifestyle Clinical information (pregnancy and children) Occupational and residential history	Cognitive impairment Impaired psychomotor development Communication disorders	Bayley Scales of Infant Development, 3rd edition (BSID-III). Subtests: cognitive, language (receptive and expressive), motor (fine and gross), Social-Emotional	1 year 2 years
Ref. 1118 Furlong et al. (2017); (n = 162)	3-PBA, *trans*-DCCA, *cis*-DCCA	OPs metabolites (DEDP, DEP, DETP, DMDP, DMP, DMTP)	Socio-demographic Lifestyle Environmental factors Children data (age, Preterm birth)	Cognitive impairment Behaviour (behavioural regulation index, internalising composite)	Behavioural Assessment System for Children (BASC) Behaviour Rating Inventory of Executive Function (BRIEF)	4–5 years 6 years 7–9 years
Ref. 1152 Viel et al. (2015); (n = 205)	3-PBA *cis*-DBCA 4-F-3-PBA, *cis*-DCCA, *trans*-DCCA	OPs metabolites (six DAPs) Pb (floor dust)	Socio-demographic Lifestyle Environmental factors Exclusion criteria (for mothers and children)	Cognitive impairment	Wechsler Intelligence Scale for Children (WISC-IV). Domains: verbal comprehension, working memory	6 years

Study (Ref. id. n)	Primary exposure	Co-exposure	Confounders	Outcome	Instrument	Children's age
Ref. 1432 Watkins et al. (2016); (n = 187)	3-PBA	Pb (maternal blood)	Socio-demographic Exclusion characteristics for mothers and children	Cognitive impairment Impaired psychomotor development	Bayley Scales for Infant Development—Spanish version (BSID-IIS): MDI, PDI	2 years 3 years
Barkoski et al. (2021); (n = 201)	3-PBA, *trans*-DCCA	TCPy	Socio-demographic Maternal features	ASD and non-typical development (non-TD)	Mullen Scales of Early Learning (MSEL) Autism Diagnostic Observation Scale (ADOS)	3 years

Further details on the analysis of risk of bias in HOS

In systematic reviews, a structured approach to the assessment of internal validity is essential and often referred to as RoB analysis. A bias is a systematic error in the study which may lead to either an over- or underestimation of the 'true' effect of the exposure. As indicated in Section 2.2.3, the OHAT-NTP tool was employed in this Scientific Opinion for the RoB analysis. Seven questions were formulated for HOS, and the three questions on the accounting for important confounding and modifying variables, and confidence in the exposure and outcome assessments were identified as key questions for the RoB analysis (see Table 7).

The studies selected were analysed for RoB. Of these, 5/10 studies were judged to have probably or definitively high RoB for confounding, while for exposure assessment, all studies had probably or definitively high RoB (Table 7). Although only studies using biomonitoring data were included, none of the studies measured deltamethrin in blood, but specific or non-specific metabolites in urine. This does not allow an accurate estimate of exposure to deltamethrin based on the uncertainty in kinetics and on the correct timing of urine samples collection, and the presence of the metabolites in environmental media. Overall, the 10 studies retrieved from the systematic literature search were categorised as Tier 3 (high RoB), mainly due to the inadequate assessment of confounding, selection bias and suboptimal measurements of exposure with non-specific biomarkers (Table 7). This tier resulted from the application of an algorithm of OHAT/NTP to the seven questions addressing the RoB.

Table 7: RoB analysis of the human observational studies using the OHAT/NTP tool (NTP, 2015)

Ref. ID	Q1_Select	Q2_Confnd	Q3_Attr	Q4_Expos	Q5_Outcom	Q6_SelectvRep	Q7_Stats	TIER
129	DLRoB	PHRoB	DLRoB	PHRoB	PLRoB	DLRoB	PLRoB	3*
433	DLRoB	DHRoB	DLRoB	DHRoB	DLRoB	PLRoB	NR	3
433	DLRoB	DHRoB	DLRoB	DHRoB	DLRoB	PLRoB	PHRoB	3
480	DLRoB	DHRoB	PLRoB	DHRoB	PLRoB	PLRoB	DLRoB	3
480	DLRoB	DHRoB	PLRoB	DHRoB	PLRoB	PLRoB	DLRoB	3
876	DLRoB	PLRoB	PLRoB	DHRoB	PLRoB	PLRoB	DLRoB	3
1117	PLRoB	PLRoB	DLRoB	PHRoB	DLRoB	PLRoB	DLRoB	3
1118	DLRoB	PLRoB	PHRoB	DHRoB	PLRoB	PLRoB	PLRoB	3
1152	DLRoB	DLRoB	PHRoB	PHRoB	DLRoB	DLRoB	PLRoB	3
1432	DLRoB	PLRoB	PLRoB	DHRoB	PLRoB	PLRoB	PLRoB	3
1432	DLRoB	PLRoB	PLRoB	DHRoB	PLRoB	PLRoB	PLRoB	3
3034541	DLRoB	PLRoB	PLRoB	PHRoB	DLRoB	DLRoB	DLRoB	3

DLRoB: definitively low RoB; PLRoB: probably low RoB; PHRoB: probably HIGH RoB; DH: definitively HIGH RoB.
*: Indicates that studies classified as DHRoB or PHRoB for critical questions will be considered as Tier 3 (for further details see Annex A).

The results of the systematic literature review for the HOS, covering search string, studies selection and retrieval, appraisal of their internal validity, as well as data extraction, synthesis and integration, has been addressed earlier in Section 2.2.3. of this Scientific Opinion. These results are summarised in Table 8.

Table 8: Summary of the associations found between urinary pyrethroid metabolites and neurodevelopmental endpoints in human observational studies from the systematic review performed

Endpoint categories	No. studies	No. analyses	Metabolites	Age range	Associations detected
Cognitive impairment	8	14	3-PBA, *cis*-DBCA, total permethrins[(1)]	3 months–6 years	Inverse association with MDI (3 months) for 3-PBA Inverse association with MDI (1 year) for total permethrins

Endpoint categories	No. studies	No. analyses	Metabolites	Age range	Associations detected
Impaired psychomotor development	3	15	3-PBA, *cis*-DBCA	3 months–3 years	Inverse (female) and direct (♂) association with gross motor (2 years) for 3-PBA 3-PBA (↑ motor composite, 2-year ♂) *cis*-DBCA (↓motor composite, ↓gross motor, 2-year ♀)
Communication	1	14	3-PBA, *cis*-DBCA	1–2 years	Inverse association with expressive communication (2 years, ♀ for 3-PBA/♀ and all children for *cis*-DBCA)
					Inverse association with language composite (2 years, ♀ for 3-PBA / ♀ and all children for *cis*-DBCA) Inverse association with language composite (1 year, ♀ for *cis*-DBCA)
Behavioural	3	14	3-PBA, *cis*-DBCA	4–9 years	Inverse association with Internalising Composite (4–5, 6, 7–9 years for 3-PBA)
					Inverse association with Behavioural Regulation Index (4–5, 6, 7–9 years for 3-PBA)
					Inverse association with Social-Emotional (1 year, ♂ for 3-PBA)(2)
ADHD	1	1	3-PBA	2–4 years	Direct association with ADHD and ^{AD}HD > 90th-ile (2–4 years)

MDI: Mental development index.
(1): Total permethrins: S3-PBA+*cis*-DCCA + *trans*-DCCA.
(2): An inverse association was also observed for *cis*-DCCA at 1 year, which may partially account for the inverse association observed for 3-PBA.

Following the RoB analysis, studies were grouped according to the neurobehavioral or neurodevelopmental endpoints that they had examined (cognitive impairment, impaired psychomotor development, communication, behavioural and ADHD) and an uncertainty analysis was carried out with the aim to derive a level of probability that an association between deltamethrin exposure *in utero* and the specific endpoint categories might occur. Inverse statistically significant associations were observed for all endpoints studied with the exception of ADHD, which showed a direct association. However, these associations corresponded mostly to the non-specific pyrethroid metabolite 3-PBA, the biomarker of pyrethroid exposure most frequently detected. Hence, the effect cannot be exclusively attributed to deltamethrin as 3-PBA is a shared metabolite of some commonly used type I and type II pyrethroid insecticides (cypermethrin, deltamethrin, permethrin, cyhalothrin). Regarding impaired psychomotor development, inconsistencies were observed between boys and girls at 2 years of age. Although behaviour was significantly impaired across a large age range (1–9 years), this association was observed only for 3-PBA but not for *cis*-DBCA, a specific deltamethrin metabolite. Furthermore, the association found for socio-emotional behaviour with 3-PBA in 1 year old male children might be partially accounted for *cis*-DCCA, which is a pyrethroid metabolite not related to deltamethrin. The more remarkable findings potentially related to deltamethrin (using urinary *cis*-DBCA as a surrogate biomarker of exposure) corresponded to communication impairment, as an inverse association was observed for expressive communication and language composite for all children and specially for girls at 1 and 2 years of age.

3.6.4. Interpretation of results and potential mechanism underlying the associations found

Pyrethroid insecticides exert their neurotoxic action by delaying the inactivation of voltage-gated sodium channels (VGSC or Na_V). Deltamethrin, a type II pyrethroid, has been shown to alter the function of multiple Na_V isoforms in *in vitro* studies, voltage-clamp experiments and in studies addressing the mode of action in mosquito and other arthropods (Tapia et al., 2020).

Mutations in genes encoding different Na_v channels have been associated with NDDs. In addition, developmental exposure to pyrethroids contributes to the increase in these disorders as a result of disruption of Na_v isoforms. In particular, alteration in biophysical properties of $Na_v1.3$ channels may impair the prenatal development of human cortical language areas resulting in speech and oral motor dysfunction. Markedly, prenatal cortical progenitor cells and neurons do not show detectable action potentials, supporting an unexpected role for sodium channel ($Na_v1.3$) in cortical organisation and neuronal proliferation and migration in the developing brain, especially in speech and language areas (Smith et al., 2018). This observation provides biological support to the communication impairment observed in the systematic literature review performed.

