



Scientific perspectives on in vivo developmental neurotoxicity testing: study design, regulatory context, and future directions

Pragati S. Coder¹ · Richard Lewis² · Anika Christoffel³ · Phil Botham⁴ · Richard Bartlett⁴ · Megan D'souza⁵ et al. [full author details at the end of the article]

Received: 3 August 2026 / Accepted: 17 September 2026
© The Author(s) 2026

Abstract

Developmental neurotoxicity (DNT) potential of (agro)chemicals is assessed in rodents according to OECD Test Guidelines (TG) 426 (DNT) or 443 (Extended One-Generation Reproductive Toxicity; EOGRTS) incorporating a DNT cohort. While the EOGRTS evaluates reproductive and neurotoxicity endpoints across multiple life stages, key differences in study design, dose selection, and exposure duration between OECD 426 and OECD 443 can complicate data interpretation and impact hazard identification and risk assessment. This manuscript evaluates current DNT testing strategies and examines the feasibility of a combined DNT-EOGRTS approach aimed at generating data for regulatory decision making. This combined approach requires substantial modifications to the EOGRTS design, and its implementation should be approached with caution. Given the increased complexity and logistical challenges, additional cognitive assessments such as learning and memory (L&M) should only be included when scientifically justified on a substance specific basis. Retrospective analyses indicate that L&M is among the least sensitive DNT endpoints, and its inclusion in an EOGRTS does not yield a functional equivalent of a standalone DNT. The review also examines regulatory triggers and recent regulatory decisions relevant to inclusion of cognitive assessments. Finally, the status and limitations of alternative approaches, including the in vitro DNT battery and zebrafish embryo models, are discussed. Overall, OECD 426 remains the scientifically preferred approach for DNT assessment. While a combined design may be feasible, it is inherently suboptimal for DNT evaluation. Collectively, this comprehensive review provides recommendations for generating robust DNT data while supporting animal welfare-conscious, future-oriented regulatory testing strategies.

Key points

The OECD 426 is the globally accepted standard for DNT assessment. Despite interest in integrated approaches, a standalone DNT provides the most scientifically robust evaluation of developmental neurotoxicity.

Although combining a DNT assessment within an EOGRTS (OECD 443) may be feasible it is scientifically and logistically challenging. Significant modifications to study design, including dose selection, exposure duration, pup allocation, and endpoint assessment, are required to generate interpretable data.

Learning and memory (L&M) assessments should only be included when scientifically justified on a substance specific basis because L&M is among the least sensitive DNT endpoints.

An EOGRTS with cognitive endpoints is not functionally equivalent to a standalone DNT. Inclusion of L&M assessments in OECD 443 may generate data of limited interpretability for hazard identification and risk assessment.

Future DNT testing should balance scientific rigour, regulatory needs, and animal welfare. Alternative methods such as in vitro DNT batteries and zebrafish embryo models are advancing and show promise but are not yet sufficiently validated to replace in vivo developmental neurotoxicity testing.

Keywords Developmental neurotoxicity · OECD 426 · OECD 443 · EOGRTS · Learning and memory · New approach methods · DNT in vitro battery

Introduction

Developmental neurotoxicity (DNT) assessments are used to evaluate the potential adverse effects of chemicals on the developing nervous system resulting from exposures during development of the offspring in utero through to early adolescence. The developing brain can be particularly vulnerable to environmental insults, and exposure to certain substances during critical periods of brain development can have long-lasting effects on cognitive, behavioural, and neuronal function. (Henck and Morford (2010)).

The information requirements for DNT vary according to the relevant EU legislation regulating plant protection products, biocidal products or industrial chemicals, however, there are three guidelines available for conducting in vivo testing for DNT: the US EPA OPPTS 870.6300 Guideline (US EPA 1998), the OECD Test Guideline (TG) 426 for Developmental Neurotoxicity (OECD 2007), and the OECD 443 for the Extended One-Generation Reproductive Toxicity Study with the inclusion of a DNT cohort (2A/2B) (OECD 2018). Despite these standard test guidelines, design differences exist across the guidelines and the triggers for conducting these studies differ across regulatory approaches. This section provides further details on the present regulatory background and triggers for DNT studies, based on relevant legislative information and data requirements.

Scope

This publication will evaluate the scientific and regulatory frameworks for developmental neurotoxicity (DNT) testing, focusing on the strengths, limitations, and comparative utility of OECD TG 426 and OECD TG 443 with DNT cohorts, and combinations thereof. Particular emphasis is placed on study design considerations, learning and memory assessments, dose selection, exposure duration, animal allocation and data interpretation for hazard identification, risk assessment and classification purposes. This publication will also examine recent regulatory requests influencing testing requirements and the current status and future potential of alternative non-animal and new approach methodologies (e.g., in silico, in vitro etc.) for assessment of DNT. The ultimate objective of this manuscript is to provide evidence-based scientific recommendations that support robust, interpretable and animal-welfare driven DNT assessment strategies.

Regulatory background and triggers

EU biocidal products regulation ((EC) No. 528/2012)

The Biocidal Products Regulation (BPR; (EC) No. 528/2012) was updated in March 2022, such that DNT is now a standard information requirement within the core data set (CDS). The specific rules for adaptation from the standard information requirements state that a DNT study shall not be conducted if the available data:

1. Indicate that the substance causes developmental toxicity and meets the criteria for classification as toxic for reproduction category 1A or 1B: May damage the unborn child (H360D); and
2. Are adequate to support a robust risk assessment.

The BPR data requirement describes three study options that can fulfil the DNT information requirement:

1. OECD TG 426: Developmental neurotoxicity study;
2. Any relevant study (set) providing information equivalent to OECD TG 426; or
3. OECD TG 443 (Extended One-Generation Reproductive Toxicity Study, EOGRTS) with inclusion of DNT Cohort 2 and additional investigations of cognitive function.

Investigations for DNT under these study options include tests for clinical observations, motor activity, motor and sensory function, and cognitive function (including associative learning and memory, L&M), as well as neuropathological examination and brain weight measurements. Guidance from the European Chemicals Agency (ECHA) for the BPR (Part A, Volume III: Human Health; (ECHA (2022b)) provides testing details to fulfil these minimum requirements, but compared with other legislative information requirements, this guidance is particularly extensive, reflecting the absence of other ECHA guidance documents that would otherwise support applicants and the evaluating Competent Authorities (CAs) in selecting appropriate tests for cognitive function.

Alternative test methods (such as a battery of in vitro DNT assays) are not yet fully validated and are not described in the BPR requirements because an OECD guidance document on an integrated approach to testing and assessment (IATA) for DNT is still under development and was not available at the time of writing. However, the BPR guidance further states that the in vitro DNT testing battery is not considered sufficient to fulfil the minimum data requirements because it currently does not provide information equivalent to that generated by in vivo tests. While positive

in vitro findings indicating DNT properties may strengthen the overall weight of evidence (WoE), negative in vitro *results* alone do not allow conclusions on the absence of DNT properties due to inherent limitations of in vitro systems compared with in vivo studies. In addition, in vitro data alone are currently not sufficient for classification and labelling under the CLP Regulation (Regulation (EC) No. 1272/2008).

EU plant protection products regulation ((EC) No. 1107/2009)

Within the European regulatory frameworks for Plant Protection Products (PPPs), DNT is considered an effect of particular significance. More recently, draft proposals have been made to amend the standard data requirement for active substances under the PPP framework (Regulation (EC) No. 283/2013) to include DNT testing as a core testing requirement (including cognitive function). However, experimental verification of DNT endpoints is not a standard data requirement at the time of this publication. Additional information or specific studies addressing the postnatal manifestation of DNT may be required where indicated by observations in other studies or by the mode of action of the test substance. Such additional data requirements are not further specified. Nevertheless, for a considerable number of PPP active substances, DNT studies performed in accordance with OECD 426 (2007) and/or the US EPA Guideline OPPTS 870.6300 (1998) are available. The majority of these studies were triggered by US EPA data requirements. In several recent active substance reviews, the EOGRTS with inclusion of DNT Cohorts 2A/2B was accepted to address developmental neurotoxicity concerns in combination with data gaps identified in datasets containing two-generation studies conducted under the outdated (pre-2001) OECD 416 test protocol, particularly where conclusions on thyroid-mediated endocrine disruption could not be drawn.

EU REACH regulation ((EC) No. 1907/2006)

Testing for DNT became an optional information requirement under the European chemicals' legislation REACH (Regulation (EC) No. 1907/2006) a few years after adoption of the EOGRTS as OECD 443 on 28 July 2011 and following subsequent updates in 2018 and 2025.

Compared with the Two-Generation Reproduction Toxicity Study (OECD 416 (2001)), the EOGRTS offers important improvements in animal welfare through reduction and refinement within a modular study design. The adverse effects of substances can be evaluated in separate cohorts for fertility and development (Cohorts 1A/1B, where Cohort 1B optionally mated to produce a second (F2) generation),

effects on the developing nervous system (Cohorts 2A/2B, DNT), and effects on the developing immune system (Cohort 3, developmental immunotoxicity (DIT)).

No specific DNT or DIT endpoints are part of the Two-Generation Reproduction Toxicity Study (OECD 416), which was initially the standard information requirement for reproductive toxicity under REACH Annexes IX (if triggered) and X. Since 13 March 2015, the EOGRTS in its basic study design (Cohorts 1A/1B without extension to an F2 generation) has become the standard testing requirement in the EU for evaluating reproductive toxicity (in lieu of the Two-Generation Reproduction Toxicity Study) for new substances in accordance with Regulation (EU) 2015/282. OECD 416 studies initiated before that date remain scientifically relevant for regulatory assessment and decision-making. Extension of EOGRTS Cohort 1B to include an F2 generation and inclusion of Cohorts 2A/2B (DNT) and/or Cohort 3 (DIT) are required under REACH only where the specific provisions or triggers set out in the legislation are met, or where a particular concern has been identified.

A particular concern for triggering DNT Cohorts 2A/2B under REACH Annexes IX/X (8.7.3., Column 2) may include:

1. Existing information on the substance derived from relevant available in vivo or non-animal approaches (e.g., abnormalities of the central nervous system (CNS), or evidence of adverse effects on the nervous system in adult animals or following prenatal exposure);
2. Specific mechanisms or modes of action associated with (developmental) neurotoxicity (e.g., cholinesterase inhibition or relevant changes in thyroid hormone levels linked to adverse effects); or
3. Existing information on structurally analogous substances suggesting similar effects or mechanisms.