$Na_v1.6$ is highly expressed in the medium spiny neurons (MSNs) of the nucleus accumbens (NAc), a neuronal population characterised by repetitive firing properties and vulnerability to excitotoxicity. The $Na_v1.6$ channel is the primary molecular determinant of MSN firing and it is sensitive to modification by pyrethroids. Dysregulation of MSN firing in the NAc is thought to play a critical role in the pathophysiology of ADHD and other NDDs. in the NAc, acute exposure to deltamethrin induces changes in $Na_v1.6$ biophysical properties, leading to long-term inactivation (LTI) of the channel. These changes modify intrinsic excitability of MSN cells by increasing evoked action potential firing frequency and inducing aberrant action potentials with low amplitude and depolarised voltage threshold. These changes are also expected to disrupt MSN firing with durable long-term effects on the reward circuit (Tapia et al., 2020). Deltamethrin also altered the expression of MSN proteins important for NAc signalling, indicating that these cells are especially sensitive to pyrethroid exposure (Magby and Richardson, 2017; Richardson et al., 2015). These observations point out to a possible mechanism for deltamethrin toxicity involving disruption of the NAc circuitry over time, increasing the risk of ADHD and other NDDs (Tapia et al., 2020).

Because the various Na_v isoforms are expressed in different neural circuits, developmental exposure of deltamethrin may impair circuit formation resulting in a wide array of clinical manifestations depending on the dysfunctional circuits. Therefore, altered MSN of the NAc are involved in abnormalities in hyperactivity and motivational behaviour observed in ADHD, changes in striatal circuits can lead to locomotor impairment and deficit in hippocampal neurogenesis may result in cognitive (learning and memory) deficits.

Developmental exposure to deltamethrin causes dopamine (DA) dysfunction and further long-term locomotor impairment. Mice developmentally exposed to deltamethrin at doses below the NOAEL (3 mg/kg) from gestational day 0–21 showed hyperactivity associated with increased expression of the DA transporter (DAT) and the DA receptor D1 (DRD1) in the NAc, and decreased synaptic DA levels in the striatum of adult mice, which can be considered as subtle endpoints of locomotor activity (Richardson et al., 2015). Furthermore, deltamethrin-induced hyperactivity was attenuated by D1 receptor antagonists and D2 receptor agonists, indicating the role of the DA system in mediating these effects (Richardson et al., 2015). For this study, Richardson et al. (2015), only the experimental part was used for the mechanistic explanation. The epidemiological part was considered not relevant for this Scientific Opinion because it concerns children aged 6-15 with contemporary assessment of the 3-PBA metabolite. In another study, neonatal rats exposed to deltamethrin from PND 3 to 20 showed adverse long-term effects on learning, memory and startle reactivity, with some of these effects being sexually dimorphic. This study also found decreased synaptic DA levels in the NAc, reduced DRD1 mRNA expression in neostriatum, and decreased norepinephrine (NE) levels in the hippocampus (Pitzer et al., 2019). While the administration of deltamethrin throughout gestation and lactation exhibited decreased VGSC subunit mRNA expression in adult mice (Magby and Richardson, 2017), the study of Pitzer et al. (2019) did not find alterations in *VGSC* gene expression. DAergic dysfunction is a likely mode of action for the persistent hyperactivity also observed in the zebrafish exposed to deltamethrin during the development (Kung et al., 2015). The behavioural abnormalities mentioned previously are similar to those observed in children with ADHD, including elevated DAT levels, increased locomotor activity, impulsive-like behaviours, and deficits in working memory and attention, as well as a male sex preference of these effects (Richardson et al., 2015).

Studies performed in adult mice treated with deltamethrin or permethrin also found increased DAT levels. Since DAT can greatly affect the vulnerability of DAergic neurons to neurotoxicants, upregulation of DAT may increase the susceptibility of these neurons to toxic insult (Elwan et al., 2006). Brain imaging studies in ADHD children have found abnormalities of frontostriatal circuits modulated by DA (e.g. increased DAT levels) (Krause, 2008); however, there is still controversy on the potential role of the DAT in ADHD.

Overall, these data provide a mechanistic basis to suggest that developmental pyrethroid exposure is a significant risk factor for ADHD (Richardson et al., 2015). These experimental findings provide support to epidemiological evidence indicating that either prenatal (Dalsager et al., 2019) or postnatal (Richardson et al., 2015) exposure to pyrethroids, assessed by urinary non-specific pyrethroid metabolites, is associated with an increased risk of ADHD diagnosis.

Co-exposure to deltamethrin and stress during neurodevelopment has been reported to impair DA function. A hypermethylation of the glucocorticoid receptor gene (*NR3C1*) has been observed in the midbrain of C57/BL6N male mice in response to the exposure to both deltamethrin and corticosterone during development. These findings suggest possible connections between multiple environmental exposures impacting on the DA system and the hypothalamic–pituitary–adrenal axis via changes in DNA methylation and provide new insight about epigenetic effects in adulthood after exposure during neurodevelopment (Vester et al., 2020).

Conversely, prenatal exposure of rats to a relatively high dose of deltamethrin (9 mg/kg) altered learning and memory abilities in offspring and decreased the levels of some *N*-methyl-D-aspartate receptor (NMDAR) subunits, brain derived neurotrophic factor (BDNF), Tyrosine kinase B (TrkB) receptor, and phosphorylated cAMP response element binding protein (pCREB) in the hippocampus of offspring rats (Zhang et al., 2018). However, no significant changes in these protein expressions were found for lower doses. This study concluded that deltamethrin caused cognitive impairment that is likely to be a result of an extended opening of the calcium channel, leading to overactivation of NMDAR and further neural damage through NMDAR/BDNF signalling.

Deficits in learning and memory are often associated with disruption of hippocampal neurogenesis, which is regulated by numerous processes, including precursor cell proliferation, survival, migration and differentiation to mature neurons. Short-term administration of deltamethrin to mice caused significant deficits in adult hippocampal neurogenesis, which was associated with impaired learning and memory (Hossain et al., 2020). Treated animals showed decreased cellular proliferation in the dentate gyrus of the hippocampus as evidenced by a decrease in nestin-expressing neural progenitor cells, a reduction in the expression of doublecortin (an early neuronal differentiation marker) and a reduction in total number of granule cells (Hossain et al., 2020).

The results of the systematic review and of the *in vitro* testing battery (Masjosthusmann et al., 2020) showed that deltamethrin reduced the neural network formation in Rat (rNNF) and the oligodendrocyte numbers at fairly low concentrations (BMC_{50} 0.5 and BMC_{30} 0.6 μmol/L, respectively), whereas the effect on human NNF (hNNF) and neuronal crest cells (NCC) migration was seen at much higher concentrations (BMC_{50} 3.9 and BMC_{25} 18.4 μmol/L, respectively). While Na_v are the primary neuronal target of pyrethroid insecticides and play a relevant role in developmental neurotoxicity, the reduction in oligodendrocytes might be related to the high sensitivity of their precursor cells to interference of pyrethroids with Na_v (neurons and oligodendrocyte precursor cells might have differential Na_v subunit expression) and also to pyrethroids-induced oxidative stress. Oligodendrocytes are specifically sensitive towards oxidative stress as they are poor in antioxidant defences and their differentiation and maturation is inhibited by these reactive species (Masjosthusmann et al., 2020).

The results of the systematic review for *in vivo* and *in vitro* endpoints and the outcome of the IVB were used in this Scientific Opinion to postulate and develop a stressor-based AOP network (using deltamethrin as a stressor substance). The AOP conceptual framework gives the opportunity to contextualise the mechanistic information and increase confidence in the *in vivo* AO and provide a plausible mechanistic link to the outcome of the HOS.

3.6.5. Overall uncertainty analysis of the available human observational studies on *in utero* exposure to deltamethrin and neurodevelopmental outcomes in children

For most endpoints and studies, the level of probability for an association was rated 0–10%. Only for the endpoint category communication in a single study (Eskenazi et al., 2018) the probability level was rated 10–33% for an association. No higher probability could be awarded principally because of remaining uncertainties over the appropriateness of the time-point for urine sampling and no assessment of co-exposures.

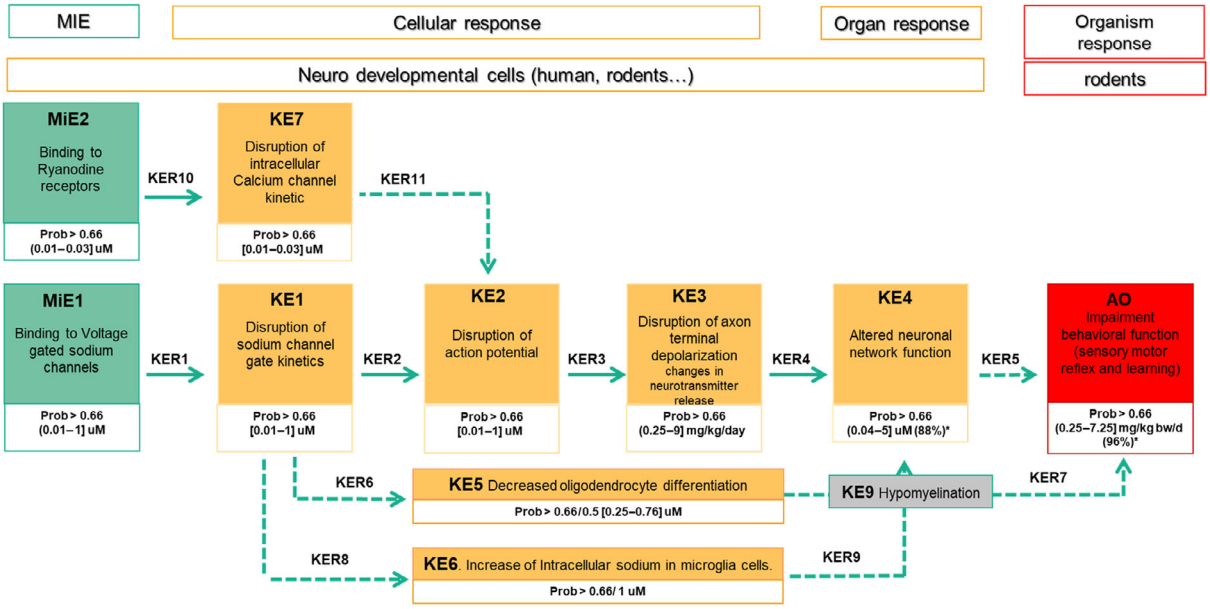
Measurement of the non-specific 3-PBA metabolite does not allow differentiation between exposures that result from deltamethrin or from other pyrethroids. In addition, the presence of 3-PBA in maternal urine samples during gestation does not represent direct evidence of exposure to parent pyrethroid compounds. The metabolite 3-PBA has a longer half-life compared with the parent

compound, it is persistent and refractory to degradation in the natural environment and therefore can accumulate in the environment ultimately reaching agricultural products. Since the measurement of 3-PBA in urine reflects exposure to both parent compounds and non-toxic preformed metabolites in the environment, this exposure metric can result in misclassification of past exposures (Burns and Pastoor, 2018). Therefore, results should be interpreted with caution and this was considered by the working group experts in the RoB and uncertainty analysis.

Despite all these limitations, the available epidemiological evidence linking prenatal exposure to deltamethrin with NDDs is supported by experimental data from animal models of some neurodevelopmental outcomes and consistent with the biology of NDDs. This observation is also supported by the work performed in the context of this Scientific Opinion. The available evidence structured in the AOP conceptual framework indicates that a plausible mechanistic link between exposure to deltamethrin and NDDs exists. However, the current (low) doses of exposure during pregnancy may not be enough for triggering the whole process leading to NDDs and additional co-exposure to other chemicals, environmental factors or maternal lifestyles may be needed, apart from the role of specific genetic mutations in Na_v. However, it can be anticipated that the higher the exposure to deltamethrin the less the level of additional factors to trigger NDDs.

3.7. Postulation of the adverse outcome pathway network and Bayesian network analysis

Deltamethrin is a type II pyrethroid and pyrethroids have been used as insecticides in agricultural and home formulations for decades. The primary mode of pyrethroid action in both insects and mammals is disruption of VGSC function. There is indeed a considerable amount of information supporting the involvement of VGSC in the mode of action of acute pyrethroid neurotoxicity; nevertheless, the potential role of VGSC for DNT of pyrethroids, including deltamethrin, is not well understood. There is no AOP for acute neurotoxicity and DNT; however, Shafer et al. (2005) published a critical review focusing on issues of mode of action and age-dependent neurotoxicity and DNT of pyrethroids as related to risk decisions, highlighting critical KEs which are likely to be common to both kinds of toxicity and for which a comprehensive biological knowledge is available. Therefore, the biological knowledge identified in the systematic review was used to postulate the AOP using deltamethrin as a chemical stressor. The outcome of the DNT-IVB (Masjosthusmann et al., 2020) was included as an additional step for the identification of KEs relevant for neurodevelopmental processes. Eventually, this resulted in a postulation of an AOP network. Mapping of the KEs in the postulated AOP network was performed based on the biological knowledge and through an uncertainty analysis and EKE (see Section 3.5). Figure 9 represents the graphical version of the postulated AOP network and details are given in Annex M. In Figure 10, the AOP was postulated by including the DNT KEs assessed with a probability of higher than 66% to be affected by deltamethrin exposure *in vitro* or *in vivo* during the UA of the full body of evidence for DNT hazard of deltamethrin retrieved and appraised. An exception was made for the endpoint 'neural crest cell migration' that was positive at concentrations too high to be aligned in terms of dose concordance. The relationships between key events (KERs, also including the MIE and AO) were built based on: (1) classification of the evidence as molecular, cellular, organ and organism responses; and (2) scientific knowledge for inferring the probable change in, or state of, a downstream KE from the known or measured state of an upstream KE. Solid lines indicate adjacent KEs for which robust knowledge and information are available; dash lines indicate non-adjacent KEs for which biological plausibility and/or empirical support is more limited. However, some considerations are worth noting here.



Dashed lines indicate non-adjacent KEs for which biological plausibility and/or empirical support is less certain. A round or squared parenthesis indicates that the extreme is excluded or included, respectively.

Figure 10: Postulated AOP network

The string of adjacent KEs from the MIE1 to KE4 represents the most robust pathway because of the considerable amount of biological knowledge and available information for pyrethroids. Data for deltamethrin were obtained from the systematic literature review and the outcome of the DNT-IVB. The biological knowledge and available information for pyrethroids others than deltamethrin are included in the Annex M (see also table 9 in Annex M). Indeed, knowledge and data for other pyrethroids were considered to inform the reliability and relevance of the assay results for deltamethrin. The *in vitro* data for effects of other pyrethroids on the MIE and early KEs is supported by decades of research; however, was not systematically reviewed in this document. Data from later KEs was available from the same reports scrutinised for deltamethrin (e.g. Masjosthusmann et al., 2020; Frank et al., 2017). Inclusion of all available *in vivo* AO data for the broad chemical class of pyrethroids was beyond the scope of this project and would require a systemic review. However, an initial consistency comparison was made with all available OECD 426 studies for others pyrethroids; data on the AO from the OECD 426 studies has been therefore retrieved from EFSA, JMPR and US EPA reports (see table 9 in Annex M). Any additional work will also need to focus on other critical experimental variables, including kinetics modelling of embryo/fetus and pup brain exposure. Although these considerations are making the comparative effort unlikely, there is evidence that consistency exists for the mechanistic data (MiE1, MiE2, KE1 KE2, KE3, KE4, KE5); indeed, the mechanistic characteristics are similar across the pyrethroids chemical class, making the postulated sequence of KEs also relevant for other pyrethroids. Additional work on the full chemical class has the potential to also support the current expert-based probability assignment to the most downstream KERs 5 and 7.

The KER5, linking KE4 (altered neuronal network function) to the AO (impaired behavioural functions) is a non-adjacent KER, meaning that a large uncertainty exists in this relationship. This higher level of uncertainty is based on: 1) a lack of empirical data on possible intermediate KEs between KE4 and the AO and uncertainty on the biological linkage between behavioural-based outcomes and the underlying anatomy and physiology in the brain; and 2) the practical limitations and uncertainty inherent to *in vivo* AO measurements, which are difficult to link to human AOs because of the limited translational value of neurodevelopmental animal models (see Annex M, Table 5.5.-1 KER5 and KER7). However, there is consensus that chemical-mediated alterations of cellular events critical to normal neurodevelopment can result in adverse neurological development (Bal-Price et al., 2018; Masjosthusmann et al., 2020; Mundy et al., 2015). Furthermore, some regulators stressed that behavioural endpoints observed in rat studies only provide an indication of a potential general DNT in humans rather than any more specific functional impairment and these indicators cannot be comprehensive for humans. Conceptually the same appears to be true for *in vitro* KE4 data (Paparella

et al., 2020). Adverse effects on oligodendrocyte differentiation (KE5) are well documented using *in vitro* methods; however, the relationship between this KE and the AO of altered behavioural function is more uncertain for the same reasons given above. Because it is biologically plausible that additional KEs would occur between impaired oligodendrocyte differentiation (KE5) and the AO, a shadow KE (KE9, hypomyelination) was added to the AOP network. This is because oligodendrocyte precursors are the predominant form of the oligodendroglial lineage in human cerebral white matter until 28 weeks of gestation, the intrinsic vulnerability of these cells is considered as central to the pathogenesis of periventricular leukomalacia observed in some premature infants that later develop cognitive and behavioural deficits (Masjosthusmann et al., 2020; Rezaie and Dean, 2002). However, although this is biologically plausible, no information was gathered for deltamethrin and this KE was not considered further.

Data and knowledge for other non-adjacent KERs are available for deltamethrin, resulting in additional putative AOP strings; however, information from other pyrethroids supporting these strings is lacking and a dedicated systematic review would be necessary.

Overall, the development of this AOP network identified two major uncertainties. The first one lies in the network itself as potentially important KEs could be missed. This is due to the decision to include only data and KEs that have been studied for the stressor deltamethrin. This can be reduced by including additional KEs and KERs from stress-agnostic biological knowledge. The second is due to a knowledge gap that drives uncertainty in the last two KERs: the lack of empirical data that provide correlative and/or causal relationships between disturbed neuronal network function (KE4) and/or hypomyelination (as a result of KE5) and altered behavioural function (AO). The knowledge gap for linking KE4 and the AO may be more difficult to address and has been identified in several other AOPs with adverse neurodevelopmental outcomes. This could be partly addressed by considering the limitations and uncertainties inherent to *in vivo* AO measurements and *in vitro* KE testing the kinetic extrapolation and human exposure (Paparella et al., 2020; Paul-Friedman et al., 2019).

To characterise the certainty within this putative AOP network and the KER, a BN approach was developed based on available data and expert knowledge for deltamethrin. More details on the BN approach and all the results of the elicitation process can be retrieved in Annex B and M. The outcome of the BN was also used to give a probabilistic quantification of the overall WoE for the postulated AOP network and to provide a quantitative estimation of the contribution of each event in the AOP to reduce the uncertainty in the AO occurrence. This allowed the estimation of the impact of the mechanistic knowledge within the putative AOP network.

In line with the methodologies chapter (see Section 2.2.5), several types of probabilities associated with the BN structure can be used to infer conclusions on the KEs and the triggering stressor: 1) conditional probability distributions (CPDs) for individual or combined KERs, 2) marginal probability estimation for the AO to occur, 3) impact of KEs on the marginal probability for the AO to occur and 4) joint probability estimation for all KEs being activated.

An expert judgement approach was adopted to estimate the parameters of the CPDs. This choice was made necessary by the lack of large data sets providing empirical data for large portions and/or the whole pathway. A summary of the CPDs is reported in Figure 11.