There have also been instances where additional cognitive function tests (learning and memory) were requested by ECHA as part of the testing decision for EOGRTs with the DNT cohort (Cohort 2). In several cases, such requests were subsequently annulled by the ECHA Board of Appeal (BoA), as they extended DNT testing requirements beyond those specified in OECD 443 (see Key Endpoints) (Beekhuizen et al. (2024), Salverda (2024)).

Study designs, key endpoints and limitations

Guideline study designs

The aim of a developmental neurotoxicity assessment is to investigate effects on the development and function of the nervous system resulting from repeated exposure to a substance in utero and during early postnatal development. This section focuses on study designs (per the aforementioned test guidelines), key endpoints, assessment parameters, and associated limitations. The preferred test species for DNT assessment is the rat.

Standalone DNT (OECD 426 & OPPTS 870.6300)

In a standalone DNT study conducted according to OPPTS 870.6300 or OECD 426, the test substance is administered daily to time-mated females from implantation (e.g., Gestation Day 6 when sperm positive day is Gestation Day 0) through gestation and parturition and continuing until weaning of the offspring (generally Lactation Day (LD)/Postnatal Day (PND) 21), resulting in approximately 36–37 total days of maternal (perinatal) exposure. The test substance is expected to be administered by the route most relevant to potential human exposure, generally oral (e.g., gavage, dietary, or drinking water), although other routes of exposure (e.g., dermal or inhalation) may be used depending on substance properties and/or known human exposure routes. Pregnant dams are evaluated during gestation and lactation to provide a basis for comparison with their offspring. At or just prior to weaning, offspring are randomly selected for DNT evaluation beginning at weaning and again as young adults. Evaluation of at least 20 litters per dose group at each age is recommended. F1 offspring are indirectly exposed in utero and potentially via maternal milk during the preweaning period. Direct dosing of F1 offspring is not typical but should be considered where transfer through maternal milk cannot be demonstrated (see *Direct Dosing of F1 Rat Pups*). Recommended assessments include physical and developmental landmarks, and functional/behavioural endpoints including at least two measures of behavioural ontogeny, and assessments of motor activity, sensory function, and L&M assessments at a minimum of two ages.

Table 1 Selection of F1 Rat Pups, by Cohort, per OECD 443 (EOGRTS)

F1 cohorts	Males/group	Females/group
F1 Cohort 1A (Reproductive)	20	20
F1 Cohort 1B (Reproductive)	20	20
F1 Cohort 2A (DNT)	10	10
F1 Cohort 2B (DNT)	10	10
F1 Cohort 3 (DIT)	10	10

EOGRTS with a DNT cohort (OECD 443) design

When DNT is evaluated as part of an EOGRTS, the test substance is administered to adult F0 males and young adult nonpregnant F0 females for 2 to 10 weeks prior to mating. Dosing continues through mating, gestation, lactation, and until weaning of the offspring (generally LD/PND 21), at which time the parental generation is euthanized. At or just prior to weaning, F1 offspring are selected for direct dose administration post-weaning and randomly assigned to cohorts for evaluation of reproductive endpoints (Cohorts 1A/1B), developmental neurotoxicity (2A/2B), and developmental immunotoxicity (Cohort 3). Assignment of at least 20 rat pups per group (10/sex/group, representing as many litters as possible) is recommended for DNT assessments in Cohorts 2A/2B. Table 1 presents the allocation of F1 rat pups to these cohorts on an EOGRTS study.

The F1 offspring in an EOGRTS are indirectly exposed to the test substance in utero during gestation and potentially via maternal milk during the preweaning period. Following weaning, F1 offspring are directly exposed to the test substance until euthanasia via the same route of administration as the parental generation. Recommended assessments include physical and developmental landmarks and functional/behavioural endpoints, including a functional observational battery, motor activity, and sensory function assessed at a single interval. Unlike OECD 426, L&M investigations are not required for Cohort 2A/2B under the published test guideline (OECD 2025b).

In recent years, ECHA has requested inclusion of L&M testing for F1 animals (e.g., Morris water maze, radial arm maze for allocentric tasks; Cincinnati water maze for egocentric tasks) in both EOGRTS compliance check (CCH) decisions and substance evaluations under REACH, typically at adolescence (PND 25 ± 2 days) and young adulthood (PND 60 and older). These requests have been previously justified based on assumptions that perturbations in thyroid hormone signalling impact cognitive abilities or based on concerns arising from structurally analogous substances.

However, a critique by Arts et al. (2023) concluded that the publications relied upon by ECHA do not substantiate that decreases in circulating thyroxine (T4) levels result in impaired L&M in rats, and that alterations in thyroid hormones alone are not sufficient to justify inclusion of L&M testing in an EOGRTS. This conclusion aligns with the US EPA's retrospective review of 69 DNT studies (Raffaie et al. (2010)), which demonstrated that L&M endpoints rarely determined the point of departure, were less sensitive than locomotor activity or auditory startle, and never served as the sole basis for lowest observed adverse effect level (LOAEL) determination. This analysis informed the OECD DNT expert group decision during adoption of the

OECD 443 to exclude cognitive tests for L&M (Piersma et al. (2012)). Similarly, a more recent curation of 104 DNT studies by the European Food Safety Authority (Mangas et al. (2025)) ranked locomotor activity as the most common endpoint used to derive the no observed adverse effect level (NOAEL), followed by: changes in brain morphometry, effects in acoustic startle response, decreased brain weight, brain cholinesterase (ChE) inhibition, and lastly effects on L&M. It is important to note that this analysis by Raffaele et al. (2010) was based solely on guideline DNT studies where L&M assessments were performed during the F1 non-dosing period, following cessation of maternal exposure. This dosing paradigm allows exposure of rat pups prior to weaning, during a period of active neurodevelopment (Semple et al. (2013)) yet avoids any confounding impact due to acute effects of test substance exposure on motor function, which is often necessary for an animal to perform in the majority of L&M tasks (see [Route of administration and exposure duration](#)). Taken together, these retrospective data analyses demonstrate that inclusion of L&M assessments, the least sensitive DNT endpoint, is unlikely to improve the sensitivity of DNT assessments beyond endpoints already included in the EOGRTS.

Combination EOGRT-DNT study design

Requests for inclusion of L&M assessments within an EOGRTS overlook a fundamental design principle: the DNT cohort in 443 was never intended to function as a full standalone DNT study. In the original proposed design by Cooper et al. (2006), the DNT module was deliberately limited in scope, using half the number of rat pups typically assigned to functional/cognitive testing in a standalone DNT and requiring functional assessments only in adult F1 animals, with the exception of auditory startle. L&M testing was explicitly not included. The intent was that these limited functional and cognitive assessments in the EOGRTS would serve as a reliable screen to identify potential developmental neurotoxicity and inform the need for follow-up (Tier 2) testing. With the adoption of OECD 443 in 2011, however, the EOGRTS became the highest-tier reproductive and developmental toxicity study, integrating reproductive, developmental, neurotoxicity, and immunotoxicity endpoints into a single complex study design, none of which underwent separate OECD validation because all were established regulatory methods prior to adoption of OECD 443 (Fegert et al. (2012)). The technical and logistical challenges associated with this integrated testing approach have been extensively discussed (Fegert et al. (2012); Beekhuijzen et al. (2016a), Beekhuijzen et al. (2016b); Moore et al. (2016); Saghir and Dorato (2016); Schulz et al. (2014); Oldenburger et al. (2022); Schneider et al. (2011)),

and inclusion of additional functional assessments will only further increase study complexity.

Paragraph 51 of OECD 443 states “*if existing information indicates the need for other functional testing (e.g., sensory, social, cognitive), these should be integrated without compromising the integrity of the other evaluations conducted in the study.*” It is important to recognise that conducting L&M testing at a single age using a limited cohort size as included in an EOGRTS (Cohort 2A, 10/sex/group) is scientifically less robust than a standalone DNT (20/sex/group). Even augmenting group sizes cannot overcome the limitations posed by a combination study (e.g., dose selection, exposure duration, pup allocation, and endpoint assessment) despite the ethical advantages from a 3Rs perspective and the likelihood of compromising the quality and integrity of the core evaluations integral to both study designs.

Clear differences in the basic study designs and regulatory guidance documents for an EOGRTS and a standalone DNT underscore why a combination EOGRT-DNT study is likely to pose logistical and interpretive challenges. A side-by-side comparison of an EOGRTS, a standalone DNT and a Combination EOGRT-DNT study are presented in Table 2.

Dose selection criteria differ fundamentally

Dose levels for a standalone DNT are selected so as to induce some maternal toxicity, but the high dose is chosen in a manner that does not induce excessive offspring toxicity. In contrast, EOGRTS dose levels are chosen with the aim to induce some parental systemic toxicity or be high enough to elicit deficits in sexual function and fertility. Furthermore, the guidance is clear that EOGRTS dose levels should not be lowered to ensure offspring survival (see [Dose Level Selection](#)). It is well established that maternal toxicity, particularly at the level recommended for the EOGRTS (at or just below maximal tolerated dose), has an impact on developmental, reproductive, as well as DNT and developmental immunotoxicity endpoints (Ladics et al. (2005)).

F1 pup group sizes and statistical power differ

Only 20 rat pups per group (10/sex/group) are allocated to DNT assessments in an EOGRTS. In contrast, at least 40 rat pups per group (20/sex/group) are allocated to each subset on a standalone DNT (see Table 2; and *Selection and Allocation of F1 Rat Pups*). This difference in group size has implications for statistical power and the ability to detect treatment-related effects, particularly for variable functional and behavioural endpoints.