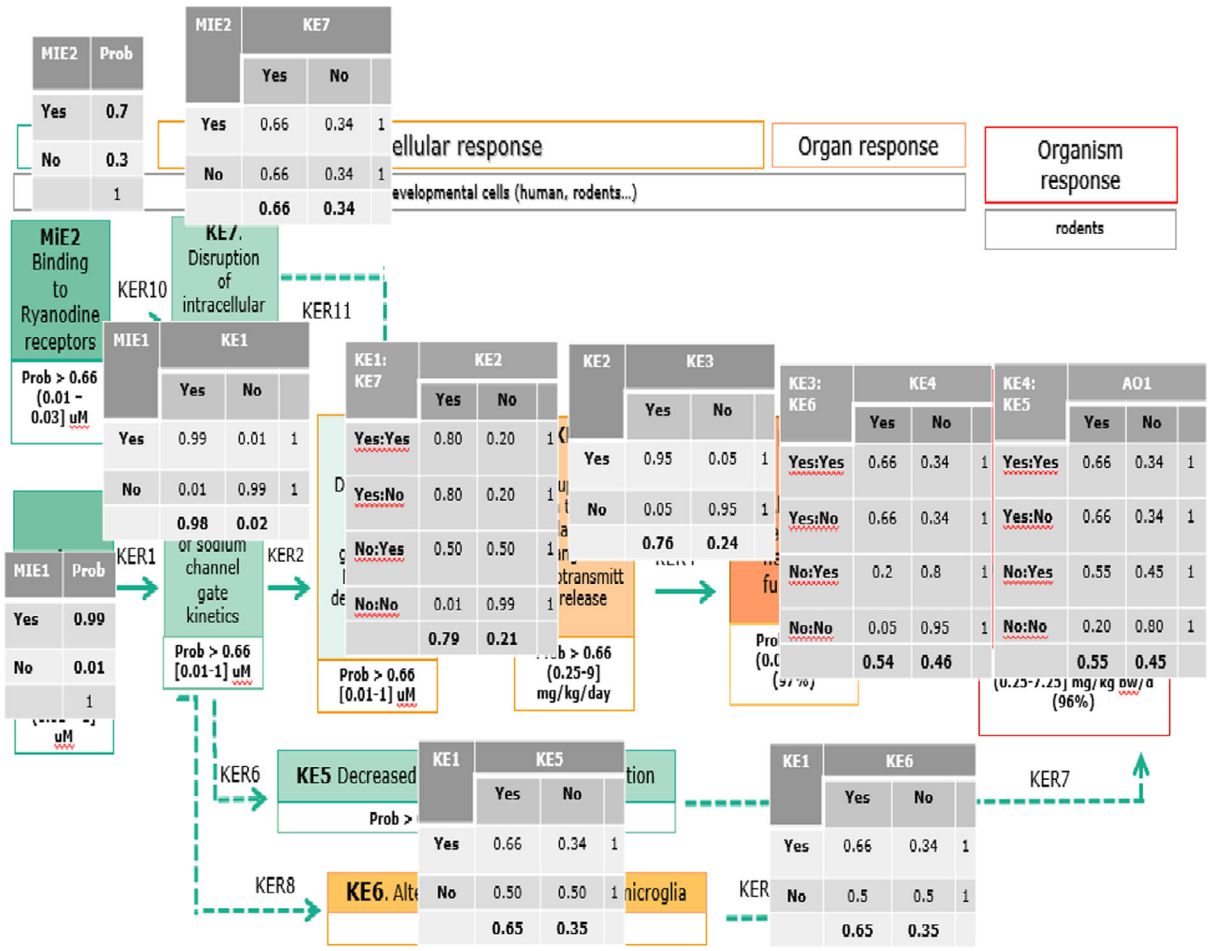


Figure 11: Conditional probability distributions in the postulated AOP network. Bold numbers in each table represent marginal probabilities

The analysis indicates that the evidence for KER5 and KER7, leading to the AO 'altered behavioural function' in rodents, contains relatively high uncertainty within the network. This is in the same range (0.66) as for the other non-adjacent KERs (e.g. KER9 and KER4), but clearly lower compared with the conditioned probabilities for KERs within the string describing the best established KEs for deltamethrin and pyrethroids, i.e. binding to VGSC ultimately leading to inhibition of neuronal network function (MIE1-KEs1-2-3-4).

Integrating the expert-derived conditional probabilities over all KERs by using the BN approach provides marginal probabilities for the AO to be activated/not activated. In this specific stressor-based AOP, those probabilities are derived under the assumption that exposure to deltamethrin occurred. The marginal probability distributions for all the MIEs/KEs/AO in the AOP network are reported in Table 9.

Table 9: Marginal probabilities for all the MIEs/KEs/AO in the network

MIE/KE/AO		Probability	
		To be activated	**To be not activated**
MIE1	Binding to VGSC	0.99	0.01
MIE2	Binding to ryanodine receptors	0.7	0.3
KE1	Disruption of sodium channel gate kinetics	0.98	0.02
KE2	Disruption of action potential generation; membrane depolarisation	0.79	0.21
KE3	Disruption of axon terminal depolarisation; changes in neurotransmitter release	0.76	0.24

MIE/KE/AO		Probability	
		To be activated	To be not activated
KE4	Altered neuronal network function	0.54	0.46
KE5	Decreased oligodendrocyte differentiation	0.66	0.34
KE6	Increase of intracellular sodium in microglia cells	0.66	0.34
KE7	Disruption of intracellular Ca channel kinetics	0.66	0.34
AO	Impairment behavioural function	0.55	0.45

The analysis of the marginal probabilities gives an indication on the uncertainty propagation along the pathway since the probability to be activated is generally lower for KEs farer from the MIEs. It also indicates that all the KEs as well as the AO have a probability to be activated greater than 0.5 leading to the conclusion that their activation is more probable than not when exposure to deltamethrin occurs.

It should be noted that in some cases (e.g. AO), the resulting marginal probability is lower than 0.66. This is because the initial probability of 0.66 was used to screen KE that could be considered in the AOP. To define the final AOP structure though, the working group experts had to consider additional sets of evidence to support the KER. The additional evidence was used to assess the three criteria required in the AOP framework (biological plausibility, essentiality, empirical evidence). Initial certainty on the causal association of KEs with deltamethrin is not the same as certainty in the KERs in the AOP sequence, as also reflected in the estimation of the marginal probability.

Furthermore, by assessing the impact of generating firm evidence for any of the KEs (setting probability of occurring at 1 or 0) on the marginal probabilities for the AO, it can be shown that the most downstream KEs are most influential, i.e. KE4 (altered neuronal network function), followed by KE5 (decreased oligodendrocyte differentiation). Therefore, in the context of an IATA, testing these KEs appears most important.

The conditional probabilities for the individual KEs to occur under the condition that the upstream KEs were activated, were used to estimate the joint probability that all events (MIEs, KEs, AO) in the network are concurrently activated/occur. This probability is 6.5%, which is well above the average conditional probability of 0.76% (see Annex B and M for details). This was achieved by resorting to the Bayesian theorem and property of the nodes in a BN to be conditionally independent from all non-descendent nodes given the ascendant nodes (conditioning nodes). Since the joint probability for all KEs and the AO to be activated within a network depends also on the number of nodes within the network, it is meaningful to consider the conditional probabilities for KE being activated given the connected upstream KEs also being active, averaged over all KEs in the network. The estimate of the joint probabilities with inclusion of number of nodes and per-node average conditional probability of the downstream KEs to occur given the activation of the connected upstream KE(s) is reported in Table 10.

Table 10: Probabilities for all KEs and the AO being activated, number of nodes and per-node average conditional probably of the downstream KEs to occur given the activation of the connected upstream KE(s) for the AOP network and the linear strings

	AOP network	AOP1: MEI1-KEs1-2-3-4-AO	AOP2: MEI2-KEs7-2-3-4-AO	AOP3: MIE1-KEs1-5-AO	AOP4: MIE1-KEs1-6-4-Ao	MIE1-KEs1-2-3-4
Joint probability	6.53	32.45	15.18	39.41	23.5	49.16
Average conditional prob/node	0.76	0.83	0.73	0.79	0.75	0.87
Number of. Nodes*	10	6	6	4	5	5

*: In a Bayesian network (BN) node represents random variables.

The analysis shows that the mechanistic knowledge gathered in this putative AOP network helps explain and support the experimental observations for deltamethrin (all KEs and AO positive) to the following extent: (1) for the complete putative network to a moderate degree (0.76 probability per

node), (2) for the AOP string including only the best documented KERs for deltamethrin (MIE1-KEs1-2-3-4 AO), to a higher degree (0.83 probability per node) and (3) for the best established string supporting that deltamethrin and other pyrethroids affect VGSC ultimately leading to altered network function (KE5) to the highest degree (0.87 probability per node). The latter information would be relevant if KE4 (neuronal network formation) was recognised as immediately useful for regulatory toxicology.

3.8. IATA case studies

The systematic review, the integration of the DNT-IVB and the postulation of the AOP network culminated in the IATA. Indeed, the main purpose of developing the IATA case studies in this Scientific Opinion was to show the applicability of an IVB for DNT in the context of the European Pesticide Regulation (EU) 283/2013 and 1107/2009 (European Commission, 2009, 2013). The two case studies support the validity of the approach taken; therefore, the IATA addresses the ToR proposed by the EFSA PPR Panel. The active substances used in the case studies have a diversified qualitative toxicological profile, one being a known neurotoxic chemical with a neurotoxic pesticidal mechanism (deltamethrin) while the second one is not (flufenacet). Therefore, the two case studies allow a comparative evaluation of the DNT-IVB, while assessing the impact of the mechanistic understanding in two very different scenarios.

For the case study using flufenacet, an OECD TG 426 study was available and conducted to dismiss a DNT concern because of the observation of clinical signs possibly of a neurotoxic nature observed in a non-rodent species and changes in thyroid hormones observed in rodents (Flufenacet Draft Renewal Assessment Report prepared according to the Commission Regulation (EU) No 1107/2009, 2017). The IATA case study indicates that, based on all the evidence from the literature and the DNT-IVB results, and accounting for the uncertainties, no DNT hazard can be identified from the mechanistic evidence retrieved, supporting the conclusion that flufenacet is not a developmental neurotoxicant. This conclusion is aligned with a lack of adverse DNT outcomes in the available *in vivo* study. *In vitro*, flufenacet is affecting oligodendrocyte differentiation in a non-specific manner, meaning that the observed effect at the highest concentration cannot be clearly separated from a cytotoxic effect.

Since deltamethrin caused acute and repeated neurotoxic effects after oral administration, a regulatory guideline study (OECD TG 426) was conducted for deltamethrin. However, other *in vivo* experimental studies available in the open literature yielded apparent equivocal results compared with the OECD TG 426 study, although the study design and the doses administered differed between them. Moreover, epidemiological data showed a potential DNT concern for the pyrethroid chemical class. This led to the hypothesis that the integration of mechanistic data, using *in vitro* assays relevant for a normal neurodevelopment, could reduce the uncertainty on the DNT hazard characterisation of deltamethrin.

The core element of the case studies, particularly that for deltamethrin, was an integrated and extensive uncertainty analysis of *in vivo* and *in vitro* lines of evidence with the HOS used as supporting evidence. This allowed for the postulation of an AOP network only for deltamethrin case study. Since no AOP was available for inhibition of VGSC, an AOP network was developed. This consisted of two MIEs leading to altered behavioural function which included data in which, based on working group experts' consensus judgement, there was more than a 66% probability of a causal association with the MiE/KEs/AO, together with a range/probability distribution expressing the uncertainty on the lowest concentration/dose triggering the causal relationship. The strength of the KER was quantified using a Bayesian network analysis approach. It was found that the marginal probability of any MIE/KE/AO to occur in an AOP string in the AOP network was always above 0.50, indicating that it is more likely than not that the MIE, KE or AO would be triggered by deltamethrin at the triggering concentration/dose. A gap, driving the uncertainty, was the lack of empirical support providing biological understanding of what is occurring between the last KEs immediately before the AO. How critical the knowledge gap is, may be contextualised by considering also the uncertainties in the AO measurement and the ultimate use of the *in vitro* data. However, the dose/concordance of the AOP was underpinned by applying a physiologically based pharmacokinetic (PBPK) model for deltamethrin and performing reverse dosimetry (see Annex M). Furthermore, it contextualised the equivocal *in vivo* effects among different studies by indicating that only direct dosing of pups (as opposed to exposure via dams) would achieve exposure high enough to trigger the AOP network. DNT was observed also with gavage exposure of dams only during gestation, but no kinetic data are available as to estimate the achieved embryo/fetal brain concentrations. Nevertheless, the assessment of the kinetics in the study in which deltamethrin

was administered to pups by gavage can support regulatory decision making, since it established the certainty of a dose range that potentially can cause the adverse outcome.

Overall, the case study shows the applicability of the DNT-IVB for hazard identification and characterisation and, in the context of these case studies, its specificity and sensitivity. Furthermore, the case study illustrates the usefulness of the postulation of an AOP network and probabilistic quantification of the WoE with the potential of aiding regulatory decision making. The overall process and the mechanistic understanding also increased the ability of interpreting the HOS by providing a plausible mechanistic link to a human-relevant AO, thus supporting the contextualisation of these studies in the risk assessment process.

The IATA case studies are reported in Annexes N and O. Although a standard template for IATA is not available, the case studies were developed using the template recommended by OECD, which was adapted when needed.

3.9. Deviations from the protocol

For *in vivo* and *in vitro* studies, the tier of RoB was used as an exclusion criterion for data extraction and data analysis (i.e. data extraction was performed only for studies allocated in Tiers 1 and 2 of the overall RoB).

Two deviations from the protocol have been applied to the elicitation process used to assess the uncertainty in the questions on the association between exposure to deltamethrin and DNT endpoints (specific endpoints or category of endpoints) (table 15 in Annex B). For all the lines of evidence, except human studies, the threshold for the probability used to express uncertainty in the answer on the causal association was 0.66 (as double as probable as not) instead of 0.60. This level was chosen since it is easier to communicate to the working group experts as suggested by the EFSA Standing Working Group on Uncertainty.

Contrary to what was planned, for human evidence, the uncertainty on the association was expressed using the approximate probability scale (0–10, 10–33, 33–50, 50–66, 66–100%) instead of a full probability distribution. This was considered by the working group experts better reflecting the level of accuracy that was possible to achieve in the assessment of the uncertainty.

In the protocol it was planned to elicit full probability distributions using the Roulette method for the assessment of the lowest concentration/dose triggering a KE/AO. This method was actually applied only for two KE/AO (MEA and behavioural outcomes). For the remaining MIEs/KEs, a range of possible values (under the assumption that they were equally probable) was agreed by the working group experts based on the consideration that uncertainty was not so large and the additional value of having a full probability distribution would not have been worth the effort.

4. Conclusions

The ToR proposed by the EFSA PPR Panel were addressed in this Scientific Opinion by applying an AOP-informed IATA approach. This was carried out using a transparent and trackable evidence-based approach with a probabilistic quantification of the WoE. To ensure the transparency and reliability of the assessment for decision making, an extensive uncertainty analysis was performed, covering from the methods' development to the limitations of the conclusions drawn. A key step of the IATA iterative process was the inclusion of mechanistic information obtained by the use of NAMs. This step was accomplished by maximising the outcome of the external scientific report committed by EFSA (Masjosthusmann et al., 2020) and the inclusion of the DNT-IVB in the AOP-informed IATA process. Therefore, a postulated AOP network was developed for integrating different lines of evidence to support causal relationship between exposure to deltamethrin, one of the prototypical chemicals used in the case studies, and a DNT-related AO. For deltamethrin, the quantitative WoE assessment of the AOP network using a BN analysis resulted in an acceptable level of certainty in the AOP network proposed for the assessment. The approach taken in this IATA case study allowed conclusions to be drawn with an acceptable level of certainty in DNT hazard identification and characterisation of deltamethrin. The assessment indicates that there is robust evidence and experimental support for the sequence of adjacent KEs leading to KE4 (neuronal network formation). However, there are still gaps of knowledge and uncertainties on the relationship between KE4 and DNT-related AOs. The inclusion of additional biologically plausible KEs will decrease the level of uncertainty in the downstream KER.

For flufenacet, the approach taken (evidence-based assessment of all the evidence and inclusion of the IVB through an AOP-informed IATA framework) allowed conclusions to be made that flufenacet is

not a developmental neurotoxicant. The inclusion of the mechanistic information, mostly derived from the DNT-IVB, provides an example for the use and application of the DNT-IVB for single substance with a non-neurotoxic mode of action and supports the sensitivity of the DNT-IVB.

Importantly, it is evident from the uncertainty analysis of AO and its KERs with the upstream KEs that the use of NAMs will need to be based primarily on biological plausibility and data correlation to apical endpoints from the rodent tests should be taken with caution. A thorough analysis of this relationship will allow moving forward towards next-generation risk assessment. This is especially true for regulatory fields including complex endpoints, like DNT, with relatively few available and reliable *in vivo* data and practical limitations for their generation and validation. The EFSA PPR Panel therefore concluded that the two IATA cases studies should be included in the OECD guidance on use and interpretation of DNT *in vitro* assays and that the methodologies applied in their development were appropriate to address the ToR and the regulatory problem formulation.

5. Recommendations

For the AOP-informed IATA methodology:

- IThe stressor-based approach experience used for the postulation of the AOP network should be expanded to the pyrethroid chemical class, including evidence for the DNT AO when available. It is also recommended to include more chemically agnostic, biologically plausible KEs, to reduce the uncertainty linking downstream KEs with the DNT AO.
- The PPR Panel recommends submitting the postulated AOP/AOPs to the OECD AOP programme to further support a regulatory uptake. In doing this, the PPR Panel recommends that the uncertainties identified in the postulated AOP network should be further explored and possibly resolved by including additional KEs based on biological knowledge of the pathway.
- A transparent approach should be applied in developing IATA using an evidence-based approach and a quantitative estimation of the WoE through a probabilistic expression of the uncertainties.
- The exposure estimation to pyrethroid pesticides during nervous system development in HOS should be improved by using reliable methods (i.e. measuring the active substance or its specific metabolite/s in blood or biomarkers in urine at several times of development during pre and postnatal periods, and application of PBPK models to biomonitoring data (Pletz et al., 2020)) to decrease the uncertainty on the current exposure levels and on the causal association with the DNT outcomes.
- The AOP-informed IATA was considered a valuable approach for the DNT hazard characterisation. Paucity of available AOPs is a limitation and development of DNT-related AOPs is recommended. Here, the DNT-IVB data could serve as starting points for inside-out AOP developments.
- The IATA case study indicates the relevance of kinetic data. PBPK models are cornerstones to describe the relationship between external exposure and target tissue dose. Pharmacokinetic (PK) information in animal models are therefore critical and recommended for a correct contextualisation of the dose concordance assessment in the AOP-informed IATA and for moving from the step of hazard to risk characterisation. For DNT, this should include the relationship between maternal and fetal compartments and the involvement of oral/lactation, inhalation and dermal exposure of the new-born.

For the implementation of the DNT-IVB in the risk assessment:

- The disruption of the neuronal network formation has a high impact in the context of this postulated AOP network. Relevance and reliability of the assays is supported by similarity in the outcome using different test systems. Considering that one of the assays is based on a human cell system, the PPR Panel recommends testing more chemicals, including additional pyrethroids, in the human-derived cell system to confirm the outcomes observed with the rat *in vitro* assay.
- The DNT-IVB indicated that the oligodendrocyte differentiation/maturation assay is a sensitive endpoint for which specificity is still uncertain. Data on this endpoint for pesticides and chemicals in general is sparse but biological plausibility for a DNT AO is high. It is therefore recommended that additional work should be performed to characterise this relevant endpoint, particularly for modes of action converging on oligodendrocyte toxicity.