Table 2 Study Designs for EOGRT (OECD 443) and Standalone DNT (OECD 426) Studies compared with options for conducting a combined EOGRT and DNT (OECD 443 + 426). Examples of common animal assignments are shown (see text for descriptions of study complexity and limitations)

Investigations	OECD 443			OECD 426			OECD 443 + 426		
	Age	Rat pups /Sex /Group	Age	Rat pups /Sex /Group	Rat pups /Sex /Group	Age	Rat pups /Sex /Group	Rat pups /Sex /Group	
		(Cohort)		Example 1	Example 2		(Cohort)	Example 1	
Dosing Period	Premating – Gestation Day 0 – Lactation Day 21 (Dams Only)			Gestation Day 6 – Lactation Day 21 (Dams Only)			Premating – Gestation Day 0 – Lactation Day 21 (Dams);		
	Day 0 – Lactation Day 21 (Dams);			No Direct Exposure for F1 Rat pups			PND 21-Euthanasia (All F1 cohorts)		
	PND 21-Euthanasia (All F1 cohorts)								
F1 generation developmental assessments									
<i>Reproductive assessments</i>									
Sexual Developmental Landmarks									
Onset of Estrus	>PND 24	60 m/60f	>PND 24	20 m/20f	20 m/20f	>PND 24	60 m/60f	60 m/60f	
	>PND 24	20 m/20f (1A)	–	–	–	>PND 24	20 m/20f (1A)	20 m/20f (1A)	
Estrous Cyclicity	>PND 75	20 m/20f (1A)	–	–	–	>PND 75	20 m/20f (1A)	20 m/20f (1A)	
Reproductive Assessments	>PND 91	20 m/20f (1A)	–	–	–	>PND 91	20 m/20f (1A)	20 m/20f (1A)	
		20 m/20f (1B)					20 m/20f (1B)	20 m/20f (1B)	
<i>Developmental immunotoxicity assessments</i>									
T-Cell Antibody Response Assay (TDAR)	>PND 56	10 m/10f (3)	–	–	–	>PND 56	10 m/10f (3)	10 m/10f (3)	
Immunophenotyping	PND 91	10 m/10f (1A)	–	–	–	PND 91	10 m/10f (1A)	10 m/10f (1A)	
<i>Developmental neurotoxicity assessments</i>									
Weanling Brain Weight (post fixation)	PND 21/22	10 m/10f (2B)	PND 21/22	10 m/10f	10 m/10f	PND 21/22	10 m/10f (2B)	10 m/10f (2B)	
Weanling Gross Brain Measurements (post fixation)									
Weanling Neuropathology (immersion or perfusion fixation)									

Table 2 (continued)

Investigations	OECD 443			OECD 426			OECD 443 + 426			
	Age	Rat pups / Sex / Group (Cohort)	Age	Rat pups / Sex /Group Example 1	Rat pups / Sex /Group Example 2	Rat pups / Sex /Group Example 3	Age	Rat pups / Sex /Group (Cohort) Example 1	Rat pups / Sex /Group (Cohort) Example 2	Rat pups / Sex /Group (Cohort) Example 3
Behavioural Ontogeny	NA	NA	Pre-weaning	20 m/20f	20 m/20f	20 m/20f	PND 11, 21, 35, 49 and 65 (± 5 days)	20/sex (1A, 2A)	20 m/20f (1A)	20 m/20f (1A)
Motor Activity & Functional Observational Battery	PND 69 (± 6 days)	10 m/10f (2A)	PND 25 (± 2 days) PND 65 (± 5 days)				PND 24 (± 1 day) PND 69 (± 5 days)			
	PND 24 (± 1 day)		PND 25 (± 2 days) PND 65 (± 5 days)		20 m/20f		PND 24 (± 1 day) PND 65 (± 5 days)			20 m/20f (2A)
	NA		PND 25 (± 2 days) PND 65 (± 5 days)		20 m/20f**	20 m/20f	PND 25 (± 2 days) ≥PND 60		20 m/20f (2A)**	20 m/20f (1C)**
			~PND 70 (± 5 days)	10 m/10f*	10 m/10f*	10 m/10f*	PND 82 (± 7 days)	10/sex (2A)*	10/sex (2A)*	10/sex (2A)*
Adult Brain Weight (post fixation)	PND 82 (± 7 days)									
Adult Gross Brain Measurements (post fixation)										
Adult Neuropathology (Perfusion fixation)										
Brain Morphometry (Quantitative)										
Animals Required for DNT Assessments	–	20 m/20f		30 m/30f	50 m/50f	70 m/70f		30 m/30f	50 m/50f	70 m/70f
Total Animals Selected	–	70 m/70f	–	30 m/30f	50 m/50f	70 m/70f	–	70 m/70f	80 m/80f	100 m/100f

* 10 m/10f adult animals selected for neuropathology; animals assigned to motor activity assessments

** If the same test of L&M is included at 2 ages, animals assigned to behavioural ontogeny, motor activity and auditory startle may also be evaluated for L&M at one age

Dosing paradigms are not comparable

Different dosing paradigms are used for a standalone DNT study compared to an EOGRTS where direct exposure of F1 rat pups to the test substance continues beyond weaning until scheduled euthanasia of adult F1 animals. In contrast, F1 rat pups on a standalone DNT are not directly exposed to the test substance (see [Route of administration and exposure duration](#)), which can result in stark differences in pup growth and development. It may be prudent to consider a dosing paradigm similar to a standalone DNT for the EOGRT-DNT cohort i.e., ending exposure at weaning, consistent with standalone DNT designs, may better distinguish true DNT originating during development from effects of continued post-weaning exposure of the EOGRTS-DNT cohort into adulthood (contrary to the RAC position that effects in adults are considered DNT if there was any evidence of developmental exposures).

Collectively, these design differences mean that adding L&M testing to an EOGRTS does not create the functional equivalent of a standalone DNT and may yield equivocal, low power data unsuitable for hazard identification and risk assessment. A true combined EOGRT-DNT study would require substantial redesign—larger group sizes, multiple ages of assessment, and modified dosing schedules—to achieve comparable scientific robustness.

Several recipients of final decision letters under the REACH compliance check process requesting investigations of L&M function in the EOGRTS filed appeals (Cases A-002-2022 and A-003-2022, A-009-2022, A-006-2023, A-007-2023 and A-008-2023) with the ECHA BoA based on the premise that such investigations are not an information requirement under REACH Annexes IX/X. This position was upheld by the BoA in Cases A-002-2022, A-003-2022 and A-009-2022. The BoA further confirmed that paragraph 51 of OECD 443 serves to ensure that the integrity of evaluations conducted in Cohorts 2A/2B of an EOGRTS is not compromised by any additional functional testing but does not establish an obligation to carry out such testing. As a consequence of these rulings, ECHA has revised its approach and has confirmed that L&M function testing will not be requested in OECD 443 studies in future CCH or TP evaluations under REACH. However, despite the shift in ECHA's stance on L&M assessments on EOGRTS conducted under REACH, L&M assessments remain standard endpoints in OECD 426 and EOGRTS studies conducted under ECHA guidance for biocidal products (ECHA 2022a, b) in stark contrast to both EU REACH and PPP information requirements.

Study design elements, key endpoints and limitations

The aim of DNT testing remains the investigation of effects on the developing nervous system. These observed effects can be functional (eliciting neurochemical, neurophysiological, or neurobehavioural deficits in the central or peripheral nervous systems) or structural (impacting neuroanatomy). Here, we compare and contrast key elements of study design, study endpoints, and associated limitations and challenges across DNT study designs.

Dose level selection

OECD 426 states that “*the highest dose level should be chosen with the aim to induce some maternal toxicity (e.g., clinical signs, decreased body weight gain (not more than 10%) and/or evidence of dose-limiting toxicity in a target organ). The highest dose level should be the maximum dose which will not induce excessive offspring toxicity, or in utero or neonatal death or malformations, sufficient to preclude a meaningful evaluation of neurotoxicity.*”

This approach is broadly consistent with and supported by well-established conventions, such as those proposed by HESI (Brannen et al. (2011), Beyer et al. (2011)), which concluded that ‘*A decrease in body weight gain during the treatment period in the dose-range finding study of 10 to 15% should be justifiable as a suitable high dose in reproductive toxicity studies.*’

By contrast, for an EOGRTS, ECHA advice on dose level selection ECHA (2022a) states that “*the highest dose level should be chosen with the aim to induce some systemic toxicity but not death or severe suffering.*” Furthermore, “*the study should be designed to ensure adequate assessment of the effects on sexual function and fertility, the dose levels should not be reduced to get enough offspring for the assessment of developmental toxicity.*”

Taken together, the OECD guidance and ECHA advice (both intended to be used for DNT assessments) are incongruent in dose level selection philosophy. The standalone DNT is designed to assess adverse effects on the developing nervous system and as such specifies that the highest dose should not induce excessive offspring toxicity. In contrast, the intent of the EOGRTS is to provide ‘*an evaluation of reproductive and developmental effects that may occur as a result of pre- and postnatal chemical exposure as well as an evaluation of systemic toxicity in pregnant and lactating dams and young and adult offspring*’ (OECD (2025b)). In recent years, however, the dose selection for EOGRTS conducted under REACH has increasingly focused on CLP considerations, prioritizing assessment of parental sexual function and fertility over offspring survival. This shift

may compromise postnatal developmental evaluations in the offspring, including DNT assessments and assessments of sexual function and fertility in the second generation. These dose selection principles are intended to be applied even when DNT (Cohort 2A/2B) and developmental immunotoxicity (Cohort 3) assessments are included in the study design. Because neurobehaviour can be impacted by dysfunction in other organ systems (e.g., liver, kidneys, endocrine), excessive systemic toxicity due to high-dose exposure in offspring can manifest as secondary non-specific changes in behaviour and/or cognitive abilities (Lewis et al. (2024)) that confound interpretation of DNT assessments. Consequently, instead of providing information that may support a WoE approach for interpretation of observed DNT findings, prioritisation of sexual function and fertility detracts from the original goal and design of the EOGRTS as a comprehensive, integrated evaluation of fertility, reproduction, and pre- and postnatal development (Cooper et al. (2006)).

This concern is not merely theoretical. Several reviews and case-based critiques have documented EOGRTS where high-dose selection elicited marked developmental toxicity (Beekhuijzen et al. (2024)) (Lewis et al. (2024)). Because neurobehaviour can be secondarily altered by systemic toxicity or non-nervous system toxicity in a non-specific manner, high dose systemic stressors in dams and/or offspring can produce changes in activity which may be erroneously attributed to DNT. For example, changes in the musculoskeletal system may be misinterpreted as neurotoxicity. Such potential confounding of DNT testing may be difficult to detect in the absence of overt indicators such as in utero and neonatal death or macroscopically apparent malformations. In addition, rat pups are dependent on dams for maternal care, milk quality and quantity, and thermal homeostasis (Tyl et al. (2008), Desesso (1987), Desesso et al. (2008)). Prenatal stress and food deprivation have been associated with long-lasting effects on neurotoxicity endpoints, including cognitive, behavioural, and brain measures (Albee et al. (1987), Garman et al. (2016), Charil et al. (2010), Weinstock (2008), Markham et al. (2010)). Thus, dose levels that are optimal for identifying primary functional and behavioural effects on the developing nervous system may differ from those selected for evaluation of fertility, reproduction, or general systemic toxicity. Conversely, dose selection optimised for general reproductive or pre- and postnatal developmental toxicity may introduce secondary non-specific effects that are not truly indicative of DNT.

Route of administration and exposure duration

Per OECD 426, the route of administration for the DNT will generally be oral (e.g., gavage, dietary, via drinking water),

but other routes (e.g., dermal, inhalation) may also be used. Similarly, the EOGRTS is designed for the administration of the test chemical orally, although the design can be modified for other routes of administration. In both cases, the selected dose route should be the one that is most relevant for human exposure.

There are, however, distinct differences in the overall exposure duration between a standalone DNT and an EOGRTS where the extended duration of exposure in the EOGRTS beyond weaning adds uncertainty to the interpretation of neurotoxicity detected in adulthood. This dosing paradigm makes it difficult to differentiate between DNT, acute adult neurotoxic effects or effects due to repeated exposures during adulthood (without developmental exposures). In both designs, F1 rat pups can be directly exposed to the test chemical during the preweaning period via dermal contact from the dam (for dermal or inhalation routes of exposure). For studies where the test chemical is administered in the diet, offspring can also be directly exposed via self-weaning and direct intake of feed beginning in about the third week of the preweaning period once the eyes are open and the rat pups are active and strong enough to access the maternal food jars. It needs to be noted that in instances where there are secondary sources of exposure, the actual dosage received by the F1 offspring can usually not be quantified. Thus, the evidence of developmentally neurotoxic effects in the offspring may depend on the kinetic profile of each individual compound in the maternal, as well as juvenile, animals. For example, for compounds where secretion into milk is high, rat pups might receive higher dosages (per kg body weight) during the pre-weaning period compared to directly dosed adult animals. This variable is difficult to control on a study-by-study basis and adds complexity when comparing sensitivity for neurotoxic effects between adults and rat pups.

Recently, the ECHA Biocides Working Group requested that OECD 426 (DNT) studies include extended dosing beyond weaning through to approximately PND 60. This emerging expectation reflects a shift in regulatory interpretation of the guideline; however, it also raises important questions regarding the scientific justification and added value of such an extension. The original design of the guideline was intended to capture the most sensitive windows of neurodevelopment, namely gestation and early postnatal life, when the brain is undergoing rapid growth, differentiation, and organisation. These stages are widely regarded as the periods of greatest vulnerability to chemical insult, and it is therefore not self-evident that extending exposure into later developmental phases provides additional sensitivity for hazard identification.

A key issue is the distinction between developmental neurotoxicity and general or frank neurotoxicity in adults.

Continued dosing through PND 60 increases the likelihood that any observed effects may reflect direct pharmacological or toxicological action in a more mature organism, rather than the effects observed due to disruption of developmental processes. This has the potential to blur the conceptual boundary that the OECD test guideline seeks to maintain between developmental and adult toxicity endpoints. As a result, interpretation of findings becomes increasingly more complex, particularly in cases where effects are only observed after prolonged exposure, raising questions as to whether such outcomes fall within the intended scope of a DNT study or would be more appropriately addressed in repeat-dose or dedicated neurotoxicity studies (as performed under OECD 408 or 424).

While it is acknowledged that aspects of brain maturation, such as synaptic refinement and functional specialisation, continue beyond weaning, there remains limited empirical evidence demonstrating that extending dosing to PND 60 consistently identifies adverse effects that would not have been detected under standard pre-weaning exposure paradigms or in standalone repeat-dose toxicity studies. The extent to which this later developmental window represents a uniquely sensitive period for chemical perturbation remains a matter of scientific debate. In this context, a more evidence-based and case-specific justification for extended dosing may be warranted, rather than its routine inclusion, to ensure that DNT study designs remain both scientifically grounded and proportionate to their regulatory purpose.

The request for extended dosing also introduces practical and ethical considerations. Prolonging the exposure period increases study duration (and numbers of procedures per animal) and resource requirements, without clear and consistent evidence that the additional data materially influences regulatory decision-making. At the same time, evolving expectations from regulatory bodies (such as ECHA and EFSA) in the absence of formal revisions to the core OECD test guidelines, introduce regulatory uncertainty and potential inconsistency in study design requirements across jurisdictions.

Direct dosing of F1 rat pups

For substances that are poorly secreted or not secreted into maternal milk, both OECD 426 and 443 recommend direct dosing of preweaning rat pups. Such evidence can be retrieved from pharmacokinetic data, offspring toxicity or changes in specific biomarkers (e.g., maternal milk concentrations). Direct dosing of rat pups, typically conducted via oral (gavage), can be included in either study type beginning as early as PND 4; however, the scientifically appropriate starting age depends heavily on CNS developmental staging. Current literature indicates that exposure beginning

around PND 10 captures the major CNS developmental window most comparable to human neonates. Specifically, Schmitt and Barrow (2022) show that CNS ontogeny in rodents is sufficiently aligned with human neonatal development by PND 10, and that dosing before PND 7 is rarely necessary except in exceptional mechanistic cases (Schmitt and Barrow 2022). Similarly, Semple et al. (2013) emphasized that rodent PND 7–10 corresponds to the term equivalent human infant for many neurodevelopmental processes. Taken together, these manuscripts support initiation of direct dosing of rat pups around PND 10, or earlier in rare occasions, when warranted.

When preweaning rat pups are directly dosed, exposure of lactating dams can be stopped as the offspring are directly exposed. However, this change is not compatible with the EOGRT study design where dams continue to nurse all F1 rat pups including those assigned to other cohorts. Nursing and nesting behaviour are integral components of postnatal developmental assessments on a reproductive toxicity study and modifying maternal exposure would disrupt the validity and functional endpoints required by the guideline.

On studies conducted via diet or drinking water routes of administration, there is also the risk of overdosing both the offspring and the dams during the second half of lactation due to the increasing caloric demands of lactation (on dams) and the high food/water efficiency (of rat pups) during these periods. Most laboratories choose to adjust dietary or water concentrations during these periods to ensure that target doses are maintained. Lastly, it is important to note that direct dosing of young rat pups via oral gavage is a stressful procedure that poses the risk of intubation trauma and overdosing due to direct exposure and exposure via the mother's milk. As such, direct dosing could lead to a different hazard assessment of chemicals depending on the choice of the route of administration, as discussed by Arts and Beekhuizen (2020). Often, (agro)chemicals are used in industrial or professional settings where direct exposure to neonates is unlikely. Thus, although direct dosing of rat pups is feasible from a technical standpoint and several industry laboratories have experience with the technique, this does not automatically imply that it is the most relevant approach for this study type or population and should therefore only be considered in exceptional cases and, when utilised, should be justified on a case-by-case basis.

Selection and allocation of F1 rat pups

Selection and allocation of F1 offspring prior to or at the time of weaning can influence overall data quality and study logistics, regardless of which test guideline is being followed. Table 1 presents the schema for assignment of F1 rat pups to various cohorts on an EOGRTS study. If all 3

cohorts are included on the study, as many as 70 rat pups per sex per group may be selected for F1 generation assessments, although only 20 rat pups per group (10/sex/group, representing as many litters as possible) are allocated for Cohort 2A and 2B for DNT assessments. Per OECD 443, the size of each litter may be standardised to yield, as nearly as possible, five males and five females per litter, to ensure that no littermates are selected for a given cohort. Even in rat strains of good fecundity, not all litters meet the expected size or sex distribution, and thus, most labs initiate these studies with 22–25 breeding pairs in the F0 (P) generation to ensure sufficient offspring, although up to 30 breeding pairs may be used for certain strains of rat. The latter somewhat diminishes the advantage of the OECD 443 study with respect to reducing animal use. Additional limitations in offspring numbers and sex ratio may arise when decreased fertility or substantial offspring mortality is noted in treatment groups, as is expected with the current ECHA dose selection recommendations for this study (see [Dose Level Selection](#)). The design of the EOGRTS (DNT cohort) was intended to provide only an initial assessment of DNT compared to the standalone OECD 426, due in part to the limited assessments and smaller group sizes in this cohort (Cooper et al. (2006)). These differences are illustrated in a side-by-side comparison of these study types in Table 2.

Various approaches can be used for selection and allocation of F1 offspring on a standalone DNT. Similar to the EOGRTS, litter sizes are standardised to yield uniform litter sizes, with equal numbers of male and female rat pups, where possible. In Appendix 1 of the published OECD 426, three different illustrative examples are provided to show the flexibility in animal assignments on this study. Animals can be assigned such that the same animals are tested at each interval for all behaviours tested, where appropriate (e.g., it is not appropriate to test the animals more than once in a test for L&M as the learning during the second interval of testing may be influenced by learning which occurred during the first interval). It is also possible to allocate the animals such that the various neurobehavioural tests are divided across 2 or 3 sets of animals, which may have the advantage of reducing interference across the various types of testing (i.e., auditory startle vs. motor activity vs. L&M). There are advantages and disadvantages to each of these approaches. Because neurobehavioural testing of animals at younger ages is often staggered over a 4–7-day period based on parturition stagger (i.e., animals bred and birthed over several days) in order to test all animals at the same developmental stage and ontogeny, the selection of different subsets of animals for each neurobehavioural test offers the additional advantage of removing overlapping testing periods. While some testing facilities have sufficient personnel and procedures in place to conduct longitudinal assessments

in a single subset of animals, these are generally considered more challenging than testing different subsets of animals.

When considering a combined EOGRT-DNT study, it is important to recognise that there are limitations in numbers of F1 offspring available for the inclusion of a full DNT study in combination with the existing design. With creative study design (selection of sufficient breeding pairs and pup allocations) and careful dose selection, an integrated study design may eliminate the need for the additional 100–120 time-mated female rats (and their offspring [1200–1600]). However, due to the need for pup allocations for reproductive toxicity assessments (Cohort 1A/1B, mandatory) and developmental immunotoxicity assessments (Cohort 3, optional) on the EOGRTS portion of a combined EOGRT-DNT study, one may be limited to allocation of only a single subset of rat pups to postweaning functional and neurobehavioural testing, i.e., it may only be possible to perform longitudinal assessments on the same subset of animals (Table 2, Example 1 (longitudinal assessments); Examples 2 and 3 (additional subsets)). This is an important consideration because Cohort 1A/1B is the primary reproductive cohort in an EOGRTS as the primary goal of this study is to provide an assessment of potential effects on sexual function and fertility. It is prudent to not assign these animals to additional testing since every additional procedure has the potential to induce additional stress on the animals and increases the likelihood of compromising the primary purpose of the study. This is especially important because several of the endpoints evaluated on Cohort 1A/1B (e.g., landmarks of sexual development, cyclicity, thyroid hormone levels, sperm parameters and male and female reproductive performance) have been demonstrated to be sensitive to the effects of excessive stress (Pollard et al. (1975); Suárez et al. (1996); Lerman et al. (1997)).

Behavioural ontogeny and functional assessments

Under OECD 426, at least two measures of behavioural ontogeny must be evaluated ideally in the same animals during both pre-weaning and adulthood using endpoints such as righting reflex, negative geotaxis, or motor activity. Although the EOGRTS does not specifically request behavioural ontogeny, these endpoints are often included with home cage and open field observations conducted during the assessment of a functional observational battery (FOB), which is required at young adulthood.