- It is recommended to map the results of the DNT-IVB vs the available DNT *in vivo* data for pesticide active substances using a systematic approach aiming to assess the applicability domain of the IVB versus the current regulatory accepted *in vivo* study.
- Considering the strength of the mechanistic understanding in the IATA case studies, the inclusion of a DNT-IVB is recommended for DNT hazard characterisation of pesticide active substances. This can also be triggered when concerns exist from *in vivo* experiments or HOS. Similarly, the outcome of the DNT-IVB should be considered as a trigger for additional DNT investigations. The PPR Panel also supports to develop a GD explaining the uses and interpretation of the DNT-IVB data for regulatory toxicology assessments in line with the OECD programme. For this purpose, available information on the limitations and uncertainties of animal tests assessing DNT should be considered to the same degree as for the IVB.
- The PPR Panel considers that the potential of NAMs should be fully exploited to allow efficient regulation with a long-term sustainable socioeconomic return of investment in terms of environmental health protection. Therefore, NAMs should be used in next-generation risk assessment (OECD GD 275; OECD, 2017; Dent et al., 2018; Desprez et al., 2018; Smith et al. 2016; Hatherell et al., 2020)
- The use of corrected nominal concentrations in the *in vitro* testing should be considered for estimating the effective concentration to be used for exposure extrapolation. The uncertainty analysis provided in this scientific opinion for the *in vitro* test indicated that several points should be considered. These include, but are not limited to, predicted or empirical data on partitioning of the chemical with plastic, lipid and protein, intracellular concentration and accumulation where the chemicals are added multiple times to the test system.
- Based on these principles, implementation of the DNT-IVB methods for DNT endpoints assessment as a potential data requirement for pesticide active substances, and possibly for other chemicals, should be discussed with risk managers and stakeholders in the appropriate context.

6. Documentation as provided to EFSA

Establishment of an *a priori* protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of developmental neurotoxicity. October 2020. EFSA external report available online at: https://www.efsa.europa.eu/en/supporting/pub/en-1938

References

ACE (America's Children and the Environment), 2019. Health: Neurodevelopmental Disorders – Report Contents. 3rd Edition, Updated in August 2019. Available online: https://www.epa.gov/sites/production/files/2019-07/documents/ace3-neurodevelopmental-updates_0.pdf

APS (American Psychiatric Association), 2013. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition. American Psychiatric Association, Arlington, Virginia.

Aschner M, Ceccatelli S, Daneshian M, Fritsche E, Hasiwa N, Hartung T, Hogberg HT, Leist M, Li A, Mundy WR, Padilla S, Piersma AH, Bal-Price A, Seiler A, Westerink RH, Zimmer B and Lein PJ, 2017. Reference compounds for alternative test methods to indicate developmental neurotoxicity (DNT) potential of chemicals: example lists & criteria for their selection & use. Altex.

Bal-Price A and Fritsche E, 2018. Developmental neurotoxicity. Toxicology and Applied Pharmacology, 354, 1. https://doi.org/10.1016/j.taap.2018.07.016

Bal-Price A, Hogberg HT, Crofton KM, Daneshian M, FitzGerald RE, Fritsche E, Heinonen T, Hougaard Bennekou S, Klima S, Piersma AH, Sachana M, Shafer TJ, Terron A, Monnet-Tschudi F, Viviani B, Waldmann T, Westerink RHS, Wilks MF, Witters H, Zurich M-G and Leist M, 2018. Recommendation on test readiness criteria for new approach methods (NAM) in toxicology: exemplified for developmental neurotoxicity (DNT). ALTEX – Alternatives to Animal Experimentation, 35, 306–352. https://doi.org/10.14573/altex.1712081

Barkoski JM, Philippat C, Tancredi D, Schmidt RJ, Ozonoff S, Barr DB, Elms W, Bennett D and Hertz-Picciotto I, 2021. In utero pyrethroid pesticide exposure in relation to autism spectrum disorder (ASD) and other neurodevelopmental outcomes at 3 years in the MARBLES longitudinal cohort. Environmental Research, 194. https://doi.org/10.1016/j.envres.2020.110495

Blair A, Stewart P, Lubin JH and Forastiere F, 2007. Methodological issues regarding confounding and exposure misclassification in epidemiological studies of occupational exposures. American Journal of Industrial Medicine, 50, 199–207. https://doi.org/10.1002/ajim.20281

Burns CJ and Pastoor TP, 2018. Pyrethroid epidemiology: a quality-based review. Critical Reviews in Toxicology, 48, 297–311. https://doi.org/10.1080/10408444.2017.1423463

Cardoso AR, Lopes-Marques M, Silva RM, Serrano C, Amorim A, Prata MJ and Azevedo L, 2019. Essential genetic findings in neurodevelopmental disorders. Human Genomics, 13, 31. https://doi.org/10.1186/s40246-019-0216-4

Carlsson T, Molander F, Taylor MJ, Jonsson U and Bölte S, 2020. Early environmental risk factors for neurodevelopmental disorders – a systematic review of twin and sibling studies.Development and Psychopathology, 2020, 1–48. https://doi.org/10.1017/S0954579420000620. Epub ahead of print.

Cheroni C, Caporale N and Testa G, 2020. Autism spectrum disorder at the crossroad between genes and environment: contributions, convergences, and interactions in ASD developmental pathophysiology. Molecular Autism, 11, 69. https://doi.org/10.1186/s13229-020-00370-1

Christensen K, Christensen CH, Wright JM, Galizia A, Glenn BS, Siegel Scott C, Mall J, Bateson TF, Murphy PA and Cooper GS, 2015. The use of epidemiology in risk assessment: challenges and opportunities. Human and Ecological Risk Assessment, 21, 1644–1663. https://doi.org/10.1080/10807039.2014.967039

Dalsager L, Fage-Larsen B, Bilenberg N, Jensen TK, Nielsen F, Kyhl HB, Grandjean P and Andersen HR, 2019. Maternal urinary concentrations of pyrethroid and chlorpyrifos metabolites and attention deficit hyperactivity disorder (ADHD) symptoms in 2–4-year-old children from the Odense Child Cohort. Environmental Research, 176. https://doi.org/10.1016/j.envres.2019.108533

Del Pino I, Rico B and Marín O, 2018. Neural circuit dysfunction in mouse models of neurodevelopmental disorders. Current Opinions in Neurobiology, 48, 174–182. https://doi.org/10.1016/j.conb.2017.12.013

Dent M, Amaral RT, Da Silva PA, Ansell J, Boisleve F, Hatao M and Kojima H, 2018. Principles underpinning the use of new methodologies in the risk assessment of cosmetic ingredients. Computational Toxicology, 7, 20–26.

Desprez B, Dent M, Keller D, Klaric M, Ouédraogo G, Cubberley R and Mahony C, 2018. A strategy for systemic toxicity assessment based on non-animal approaches: the Cosmetics Europe Long Range Science Strategy programme. Toxicology in Vitro, 50, 137–146.

EFSA (European Food Safety Authority), 2014. Guidance on Expert knowledge elicitation in food and feed safety risk assessment. EFSA Journal 2014;12(6):3734, 278 pp. https://doi.org/10.2903/j.efsa.2014.3734

EFSA and EBTC (European Food Safety Authority and Evidence-Based Toxicology Collaboration), 2018. EFSA Scientific Colloquium 23: Evidence integration in risk assessment: the science of combining apples and oranges. EFSA supporting publication 2018;16(3):EN-1396, 28 pp. https://doi.org/10.2903/sp.efsa.2018.EN-1396

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2013. Scientific Opinion on the developmental neurotoxicity potential of acetamiprid and imidacloprid. EFSA Journal 2013;11(12):3471, 47 pp. https://doi.org/10.2903/j.efsa.2013.3471

EFSA Scientific Committee, Benford D, Halldorsson T, Jeger MJ, Knutsen HK, More S, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, Schlatter JR, Silano V, Solecki R, Turck D, Younes M, Craig P, Hart A, Von Goetz N, Koutsoumanis K, Mortensen A, Ossendorp B, Germini A, Martino L, Merten C, Smith A and Hardy A, 2018. Principles and methods behind EFSA's Guidance on Uncertainty Analysis in Scientific Assessment. Scientific Opinion. EFSA Journal 2018;16(1):5122, 282 pp. https://doi.org/10.2903/j.efsa.2018.5122

EFSA Scientific Committee, More S, Bambidis V, Benford D, Bragard C, Hernandez-Jerez A, Bennekou SH, Koutsoumanis K, Machera K, Naegeli H, Nielsen SS, Schlatter JR, Schrenk D, Silano V, Turck D, Younes M, Fletcher T, Greiner M, Ntzani E, Pearce N, Vinceti M, Ciccolallo L, Georgiadis M, Gervelmeyer A and Halldorsson TI, 2020. Draft for internal testing Scientific Committee guidance on appraising and integrating evidence from epidemiological studies for use in EFSA's scientific assessments. EFSA Journal 2020;18(8):6221, 83 pp. https://doi.org/10.2903/j.efsa.2020.6221

EFSA (European Food Safety Authority), Martino L, Aiassa E, Halldórsson TI, Koutsoumanis PK; Naegeli H, Baert K, Baldinelli F, Devos Y, Lodi F, Lostia A, Manini P, Merten C, Messens W, Rizzi V, Tarazona J, Titz A and Vos S, 2020. Draft framework for protocol development for EFSA's scientific assessments. EFSA supporting publication 2020;EN-1843, 46 pp. https://doi.org/10.2903/sp.efsa.2020.EN-1843

Elwan MA, Richardson JR, Guillot TS, Caudle WM and Miller GW, 2006. Pyrethroid pesticide-induced alterations in dopamine transporter function. Toxicology and Applied Pharmacology, 211, 188–197. https://doi.org/10.1016/j.taap.2005.06.003

Eskenazi B, An S, Rauch SA, Coker ES, Maphula A, Obida M, Crause M, Kogut KR, Bornman R and Chevrier J, 2018. Prenatal exposure to DDT and pyrethroids for malaria control and child neurodevelopment: the VHEMBE cohort, South Africa. Environmental Health Perspective, 126. https://doi.org/10.1289/ehp2129. [Erratum. In: Environmental Health Perspective, 2020, 128, 39001.].

European Commission, 2009. Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Communities, 309, 1–50.

European Commission, 2013. Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. Official Journal of the European Union, L 93/2.