Functional assessments for motor and sensory function are required for DNT assessments. Motor activity is assessed in both a standalone 426 and EOGRTS although the latter only requires assessment at a single interval (adult animals in Cohort 2A). In contrast, standalone DNTs perform motor activity assessments repeatedly across developmental

timepoints as part of the assessment of behavioural ontogeny. In addition, the OECD 426 also requires motor and sensory function to be assessed at least once during the adolescent period and again during the adult period (Table 2). Sensory function, often evaluated using auditory startle response testing, is also only assessed once near the age of weaning (PND 24 ± 1) on an EOGRS (Table 2).

As mentioned previously, L&M assessments are an integral component of a standalone DNT and are required on individual subsets of male and female animals post-weaning and at young adulthood. OECD 443 originally did not require L&M assessments; however, under the ECHA BPR guidance for biocidal products (ECHA (2022b)), two ages of L&M assessment are now required for EOGRS conducted for biocidal products.

Though the guidelines allow some flexibility in the choice of L&M test(s), certain acceptability criteria should be met: (a) Learning should be evaluated as a change across multiple repeated learning trials (or sessions) and (b) the selected test should include a measure of either short-term or long-term memory, which cannot be interpreted without accompanying evidence of learning acquisition in the same test. Regulatory documents frequently cite the NAFTA TWG (NAFTA 2016) DNT guidance, which discusses allocentric and egocentric navigation tasks—including the Morris Water Maze (MWM) for allocentric learning, and the Cincinnati (CWM, 9 T) and Biel Water Maze (BWM, 6 T) for egocentric learning. Another test for allocentric navigation is the Radial Arm Maze (RAM) although its use in a regulatory safety context is limited. A detailed description and analysis of L&M paradigms in each of these mazes is outside the scope of this manuscript. However, in light of recent ECHA BPR requests to include L&M investigations on EOGRS studies, discussion regarding the use of these tests is warranted.

The CWM and BWM are both complex labyrinthine mazes with T-shaped cul-de-sacs, requiring sequences of left–right navigational choices to reach a hidden platform. The two mazes differ in size, design and complexity (based on the number of T-junctions). Testing can be done under visible, low, red lighting or infrared (complete darkness) conditions. Under visible light, rats may rely on both allocentric and egocentric cues, reducing test specificity; under infrared lighting, distal cues are eliminated, but testing duration increases dramatically (5 days of training under visible light vs. 15–20 days under infrared) (Vorhees and Williams 2024, 2015). Although there are currently no available reports for the operation of the BWM under infrared light, similar increases in learning duration are anticipated for animals trained in the dark. For a typical four group EOGRS with 10 males and 10 females per group, CWM testing may require >12 h/day of continuous trials, raising

concerns about diurnal activity related variability in the data (Stryjek et al. 2013) and major increases in labour and costs of the study. In contrast, the MWM requires significantly less time, and most animals learn the location of the escape platform over 16–20 learning trials (i.e., 4 to 5 days, with 4 trials/day).

While adult rats perform well in the MWM, there are limited data to support testing of juvenile rodents in this maze in a safety assessment setting. Hippocampal-dependent allocentric navigation is not reliably developed prior to PND 20 (Ehman and Moser (2006)) and juvenile testing requires methodological modifications (OECD (2008)) such as smaller pools (<5 feet diameter), adjusted platform sizes to adapt the complexity of the task, the use of warmer water (to avoid hypothermia), drying and warming rat pups between trials (to maintain body temperature), providing cued platform learning etc. (Roberts GK (2023); Ehman and Moser (2006)). Rats as young as 22–30 days have been used to evaluate L&M in smaller diameter mazes in academic settings (Yang et al. (2009); Zaichenko et al. (2025); Jett et al. (2001); Goodlett and Peterson (1995); Jett et al. (1997)), however, performance at this age is variable (Vorhees and Williams 2024) and there is limited evidence for use of a MWM at this age in a safety assessment setting. Similarly, there is limited evidence for use of a standard CWM (9–10 Ts) at this age in a safety assessment setting. Jablonski et al. (2019) demonstrated that a 10 T CWM was too difficult for PND 30 rats (none of the rats found the goal) and the same age performed poorly even when the maze was simplified by blocking off part of the maze to only 6Ts. In this scenario, rat pups were able to find the goal platform, however, even control rats failed to reach full proficiency even following 15 days of training and Jablonski et al. concluded that rats at this age (PND 30) may not have reached their full egocentric cognitive ability. However, there are reports that demonstrate successful performance of rats at this age in a BWM (6 T) in a safety assessment setting (Stump et al. (2010); Faber et al. (2007); Cruzan et al. (2005); Biesemeier et al. (2011); Butenhoff et al. (2009); Sheets et al. (2015)). Although the exact reason for the difference in performance between a 6 T CWM as tested by Jablonski and a 6 T BWM as tested by these authors is unclear, these differences could have arisen from a variety of variables including but not limited to room size and dimensions, construction differences between a partially blocked (10 T > 6 T) CWM and a 6 T BWM, the presence of a longer straight channel in a BWM or even differences in training and testing paradigms between laboratories. Overall, due to recent regulatory requests for L&M assessments in weanlings, industry laboratories continue to test and optimize weanling L&M performance in various types of mazes. It is clear however that the type of maze and the testing paradigm should be

selected based on multiple factors, including but not limited to (a) age of animals, (b) validation of the training and testing paradigms, (c) applicability and logistical appropriateness of the test in a safety assessment setting, (d) the ability to detect deficits in L&M following exposures to positive control substances, (d) the availability of relevant historical control data, and (e) the technical expertise of the performing laboratory.

Gross brain weights, measurements and neurohistopathology

Gross brain weights, brain measurements and neurohistopathology are required DNT assessments. Typically, these are performed around the time of weaning and at study (cohort) termination. Animals selected for weanling assessments are allocated primarily for central nervous (CN) tissue (brain) neurohistopathology whereas examination of both CNS and peripheral nervous system (PNS) tissues is required for animals necropsied at study (cohort) termination. While tissues for weanlings can be immersion fixed, perfusion fixation of the nervous system tissues is required for young adults since these tissues, especially the brain, are used for morphometric (quantitative) evaluations, which include linear (or sometimes areal) measurements of specific brain regions. In addition, OECD 426 also recommends inclusion of information regarding alterations in the gross size or shape of the olfactory bulbs, cerebrum, or cerebellum. Procedures for performing linear gross brain measurements of the intact brain in both weanlings and young adults are included in the Society of Toxicologic Pathology (STP) “best practice” recommendations for DNT neuropathology assessment (Bolon et al. (2006) and as revised by Garman et al. (2016).

The extent to which deficits in body weight or body weight gains impact brain weights is a subject of considerable debate. Some agencies consider any change in absolute brain weight to be evidence of adverse effects on neurodevelopment (NAFTA 2016), regardless of differences in pup body weights. Other agencies take into consideration whether brain weight deficits occurred in the presence of substantial body weight changes (Crofton (2024); Li et al. (2017)) and may even consider absolute brain weight deficits to be non-adverse when confounded by alterations in body weight.

In a recent analysis sponsored by EFSA, Crofton (2024) surveyed neonatal body weights and brain weights from 173 publicly available DNT studies conducted according to EPA or OECD DNT guidelines. Out of 173 studies, 122 studies (70%) showed a decrease in body weights in rat pups, but only 50 of these studies (41%) reported concomitant decreased brain weights. This analysis led to the conclusion that body weight is an unreliable predictor of alterations in

brain weight and concluded that the use of a brain weight to body weight ratio is not appropriate as a common approach to dismiss brain weight as a DNT effect. This conclusion may be inadequate given that the Crofton analysis did not consider (or specify) whether the brain weights used in the analysis were from perfusion- or immersion-fixed brains (which can lead to significant variability in brain weights). There was also no discussion of whether or not maternal care was impacted on individual studies; as summarized by Garman et al. (2016), maternal care issues impacting pup growth can produce long lasting effects on brain weights. It is important to recognise that high dose exposures can impact maternal care and cause substantial initial decreases in pup body weight gain or absolute body weight. In studies where rat pups and/or dams are exposed soon after parturition, the ability of the rat pups to nurse and/or the dams’ ability to care for their offspring can be impacted by systemic toxicity. This is an especially important consideration for studies such as a EOGRTS or a combined EOGRT-DNT (see *Dose Selection*) where the regulatory recommendation is to select dose levels high enough to cause systemic toxicity (but not death or severe suffering). According to the best practices approach to neuropathologic assessments in DNTs, Bolon et al. (2006) state that absolute brain weight and morphometric values can be compared with relative values based on brain weight or brain volume when evaluating whether alterations might be either localised or related to a more general effect on growth.

Statistical analysis for neurobehavioural and neuropathology endpoints

The OECD 443 does not provide specific guidance on the statistical analyses for neurobehavioural and neuropathology endpoints and instead refers the reader to OECD 426 for further guidance. Both guidelines specify that the litter must be the experimental (and statistical) unit, and repeated measures should be analysed to account for non-independence of those measures. On developmental and reproductive studies, the litter (and not the pup) has been used as the statistical unit of measure for decades due to the clustering of variance among offspring in rodents (Vorhees and Williams (2020)). Experiments should be designed such that measurements on littermates are not treated as independent observations. Treating the litter as a unit allows for the use of representative rat pups within each litter for different tests (see *Selection and Allocation of F1 Rat Pups*).

Functional and cognitive DNT endpoints are typically collected from the same subjects at more than one time in relation to treatment (Maurissen and Vidmar (2017)). These repeated measures could reflect multiple intervals within a study (e.g., two or more ages of motor activity assessments)

or multiple trials within a single session (e.g., assessment of habituation across each testing interval in an auditory startle response test). Tests for L&M also have repeated measures within a session (multiple learning trials over each day or multiple days of learning). Both guidelines indicate that separate analysis of each individual interval or trial is inappropriate. For example, statistically analysing each 10-min interval for motor activity separately from the rest is inappropriate as an initial analysis of the data, because progressive activity of the animals in the enclosures is a continuum of observations.