Fedak KM, Bernal A, Capshaw ZA and Gross S, 2015. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. Emerging Themes in Epidemiology, 12, 14. https://doi.org/10.1186/s12982-015-0037-4

Fluegge KR, Nishioka M and Wilkins III JR, 2016. Effects of simultaneous prenatal exposures to organophosphate and synthetic pyrethroid insecticides on infant neurodevelopment at three months of age. Journal of Environmental Toxicology in Public Health, 1, 60–73. https://doi.org/10.5281/zenodo.218417

Frank CL, Brown JP, Wallace K, Mundy WR and Shafer TJ, 2017. From the cover: developmental neurotoxicants disrupt activity in cortical networks on microelectrode arrays: results of screening 86 compounds during neural network formation. Toxicological Sciences, 160, 121–135. https://doi.org/10.1093/toxsci/kfx169

Fritsche E, 2017. Report on integrated testing strategies for the identification and evaluation of chemical hazards associated with the developmental neurotoxicity (DNT), to facilitate discussions at the joint EFSA/OECD workshop on DNT. Series on Testing and Assessment No. 261 ENV/JM/MONO (2017) 4/ANN1 (2017).

Fritsche E, Alm H, Baumann J, Geerts L, Håkansson H, Masjosthusmann S and Witters H, 2015. Literature review on *in vitro* and alternative developmental neurotoxicity (DNT) testing methods. EFSA Supporting Publication 2015;12(4):EN-778, 186 pp. https://doi.org/10.2903/sp.efsa.2015.EN-778

Fritsche E, Crofton KM, Hernandez AF, Hougaard Bennekou S, Leist M, Bal-Price A, Reaves E, Wilks MF, Terron A, Sachana M and Gourmelon A, 2017. OECD/EFSA workshop on developmental neurotoxicity (DNT): the use of non-animal test methods for regulatory purposes. ALTEX – Alternatives to Animal Experimentation, 34, 311–315. https://doi.org/10.14573/altex.1701171

Furlong MA, Barr DB, Wolff MS and Engel SM, 2017. Prenatal exposure to pyrethroid pesticides and childhood behavior and executive functioning. Neurotoxicology, 62, 231–238. https://doi.org/10.1016/j.neuro.2017.08.005

Grandjean P and Landrigan PJ, 2014. Neurobehavioural effects of developmental toxicity. Lancet Neurology, 13, 330–338. https://doi.org/10.1016/S1474-4422(13)70278-3

Harrill JA, Freudenrich T, Wallace K, Ball K, Shafer TJ and Mundy WR, 2018. Testing for developmental neurotoxicity using a battery of *in vitro* assays for key cellular events in neurodevelopment. Toxicology and Applied Pharmacology, 354, 24–39. https://doi.org/10.1016/j.taap.2018.04.001

Hatherell S, Baltazar MT, Reynolds J, Carmichael PL, Dent M, Li H and Middleton AM, 2020. Identifying and characterizing stress pathways of concern for consumer safety in next-generation risk assessment. Toxicological Sciences, 176, 11–33.

Hill AB, 1965. The environment and disease: association or causation? Proceedings of the Royal Society of Medicine, 58, 295–300.

Hossain MM, Belkadi A, Al-Haddad S and Richardson JR, 2020. Deltamethrin exposure inhibits adult hippocampal neurogenesis and causes deficits in learning and memory in mice. Toxicological Sciences, 178, 347–357. https://doi.org/10.1093/toxsci/kfaa144

Hsueh YP, 2019. Synaptic formation, neural circuits and neurodevelopmental disorders controlled by signaling, translation, and epigenetic regulation. Development Neurobiology, 79, 2–7. https://doi.org/10.1002/dneu.22655

Krause J, 2008. SPECT and PET of the dopamine transporter in attention-deficit/hyperactivity disorder. Expert Review of Neurotherapeutics, 8, 611–625. https://doi.org/10.1586/14737175.8.4.611

Kung TS, Richardson JR, Cooper KR and White LA, 2015. Developmental deltamethrin exposure causes persistent changes in dopaminergic gene expression, neurochemistry, and locomotor activity in zebrafish. Toxicological Sciences, 146, 235–243. https://doi.org/10.1093/toxsci/kfv087

Magby JP and Richardson JR, 2017. Developmental pyrethroid exposure causes long-term decreases of neuronal sodium channel expression. Neurotoxicology, 60, 274–279. https://doi.org/10.1016/j.neuro.2016.04.002

Makris SL, Raffaele K, Allen S, Bowers WJ, Hass U, Alleva E, Calamandrei G, Sheets L, Amcoff P, Delrue N and Crofton KM, 2009. A retrospective performance assessment of the developmental neurotoxicity study in support of OECD test guideline 426. Environmental Health Perspectives, 117, 17–25. https://doi.org/10.1289/ehp.11447

Masjosthusmann S, Blum J, Bartmann K, Dolde X, Holzer A-K, Stürzl L-C, Hagen KE, Förster N, Dönmez A, Klose J, Pahl M, Waldmann T, Bendt F, Kisitu J, Suciu I, Hübenthal U, Mosig A, Leist M and Fritsche E, 2020. Establishment of an a priori protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of developmental neurotoxicity. EFSA supporting publication 2020;EN-1938, 152 pp. https://doi.org/10.2903/sp.efsa.2020.EN-1938

Mundy WR, Padilla S, Breier JM, Crofton KM, Gilbert ME, Herr DW, Jensen KF, Radio NM, Raffaele KC, Schumacher K, Shafer TJ and Cowden J, 2015. Expanding the test set: chemicals with potential to disrupt mammalian brain development. Neurotoxicolgy and Teratology, 52(Pt A), 25–35. https://doi.org/10.1016/j.ntt.2015.10.001

Nachman KE, Fox MA, Sheehan MC, Burke TA, Rodricks JV and Woodruff TJ, 2011. Leveraging epidemiology to improve risk assessment. Open Epidemiology Journal, 4, 3–29. https://doi.org/10.2174/1874297101104010003

Nakai N, Takumi T, Nakai J and Sato M, 2018. Common defects of spine dynamics and circuit function in neurodevelopmental disorders: a systematic review of findings from *in vivo* optical imaging of mouse models. Frontiers in Neuroscience, 12, 412. https://doi.org/10.3389/fnins.2018.00412.

NTP (National Toxicology Program), 2015. OHAT risk of bias rating tool for human and animal studies. Office of Health Assessment and Translation, National Toxicology Program, 2015.

NTP (National Toxicology Program), 2016. Monograph on immunotoxicity associated with exposure to perfluorooctanoic acid (PFOA) or perfluorooctane sulfonate (PFOS). National Toxicological Program, Research Triangle Park.
OECD, 2007. Test No. 426: Developmental Neurotoxicity Study, OECD Guidelines for the Testing of Chemicals, Section 4. OECD Publishing, Paris, https://doi.org/10.1787/9789264067394-en
OECD, 2016a. Series on Testing and Assessment No. 255: Guidance Document on the Reporting of Defined Approaches (DAs) to Be Used within IATA. ENV/JM/MONO(2016)28.
OECD, 2016b. Series on Testing and Assessment No. 256: Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to Be Used within IATA for Skin Sensitisation. ENV/JM/MONO (2016)29.
OECD, 2016c. Series on Testing and Assessment No. 260: Guidance Document for the Use of Adverse Outcome Pathways in Developing Integrated Approaches to Testing and Assessment (IATA). ENV/JM/MONO(2016)67.
OECD, 2017. Series on Testing & Assessment No. 275: Chemical Safety Assessment Workflow Based on Exposure Considerations and Non-animal Methods, OECD Publishing, Paris, ENV/JM/MONO(2017)27.
OECD, 2018. Test No. 443: Extended One-Generation Reproductive Toxicity Study, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris. https://doi.org/10.1787/9789264185371-en
OECD, 2020. Series on Testing and Assessment, No. 329: Overview of Concepts and Available Guidance related to Integrated Approaches to Testing and Assessment (IATA), OECD Publishing, Paris. ENV/JM/MONO(2020)25.
O'Hagan A, 2019. Expert knowledge elicitation: subjective but scientific. The American Statistician, 73(supp 1), 69–81. https://doi.org/10.1080/00031305.2018.1518265
Paparella M, Bennekou SH and Bal-Price A, 2020. An analysis of the limitations and uncertainties of *in vivo* developmental neurotoxicity testing and assessment to identify the potential for alternative approaches. Reproductive Toxicology, 96, 327–336. https://doi.org/10.1016/j.reprotox.2020.08.002
Paul-Friedman K, Gagne M, Loo LH, Karamertzanis P, Netzeva T, Sobanski T and Thomas RS, 2019. Examining the utility of in vitro bioactivity as a conservative point of departure: a case study. Toxicological Sciences.
Pitzer EM, Sugimoto C, Gudelsky GA, Huff Adams CL, Williams MT and Vorhees CV, 2019. Deltamethrin exposure daily from postnatal day 3–20 in Sprague-Dawley Rats causes long-term cognitive and behavioral deficits. Toxicoogical Sciences, 169, 511–523. https://doi.org/10.1093/toxsci/kfz067
Pletz J, Blakeman S, Paini A, Parissis N, Worth A, Andersson A-M, Frederiksen H, Sakhi AK, Thomsen C and Bopp SK, 2020. Physiologically based kinetic (PBK) modelling and human biomonitoring data for mixture risk assessment. Environment International, 143. https://doi.org/10.1016/j.envint.2020.105978
Radio NM and Mundy WR, 2008. Developmental neurotoxicity testing *in vitro*: models for assessing chemical effects on neurite outgrowth. Neurotoxicology, 29, 361–376. https://doi.org/10.1016/j.neuro.2008.02.011
Rezaie P and Dean A, 2002. Periventricular leukomalacia, inflammation and white matter lesions within the developing nervous system. Neuropathology, 22, 106–132. https://doi.org/10.1046/j.1440-1789.2002.00438.x
Richardson JR, Taylor MM, Shalat SL, Guillot TS 3rd, Caudle WM, Hossain MM, Mathews TA, Jones SR, Cory-Slechta DA and Miller GW, 2015. Developmental pesticide exposure reproduces features of attention deficit hyperactivity disorder. FASEB Journal, 29, 1960–1972. https://doi.org/10.1096/fj.14-260901
Rubinstein M, Patowary A, Stanaway IB, McCord E, Nesbitt RR, Archer M, Scheuer T, Nickerson D, Raskind WH, Wijsman EM, Bernier R, Catterall WA and Brkanac Z, 2018. Association of rare missense variants in the second intracellular loop of $Na_v1.7$ sodium channels with familial autism. Molecular Psychiatry, 23, 231–239. https://doi.org/10.1038/mp.2016.222
Sachana M, Bal-Price A, Crofton KM, Bennekou SH, Shafer TJ, Behl M and Terron A, 2019. International regulatory and scientific effort for improved developmental neurotoxicity testing. Toxicological Sciences, 167, 45–57. https://doi.org/10.1093/toxsci/kfy211
Shafer TJ, Meyer DA and Crofton KM, 2005. Developmental neurotoxicity of pyrethroid insecticides: critical review and future research needs. Environmental Health Perspectives, 113, 123–136.
Shafer TJ, Brown JP, Lynch B, Davila-Montero S, Wallace K and Friedman KP, 2019. Evaluation of chemical effects on network formation in cortical neurons grown on microelectrode arrays. Toxicological Sciences, 169, 436–455. https://doi.org/10.1093/toxsci/kfz052
Smith MT, Guyton KZ, Gibbons CF, Fritz JM, Portier CJ, Rusyn I, DeMarini DM, Caldwell JC, Kavlock RJ, Lambert PF, Hecht SS, Bucher JR, Stewart BR, Baan RA, Cogliano VJ and Straif K, 2016. Key Characteristics of carcinogens as a Basis for Organizing Data on Mechanisms of Carcinogenesis. Environmental Health Perspectives, 124. https://doi.org/10.1289/ehp.1509912
Smith RS, Kenny CJ, Ganesh V, Jang A, Borges-Monroy R, Partlow JN, Hill RS, Shin T, Chen AY, Doan RN, Anttonen AK, Ignatius J, Medne L, Bönnemann CG, Hecht JL, Salonen O, Barkovich AJ, Poduri A, Wilke M, de Wit MCY, Mancini GMS, Sztriha L, Im K, Amrom D, Andermann E, Paetau R, Lehesjoki AE, Walsh CA and Lehtinen MK, 2018. Sodium channel SCN3A ($Na_v1.3$) regulation of human cerebral cortical folding and oral motor development. Neuron, 99, 905–913.e7. https://doi.org/10.1016/j.neuron.2018.07.052
Tapia CM, Folorunso O, Singh AK, McDonough K and Laezza F, 2020. Effects of deltamethrin acute exposure on $Na_v1.6$ channels and medium spiny neurons of the nucleus accumbens. Toxicology, 440. https://doi.org/10.1016/j.tox.2020.152488