In 2016, a more specific internal guidance document was written by the NAFTA Technical Working Group (TWG) on Pesticides, comprised of scientists from the United States Environmental Protection Agency (US EPA) and Health Canada (NAFTA 2016). This guidance was written for the standalone DNT but continues to serve as a useful resource for behavioural endpoints on any study with DNT assessments, with the following caveats. It recommends that all analysis for neurobehaviour should be conducted with males and females combined which is especially helpful for EOGRTS Cohort 2A, where only one male or one female from each litter is tested (10/sex/group, representing as many litters as possible). Combining data for males and females doubles the sample size from 10/sex to a total of 20 rat pups, each representing a different litter. However, combining different factors (e.g., sex) into a single global analysis requires that the investigator ensure that the equipment settings and experimental design support pooling all the data together (Holson et al. (2008)). If equipment or equipment settings are different based on the sex or age (due to size differences), it may not be appropriate to combine data into one analysis across these factors. For example, in some laboratories the auditory startle equipment settings are optimised for males and females differently so that the full range of responses for males vs females can be detected. In such cases, it may not be appropriate to combine both sexes. To combine data for males and females in a single analysis, it is also important to counterbalance the time of testing not just across group but also across both sexes. Failure to do so is likely to confound interpretation of the sex effect or dose interaction with sex due to the divergent testing conditions (Maurissen (2010)). If there is a statistically significant sex interaction, it is important to analyse each sex separately within the framework of the original analysis to take advantage of the full power of this combined analysis, even in subsequent post-hoc analyses.

New approach methodologies (NAMs) for DNT: current status, application and limitations

The literature was reviewed for information on new approach methods (e.g. *in silico*, *in vitro* etc.) for assessment of DNT. Several alternative methods identified are presented in the Supplemental Information. These are increasingly considered within applications of next generation risk assessment (NGRA, in which NAM-derived hazard, kinetics and exposure data are combined to drive substance specific risk assessment decisions (Leist et al. (2026))), and the development of these is a clear objective of the EU Commission's Roadmap toward phasing out animal testing for chemical safety (European Commission (2026)).

No robust quantitative structure activity relationship (QSAR) models for assessing DNT could be found in the literature, and previous work has identified that there are no predictive *in silico* computational DNT models (Crofton et al. (2022) as cited in OECD (2023)). This is likely due to the lack of available *in vivo* data. For an appropriate *in silico* tool (e.g., QSAR) to be developed, a large dataset of *in vivo* studies on well-characterised substances would be needed.

The most recognised but not yet fully validated *in vitro* approach for assessing DNT is the Developmental Neurotoxicity In Vitro Battery (DNT IVB), which was developed within an international, inter-organisation programme coordinated by the OECD, with the aim of predicting DNT outcomes without the use of laboratory animals. It is acknowledged that the DNT IVB does not cover all molecular or cellular components that are involved in adverse neurodevelopmental outcomes and most assays included are based on downstream effects instead of initiating molecular events and have not yet been formally validated. In addition, there is no analysis to date correlating changes *in vitro* with *in vivo* neurodevelopmental outcomes (OECD 2023).

However, with increasing data generation using the DNT IVB and further Adverse Outcome Pathway (AOP) development these uncertainties are expected to be reduced. AOPs serve as the mechanistic link between NAM data and adverse outcomes (AOs) on which NGRA depends. Spinu et al. (2019) neurotoxicity AOP network identifies learning and memory impairment as the principal convergence point, reached from multiple initiating events including thyroperoxidase and sodium/iodide symporter inhibition, whilst also showing present constraints. The links between key events remain largely qualitative without the dose response data risk assessment requires.

The DNT IVB testing cannot currently replace *in vivo* testing but may be used to screen and inform on follow-up testing and it can also be used in WoE approaches for DNT hazard assessments (OECD 2023). OECD work continues

on new DNT in vitro assays, standardisation, methodological refinement and regulatory integration of the DNT IVB (OECD (2025a), Sachana et al. (2026)). The authors of this review note that when a substance's DNT effect is clearly linked to its interaction with a specific enzyme or receptor, in vitro assays targeting this interaction can be useful for screening chemical libraries to identify analogues that do not exhibit it. However, when such a mechanistic understanding is not available, it is challenging to make conclusions on potential toxicity.

The ECHA June 2024 and 2026 Key Areas of Regulatory Challenge (ECHA Karc (2026)) (ECHA 2024) identified a need for a better understanding of developmental and adult neurotoxicity. The KARC proposed a focus on NAMs development, including further validation and refinement of the DNT IVB battery and identification of reliable positive and negative reference chemicals for NAM validation. It also proposed research on new AOPs, further development of existing AOPs and establishing their interlink with NAMs.

The zebrafish (*Danio rerio*) embryo provides a whole organism non-mammalian model that possesses 70% concordance with human genes (Howe et al. 2013) and has been recognised as a valuable model to study DNT. Changes in zebrafish neurobehaviour and morphology (such as spontaneous swimming, learning, normal eye development) can be easily observed and paired with the analysis of gene expression, neural morphogenesis and neurobehaviour, which provides opportunity to test for potential developmental neurotoxicants (Chen et al. 2024; Fu et al. 2024; Ikeda et al. 2023). Zebrafish embryo developmental toxicity assays are reported widely in the literature, and validation activities are progressing rapidly, however the addition of neurodevelopmental parameters is less mature in the literature (Alzualde et al. 2018; Weiner et al. 2024). An international collaboration led by the Netherlands aims to harmonise behavioural testing in eleuthero embryos of zebrafish and to explore incorporation into the DNT IVB. In addition, an interlaboratory ring test is underway to validate the zebrafish embryo DNT assay using light/ dark transition behaviour testing (Massei 2024, Muriana, 2023).

It is noted that zebrafish embryo tests may not be considered by all stakeholders as an alternative method because animals are used, though zebrafish embryos up to a certain age are not considered a protected life stage under EU Directive 2010/63 on the protection of animals used for scientific purposes, which covers independently feeding larval forms.

A survey of eleven contract research organisations, international and Europe-based, was conducted by the authors to understand the capacity for alternative testing methods for DNT and their commercial availability. No laboratories responded with commercial availability to conduct

the assays included in the OECD DNT IVB, despite this being the only alternative method with regulatory recognition, at the time of this publication. DNT testing of zebrafish embryo/larvae (not older than 5 days post fertilisation) is offered commercially by at least one commercial laboratory. As mentioned above, the regulatory acceptance of this assay is in the early stages, with validation activities ongoing.

This review indicates that in vivo testing remains the default approach for assessing developmental neurotoxicity potential of substances, as in silico approaches are not yet available as in vitro approaches still require further development and validation. This should not be interpreted that in vitro and in vivo testing could be a one-to-one substitution. NGRA aims to be protective without superseding in vivo testing directly (Leist et al. (2026)), and NAMS are increasingly contributing to early screening, prioritisation and weight of evidence.

Conclusions and recommendations

As outlined in the introduction, there have been instances in which additional cognitive function testing (L&M) has been requested by ECHA as part of the EOGRS with the DNT cohort (Cohort 2), thereby extending the testing requirements beyond those intended under OECD 443. ECHA defines the EOGRS (OECD 443) as a fertility study, such that priority is given to Cohort 1 over Cohorts 2 and 3 with respect to dose level selection and pup allocation in cases of insufficient rat pups. Such design priorities can compromise the conduct and interpretation of DNT assessments. From a DNT perspective, this may result in inappropriate dose levels, unsuitable dosing paradigms, and insufficient animal numbers in Cohort 2. As these data are used for classification and labelling under REACH, as well as for risk assessment purposes, it is essential that the data generated are interpretable. A compromised Cohort 2 dataset is not consistent with good scientific practice and is therefore not recommended. However, where assessment of DNT as part of an EOGRS is a clear regulatory requirement, then the following recommendations support a more scientifically robust approach, increase the likelihood of generating interpretable data, and ensure the most appropriate use of animals.

Study design

- The OECD 426 is the globally accepted standard for DNT assessment. Despite interest in integrated approaches, a standalone DNT provides the most scientifically robust evaluation of developmental neurotoxicity.

- Although combining a DNT assessment within an EOGRTS (OECD 443) may be feasible it is scientifically and logistically challenging. Significant modifications to study design, including dose selection, exposure duration, pup allocation, and endpoint assessment, are required to generate interpretable data. Even then such an assessment of DNT would be considered scientifically sub-optimal.

Learning and memory (L&M) assessment

- L&M should only be evaluated, when scientifically justified on a substance specific basis, in an EOGRTS if there is clear evidence that such testing would resolve or further clarify an issue that would otherwise require a standalone DNT study.
- Where conducted, L&M assessments should only be conducted in adult animals and scheduled independently of final motor activity assessments to avoid overlap and ensure valid evaluation of complex behavioural patterns.
- At present, the MWM and the BWM are the most widely used and validated L&M tests available in safety assessment laboratories. Should alternate tests be recommended it is important to assess whether they meet equivalent validation requirements for use for safety assessment purposes.

Exposure duration

- To ensure the generation of reliable and unambiguous DNT data, Cohort 2A animals in an EOGRTS should not be dosed beyond weaning (PND 21), prioritising their purpose for the assessment of true DNT as opposed to general neurotoxicity.
- Whilst such animals would no longer be suitable to assess the timing of puberty or other endpoints unrelated to neurodevelopment, these endpoints can be adequately assessed using animals from other cohorts without impacting the core objectives of the EOGRTS.

Animal numbers and statistical considerations

- Animal numbers can be reduced to a manageable level when incorporating aspects of DNT into an EOGRTS. OECD 426 recommends that 20 litters per group should be used for DNT assessments. If the experimental methods and equipment settings are identical for males and females, then only 10 male and 10 female rat pups can

be selected such that there is one male or female per litter (representative of 20 litters in total per dose group).

- Combining data from both sexes into a single analysis may increase statistical power to detect treatment-related effects (Li et al. (2017)). If this approach is used, then the experimental design must balance the time of testing for males and females so that statistical analysis can combine data from both sexes together in a repeated measures analysis of variance that includes sex of animal as a factor in the analysis (Maurissen (2010)). Post-hoc analysis can be done within the framework of the combined analysis to maintain the power even if a sex vs. dose interaction is present.
- Validation would be needed of the 10 rats/sex for all behavioural tests using statistics described above. This can be achieved using existing DNT datasets and simulating selection of individual animals such that only 1 male or 1 female per litter is allocated to each assessment.

Endpoint selection

- To reduce study complexity and overlap, the functional observational battery currently required for Cohort 2A between PND 63 and PND 75 can be removed or limited to key observations such as gait, foot splay, and righting reflex.
- Routine assessment of brain morphometry and neuropathology in Cohort 2B (PND 22) could be omitted, with evaluation instead limited to adult offspring (Cohort 2A), recognising that such a study design allows evaluation of effects across all stages of life but cannot be interpreted as a specific DNT effect.

Direct pup dosing

- Direct exposure to rat pups should not be conducted routinely due to confounding effects of acute bolus doses and potential physical trauma to rat pups.
- Where unavoidable, the first dose should not be administered before PND 10, and determination of overt toxicity following direct dosing in juvenile rat pups should be completed before commencement of the study.