Terron A and Bennekou SH, 2018. Towards a regulatory use of alternative developmental neurotoxicity testing (DNT). Toxicology and Applied Pharmacology, 354, 19–23. https://doi.org/10.1016/j.taap.2018.02.002
Vester AI, Hermetz K, Burt A, Everson T, Marsit CJ and Caudle WM, 2020. Combined neurodevelopmental exposure to deltamethrin and corticosterone is associated with Nr3c1 hypermethylation in the midbrain of male mice. Neurotoxicology and Teratology, 80. https://doi.org/10.1016/j.ntt.2020.106887
Viel JF, Warembourg C, Le Maner-Idrissi G, Lacroix A, Limon G, Rouget F and Chevrier C, 2015. Pyrethroid insecticide exposure and cognitive developmental disabilities in children: the PELAGIE mother-child cohort. Environment International, 82, 69–75. https://doi.org/10.1016/j.envint.2015.05.009
Viel JF, Rouget F, Warembourg C, Monfort C, Limon G, Cordier S and Chevrier C, 2017. Behavioural disorders in 6-year-old children and pyrethroid insecticide exposure: the PELAGIE mother-child cohort. Occupational and Environmental Medicine, 74, 275–281. https://doi.org/10.1136/oemed-2016-104035
Watkins DJ, Fortenberry GZ, Sánchez BN, Barr DB, Panuwet P, Schnaas L and Meeker JD, 2016. Urinary 3-phenoxybenzoic acid (3-PBA) levels among pregnant women in Mexico City: distribution and relationships with child neurodevelopment. Environmental Research, 147, 307–313. https://doi.org/10.1016/j.envres.2016.02.025
Xue Z, Li X, Su Q, Xu L, Zhang P, Kong Z, Xu J and Teng J, 2013. Effect of synthetic pyrethroid pesticide exposure during pregnancy on the growth and development of infants. Asia Pacific Journal of Public Health, 25(4 Suppl), 72S–79S. https://doi.org/10.1177/1010539513496267
Zhang C, Xu Q, Xiao X, Li W, Kang Q, Zhang X, Wang T and Li Y, 2018. Prenatal deltamethrin exposure-induced cognitive impairment in offspring is ameliorated by memantine through NMDAR/BDNF signaling in hippocampus. Frontiers in Neuroscience, 12, 615. https://doi.org/10.3389/fnins.2018.00615

Abbreviations

3-PBA	3-phenoxybenzoic acid
ADHD	attention-deficit hyperactivity disorder
ADOS	Autism Diagnostic Observation Scale
AO	adverse outcome
AOP	adverse outcome pathway
ASD	autism spectrum disorder
BASC	Behavioural Assessment System for Children
BDNF	brain-derived neurotrophic factor
BMCs	benchmark concentrations
BMRs	benchmark responses
BPA	bisphenol A
BRIEF	Behaviour Rating Inventory of Executive Function
BSID-II	Bayley Scales of Infant Development-II
BSID-IIS	Bayley Scales for Infant Development—Spanish version
bw	body weight
CATs	critical appraisal tools
CBCL	1½-5Child Behaviour Check List for ages 1.5–5 years
cis-DBCA	*cis*-3-(2,2-dibromovinyl)-2,2-dimethylcyclopropane carboxylic acid
cMINC	Neural Crest Cell Migration Assay
CNS	central nervous system
CPDs	conditional probability distributions
DA	dopamine
DART	Digital Access to Research Theses
DAT	DA transporter
DK-EPA	Danish Environmental Protection Agency
DNT	developmental neurotoxicity
DNT-IVB	DNT *In vitro* Battery
DQ	developmental quotient
drd	DA receptor D1
DSM	Diagnostic and Statistical Manual of Mental Disorders
DST	Development Screen Test
EBSCO	Elton B. Stephens Company
EKE	expert knowledge elicitation
GD	guidance document
hiPSC	human induced pluripotent stem cells
hNPC	human neural progenitor cells
HOS	human observational studies

IC_{50}	half maximal inhibitory concentration
IUF	Leibniz Research Institute for Environmental Medicine
IVB	*In vitro* battery
IVIVE	*In vitro* to *in vivo* extrapolation'
KE	key events
KER	key event relationship
LOD	limit of detection
MDI	mental development index
MEA	microelectrode array
MIE	molecular initiating events
mRNA	messenger RNA
MSEL	Mullen Scales of Early Learning
MSN	medium spiny neuron
NA	nucleus accumbens
NAMs	new approach methodologies
NDDS	neurodevelopmental disorders
NE	norepinephrine
NeyriTox	Neurite Outgrowth of Central Nervous System Test
NMDA	*N*-Methyl-D-aspartate
NNF	neuronal network formation
NPC1 test	primary hNPC proliferation assay
NPC2 test	primary hNPC migration assay
NPC3 test	primary hNPC neuronal differentiation assay
NPC4 test	neuronal morphology (neurite length and area) of young neurons differentiated from hNPC
NPC5 test	oligodendrocyte differentiation
NR3C1	glucocorticoid receptor gene
OECD	Organisation for Economic Co-operation and Development
OHAT/NTP	The Office of Health Assessment and Translation/National Toxicology Programme
PAHs	polycyclic aromatic hydrocarbons
PBPK	physiologically based pharmacokinetic
PCBs	polybrominated diphenyl ether flame retardants
pCREB	phosphorylated camp response element binding protein
PDI	impaired psychomotor development
PeriTox	Neurite Outgrowth of Peripheral Nervous System Neurons Test
PFAS	perfluoroalkyl substances
PFOS	perfluorooctane sulfonate, perfluorooctane sulfonic acid
PPR	Panel on Plant Protection Products and their Residues
PQDT	Proquest Dissertations and Theses
PREV	EFSA Peer Review Unit
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
r/hNNF	Rat/Human neuronal network formation
RoB	risk of bias
RoB DH	risk of bias definitely high
RoB DL	risk of bias definitely low
RoB PH	risk of bias probably high
RoB PL	risk of bias probably low
SC	Scientific Committee
SDQ	Strengths and Difficulties Questionnaire
SPSF	Standard Project Submission Form
TG	Test Guideline
ToR	Terms of References
TrkB	tyrosine kinase b
TTC	threshold of toxicological concern
UA	uncertainty analysis
UKN	University of Konstanz
UKN2 test	The cMINC Neural Crest Cell Migration Assay
UKN4 test	The NeuriTox Neurite Outgrowth of CNS Neurons Test

UKN5 test	The PeriTox Neurite Outgrowth of PNS Neurons Test
US EPA	United States Environmental Protection Agency
VGSC	voltage gate sodium channel
WISC-IV	Wechsler Intelligence Scale for Children
WoE	weight of evidence

Annex A – Protocol of the Scientific Opinion

Annex A can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex B – Statistical analysis report

Annex B can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex C – List of included studies and excluded with reason deltamethrin

Annex C can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex D – List of included studies and excluded with reason flufenacet

Annex D can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex E – Outcome of the RoB deltamethrin

Annex E can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex F – Outcome of the RoB flufenacet

Annex F can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex G – Outcome of the RoB IVB (*In vitro* battery)

Annex G can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex H – Graph report *In vivo* and *in vitro* deltamethrin

Annex H can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex I – Human evidence table

Annex I can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex J – Graph report *In vivo* and *in vitro* flufenacet

Annex J can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex K – Uncertainty analysis tables for deltamethrin

Annex K can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex L – Uncertainty analysis tables for flufenacet

Annex L can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex M – AOP development and assessment (OECD template)

Annex M can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex N – IATA for deltamethrin (OECD template)

Annex N can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599

Annex O – IATA for flufenacet (OECD template)

Annex O can be found in the online version of this output (in the 'Supporting information' section): https://doi.org/10.2903/j.efsa.2021.6599