Dose selection philosophy

- Mathematically focused approaches to dose selection designed to maximise the chance of detecting a hazard by exaggerating the top dose level of an EOGRT study (e.g., Van Berlo et al. (2022)) should be avoided, as such

approaches fail to account for the integrated biology of the test organism.

- The complexity and interdependencies of DART experimental models have led to the development of scientifically robust approaches to dose level selection in DART studies including consensus limits on reductions in body weight gain, with >20% considered excessive and 10–15% reduction adequate for the highest dose level.
- Replacement of MTD-based criteria of a 10% decrease in body weight *gain* with absolute decreases in *body weight* is not compatible with the aims of DART studies. (van Berlo et al. (2022), Lewis et al. (2024))
- Definitions of severe suffering that permit mortality (e.g., >10%) are inconsistent with dose level selection approaches outlined in OECD 414, 421/422, and 443, as well as humane endpoint guidance, which indicate that death should be avoided. Any degree of mortality cannot be supported on animal welfare grounds, nor on scientific grounds due to the risk of generating non-specific, uninterpretable data.
- It is recommended that a clear and transparent justification of dose selection should be included in regulatory submissions to support interpretation by regulatory authorities.

Alternative methods

The status of New Approach Methods (NAMs) (e.g. in silico and in vitro methods) for assessment of DNT has been extensively reviewed (see Supplemental Information). These analyses confirm that in vivo testing remains the primary method for evaluating DNT potential of substances, as in silico tools are not yet established and current in vitro methods require further development and validation before they can serve as reliable alternatives for safety assessment purposes.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00204-026-04564-z>.

Acknowledgements Dorothee Funk-Weyer (BASF SE, Germany), Steward from the ECETOC Scientific Committee until February 2025, and Steffen Schneider (BASF, Germany), member of the ECETOC Scientific Perspective on in vivo DNT TF as of May 2025, are thanked for fruitful discussions.

Author contributions All co-authors contributed to the writing of this manuscript. The final manuscript is the result of all authors commenting upon and helping revise the original draft; it was reviewed and edited by all authors.

Funding This manuscript relates to work undertaken by the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC; www.ecetoc.org) Scientific Perspective on in vivo DNT Task Force. ECETOC is a not-for-profit scientific organisation that provides a col-

laborative space for scientists from industry, academia, and regulatory arenas to develop and promote trusted and practical scientific solutions to chemical safety assessment. ECETOC is financed by its membership, which are leading companies with interests in the manufacture and use of chemicals, biomaterials, and pharmaceuticals (<https://www.ecetoc.org/about-us/our-members/>). Within the ECETOC Task Forces, Task Force members do not receive compensation from ECETOC.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest ECETOC Task Force members aim to provide an objective, scientific assessment of the topic. While some members of this Task Force work for companies that produce chemicals, there are no potential conflicts of interest between these affiliations and the content of the manuscript.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Albee RR, Mattsson JL, Yano BL, Chang LW (1987) Neurobehavioral effects of dietary restriction in rats. *Neurotoxicol Teratol* 9:203–211
- Alzualde A, Behl M, Sipes NS, Hsieh JH, Alday A, Tice RR, Paules RS, Muriana A, Quevedo C (2018) Toxicity profiling of flame retardants in zebrafish embryos using a battery of assays for developmental toxicity, neurotoxicity, cardiotoxicity and hepatotoxicity toward human relevance. *Neurotoxicol Teratol* 70:40–50
- Arts J, Beekhuijzen M (2020) Is there a rationale for direct dosing of chemicals to nursing pups in the EOGRTS (OECD 443)? *Regul Toxicol Pharmacol* 113:104641
- Arts JHE, Faulhammer F, Schneider S, Salverda JGW (2023) Investigations on learning and memory function in extended one-generation reproductive toxicity studies – when considered needed and based on what? *Crit Rev Toxicol* 53:372–384
- Beekhuijzen M, Barentsen H, Marsden E, Zmarowski A, Aujoulat M, Picut C, Slotter E (2016a) Implementing the extended one-generation reproductive toxicity study (EOGRTS): important points to consider. *Crit Rev Toxicol* 46:332–347
- Beekhuijzen M, VAN Otterdijk F, Wieland W, VAN Tuyl M, Rijcken RP, Peter B, Emmen H (2016b) Update of OECD DART guidelines with endocrine disruptor relevant endpoints: practical considerations. *Reprod Toxicol* 64:64–71
- Beekhuijzen M, Richmond E, Manton J, Coder PS, Goyak K, Ghaf-fari R, Makris SL, VAN Cruchten S, Zorrilla L, Mitchell-Ryan S (2024) Review of dose setting for the extended one-generation

- reproductive toxicity studies (OECD TG 443): considerations on ECHA's dose level selection recommendations. *Regul Toxicol Pharmacol* 151:105665
- Beyer BK, Chernoff N, Danielsson BR, Davis-Bruno K, Harrouk W, Hood RD, Janer G, Liminga UW, Kim JH, Rocca M, Rogers J, Scialli AR (2011) ILSI/HESI maternal toxicity workshop summary: maternal toxicity and its impact on study design and data interpretation. *Birth Defects Res B* 92:36–51
- Biesemeier JA, Beck MJ, Silberberg H, Myers NR, Ariano JM, Radvosky A, Freshwater L, Sved DW, Jacobi S, Stump DG, Hardy ML, Stedeford T (2011) An oral developmental neurotoxicity study of decabromodiphenyl ether (DecaBDE) in rats. *Birth Defects Res B Dev Reprod Toxicol* 92:17–35
- Bolon B, Garman R, Jensen K, Krinke G, Stuart B (2006) A “best practices” approach to neuropathologic assessment in developmental neurotoxicity testing—for today. *Toxicol Pathol* 34:296–313
- Brannen K, Fenton SE, Hansen DK, Harrouk W, Kim JH, Shuey D (2011) Developmental toxicology—new directions workshop: refining testing strategies and study designs. *Birth Defects Res B Dev Reprod Toxicol* 92:404–412
- Butenhoff JL, Ehresman DJ, Chang S-C, Parker GA, Stump DG (2009) Gestational and lactational exposure to potassium perfluorooctanesulfonate (K+PFOS) in rats: developmental neurotoxicity. *Reprod Toxicol* 27:319–330
- Charil A, Laplante DP, Vaillancourt C, King S (2010) Prenatal stress and brain development. *Brain Res Rev* 65:56–79
- Chen H, Wang Y, Liang H (2024) The combined neurotoxicity of DBP and nano-TiO₂ in embryonic zebrafish (*Danio rerio*) revealed by oxidative activity, neuro-development genes expression and metabolomics changes. *Aquat Toxicol* 269:106881
- Cooper RL, Lamb JC, Barlow SM, Bentley K, Brady AM, Doerrer NG, Eisenbrandt DL, Fenner-Crisp PA, Hines RN, Irvine LF, Kimmel CA, Koeter H, Li AA, Makris SL, Sheets LP, Speijers G, Whitby KE (2006) A tiered approach to life stages testing for agricultural chemical safety assessment. *Crit Rev Toxicol* 36:69–98
- Crofton K (2024) Causation or canard: use of brain weight to body weight ratios in developmental neurotoxicity. *EFSA Support Publ* 21:9098E
- Crofton KM, Bassan A, Behl M, Chushak YG, Fritsche E, Gearhart JM, Marty MS, Mumtaz M, Pavan M, Ruiz P, Sachana M (2022) Current status and future directions for a neurotoxicity hazard assessment framework that integrates in silico approaches. *Comput Toxicol* 1(22):100223
- Cruzan G, Faber WD, Johnson KA, Roberts LS, Hellwig J, Maurissen J, Beck MJ, Radvosky A, Stump DG (2005) Developmental neurotoxicity study of styrene by inhalation in Crl-CD rats. *Birth Defects Res B Dev Reprod Toxicol* 74:221–232
- Desesso JM (1987) Maternal factors in developmental toxicity. *Teratog Carcinog Mutagen* 7:225–240
- Desesso JM, Watson RE, Keen CL, Hazelden KP, Haws LC, Li AA (2008) Analysis and integration of developmental neurotoxicity and ancillary data into risk assessment: a case study of dimethoate. *J Toxicol Environ Health A* 72:94–109
- ECHA 2022a. Advice on dose-level selection for the conduct of reproductive toxicity studies (OECD TGs 414, 421/422 and 443) under REACH
- ECHA 2022b. Guidance on the biocidal products regulation volume III: human health part A: information requirements version 2, Mar 2022. Reference: ECHA-22-H-09-EN. <https://doi.org/10.2823/433526>
- ECHA (2024) Key areas of regulatory challenge. ECHA-23-R-08-EN
- Ehman KD, Moser VC (2006) Evaluation of cognitive function in weanling rats: a review of methods suitable for chemical screening. *Neurotoxicol Teratol* 28:144–161
- US EPA (1998) Health effects test guidelines, OPPTS 870.6300, developmental neurotoxicity study. EPA 712-C-98-239
- European Commission (2026) communication from the commission—roadmap towards phasing out animal testing for chemical safety assessments
- Faber WD, Roberts LSG, Stump DG, Beck M, Kirkpatrick D, Regan KS, Tort M, Moran E, Banton M (2007) Inhalation developmental neurotoxicity study of ethylbenzene in Crl-CD rats. *Birth Defects Res B Dev Reprod Toxicol* 80:34–48
- Fegert I, Billington R, Botham P, Carney E, Fitzgerald RE, Hanley T, Lewis R, Marty MS, Schneider S, Sheets LP, Stahl B, VAN Ravenzwaay B (2012) Feasibility of the extended one-generation reproductive toxicity study (OECD 443). *Reprod Toxicol* 34:331–339
- Fu R, Liu H, Zhang Y, Mao L, Zhu L, Jiang H, Zhang L, Liu X (2024) Imidacloprid affects the visual behavior of adult zebrafish (*Danio rerio*) by mediating the expression of opsin and phototransduction genes and altering the metabolism of neurotransmitters. *Sci Total Environ* 910:168572
- Garman RH, Li AA, Kaufmann W, Auer RN, Bolon B (2016) Recommended methods for brain processing and quantitative analysis in rodent developmental neurotoxicity studies. *Toxicol Pathol* 44:14–42
- Goodlett CR, Peterson SD (1995) Sex differences in vulnerability to developmental spatial learning deficits induced by limited binge alcohol exposure in neonatal rats. *Neurobiol Learn Mem* 64:265–275
- Henck JW, Morford LL (2010) 12.19 - Developmental neurotoxicology. In: McQueen CA (ed) *Comprehensive toxicology*, Second Edition. Elsevier, Oxford
- Holson RR, Freshwater L, Maurissen J, Moser VC, Phang W (2008) Statistical issues and techniques appropriate for developmental neurotoxicity testing: a report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. *Neurotoxicol Teratol* 30:326–348
- Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, McLaren S, Sealy I, Caccamo M, Churcher C, Scott C, Barrett JC, Koch R, Rauch GJ, White S, Chow W, Kilian B, Quintais LT, Guerra-Assunção JA, Zhou Y, Gu Y, Yen J, Vogel JH, Eyre T, Redmond S, Banerjee R, Chi J, Fu B, Langley E, Maguire SF, Laird GK, Lloyd D, Kenyon E, Donaldson S, Sehra H, Almeida-King J, Loveland J, Trevanion S, Jones M, Quail M, Willey D, Hunt A, Burton J, Sims S, McLay K, Plumb B, Davis J, Clee C, Oliver K, Clark R, Riddle C, Elliot D, Threadgold G, Harden G, Ware D, Begum S, Mortimore B, Kerry G, Heath P, Phillimore B, Tracey A, Corby N, Dunn M, Johnson C, Wood J, Clark S, Pelan S, Griffiths G, Smith M, Glithero R, Howden P, Barker N, Lloyd C, Stevens C, Harley J, Holt K, Panagiotidis G, Lovell J, Beasley H, Henderson C, Gordon D, Auger K, Wright D, Collins J, Raisin C, Dyer L, Leung K, Robertson L, Ambridge K, Leongamornlert D, McGuire S, Gilderthorp R, Griffiths C, Manthavadi D, Nichol S, Barker G et al (2013) The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496:498–503
- Ikeda D, Fujita S, Toda K, Yaginuma Y, Kan-No N, Watabe S (2023) Cold-induced muscle atrophy in zebrafish: insights from swimming activity and gene expression analysis. *Biochem Biophys Rep* 36:101570
- Jablonski SA, Williams MT, Vorhees CV (2019) Learning and memory effects of neonatal methamphetamine exposure in Sprague–Dawley rats: test of the role of dopamine receptors D1 in mediating the long-term effects. *Dev Neurosci* 41:44–55
- Jett DA, Kuhlmann AC, Farmer SJ, Guilarte TR (1997) Age-dependent effects of developmental lead exposure on performance in the Morris water maze. *Pharmacol Biochem Behav* 57:271–279
- Jett DA, Navoa RV, Beckles RA, Mclemore GL (2001) Cognitive function and cholinergic neurochemistry in weanling rats exposed to chlorpyrifos. *Toxicol Appl Pharmacol* 174:89–98

- ECHA KARC (2026) Key areas of regulatory challenge, European Chemicals Agency
- Ladies GS, Chapin RE, Hastings KL, Holsapple MP, Makris SL, Sheets LP, Woolhiser MR, Burns-Naas LA (2005) Developmental toxicology evaluations—issues with including neurotoxicology and immunotoxicology assessments in reproductive toxicology studies. *Toxicol Sci* 88:24–29
- Leist M, Tangianu S, Affourtit F, Braakhuis H, Colbourne J, Cöllen E, Dreser N, Escher SE, Gardner I, Hahn S, Hardy B, Herzler M, Islam B, Kamp H, Magel V, Marx-Stoelting P, Moné MJ, Lundquist P, Ottenbros I, Ouedraogo G, Pallocca G, van de Water B, Vinken M, White A, Pastor M, Luijten M (2026) An alternative safety profiling algorithm (ASPA) to transform next generation risk assessment into a structured and transparent process. *ALTEX Alternat Animal Exp* 43:158–175
- Lerman SA, Miller GK, Bohlman K, Albaladejo V, Léonard J-F, Devas V, Clark RL (1997) Effects of corticosterone on reproduction in male Sprague–Dawley rats. *Reprod Toxicol* 11:799–805
- Lewis RW, Andrus AK, Arroyo J, Brescia S, Botham PA, Corvaro M, Daston GP, Hofmann T, Rodriguez C, Sewell F, VAN Ravenzwaay B, Wiench K, Marty S (2024) Considerations for the development of guidance on dose level selection for developmental and reproductive toxicity studies. *Regul Toxicol Pharmacol* 148:105585
- Li AA, Sheets LP, Raffaele K, Moser V, Hofstra A, Hoberman A, Makris SL, Garman R, Bolon B, Kaufmann W, Auer R, Lau E, Vidmar T, Bowers WJ (2017) Recommendations for harmonization of data collection and analysis of developmental neurotoxicity endpoints in regulatory guideline studies: proceedings of workshops presented at Society of Toxicology and joint Teratology Society and Neurobehavioral Teratology Society meetings. *Neurotoxicol Teratol* 63:24–45
- Mangas I, Spilioti E, Terron A, Panzarea M, Nepal M, Viviani B, Binaglia M, Crofton KM (2025) European Food Safety Authority database for in vivo developmental neurotoxicity studies of pesticides. *Crit Rev Toxicol* 55:587–598
- Markham JA, Taylor AR, Taylor SB, Bell DB, Koenig JI (2010) Characterization of the cognitive impairments induced by prenatal exposure to stress in the rat. *Front Behav Neurosci* 4:173
- Massei R, Christou M, Schroeder K, Ottermanns R, Domingues I, Pavlaki M, Monteiro M, Zwart E, Hessel E, Johann S, Knapen D, Vergauwen L, Huyghe K, Kjør PA, Holbech H, Molina B, Muriana A, Sebire M, & Lillicrap A (2024) Towards the standardization of the dark/light transition test with danio rerio embryos—preliminary results from an inter-laboratory ring test. SETAC Europe 34th Annual Meeting (setac eu), Sevilla. Zenodo. <https://doi.org/10.5281/zenodo.12095147>
- Maurissen J (2010) Practical considerations on the design, execution and analysis of developmental neurotoxicity studies to be published in neurotoxicology and teratology. *Neurotoxicol Teratol* 32:121–123
- Maurissen JP, Vidmar TJ (2017) Repeated-measure analyses: which one? A survey of statistical models and recommendations for reporting. *Neurotoxicol Teratol* 59:78–84
- Moore NP, Beekhuijzen M, Boogaard PJ, Foreman JE, North CM, Palermo C, Schneider S, Strauss V, VAN Ravenzwaay B, Poole A (2016) Guidance on the selection of cohorts for the extended one-generation reproduction toxicity study (OECD test guideline 443). *Regul Toxicol Pharmacol* 80:32–40
- Muriana A, Molina-Martínez B, Del Pozo A, Alzualde A, Schiavone V, Di Donato V, Sanz CM, Ibarra J, Truong L, Tanguay R, Hsieh J-H, Ryan K, Ellis L, Kluver N, Padilla S, Hill B, Shafer T, Behl M, Woodland C, Sachana M, Hessel E (2023) Harmonizing embryonic zebrafish light/dark behavior test for the oecd developmental neurotoxicity in vitro. Poster presented June 2023 at the birth defect research and prevention (BDRP) meeting
- NAFTA (2016) North American free trade agreement (NAFTA) technical working group on pesticides (TWG) developmental neurotoxicity study guidance document
- OECD (2007) Test no. 426: developmental neurotoxicity study. OECD guidelines for the testing of chemicals. Section 4. OECD Publishing, Paris
- OECD (2018) Test no. 443: extended one-generation reproductive toxicity study. OECD guidelines for the testing of chemicals. Section 4. OECD Publishing, Paris
- OECD (2008) Guidance Document on mammalian reproductive toxicity testing and assessment, OECD series on testing and assessment, No. 43, OECD Publishing, Paris
- OECD (2023) Initial recommendations on evaluation of data from the developmental neurotoxicity (DNT) in-vitro testing battery. OECD environment, health and safety publications. Series on Testing and Assessment. No. 377. 3 November 2023. ENV/CBC/MONO(2023)13
- OECD (2025a) In vitro assays for developmental neurotoxicity [Online]. Available: <https://www.oecd.org/en/topics/sub-issues/testing-of-chemicals/in-vitro-assays-for-developmental-neurotoxicity.html> [Accessed]
- OECD (2025b) Test No. 443: extended one-generation reproductive toxicity study, OECD guidelines for the testing of chemicals, section 4 [Online]. Available: https://www.oecd.org/en/publications/test-no-443-extended-one-generation-reproductive-toxicity-study_y_9789264185371-en.html [Accessed].
- Oldenburger MM, Doomen MJ, Lourens NJJ, Beekhuijzen M (2022) Unnecessary use of additional animals for determination of sexual maturation in the EOGRTS. *Reprod Toxicol* 112:14–22
- Piersma AH, Tonk ECM, Makris SL, Crofton KM, Dietert RR, van Loveren H (2012) Juvenile toxicity testing protocols for chemicals. *Reprod Toxicol* 34:482–486
- Pollard I, White BM, Bassett JR, Cairncross KD (1975) Plasma glucocorticoid elevation and desynchronization of the estrous cycle following unpredictable stress in the rat. *Behav Biol* 14:103–108
- Raffaele KC, Rowland J, May B, Makris SL, Schumacher K, Scarano LJ (2010) The use of developmental neurotoxicity data in pesticide risk assessments. *Neurotoxicol Teratol* 32:563–572
- Roberts GK, SM (2023) Specifications for the conduct of toxicity studies by the division of translational toxicology at the national institute of environmental health sciences
- Sachana M, Hogberg HT, Mangas I (2026) From concept to regulation: OECD's role in the development of the DNT in vitro battery. *Front Toxicol* Vol 8
- Saghir SA, Dorato MA (2016) Reproductive and developmental toxicity testing: examination of the extended one-generation reproductive toxicity study guideline. *Regul Toxicol Pharmacol* 79:110–117
- Salverda JGW (2024) It is the dose that makes the poison—a critical view on dose level setting in extended-one generation reproductive toxicity (EOGRT) studies. *JNST* 1(1):1–4
- Schmitt G, Barrow P (2022) Considerations for and against dosing rodent pups before 7 days of age in juvenile toxicology studies. *Reprod Toxicol* 112:77–87
- Schneider S, Kaufmann W, Strauss V, VAN Ravenzwaay B (2011) Vinclozolin: a feasibility and sensitivity study of the ILSI-HESI F1-extended one-generation rat reproduction protocol. *Regul Toxicol Pharmacol* 59:91–100
- Schulz F, Batke M, Mangelsdorf I, Pohlenz-Michel C, Simetska N, Lewin G (2014) Sensitivity of different generations and developmental stages in studies on reproductive toxicity. *Toxicol Lett* 226:245–255
- Semple BD, Blomgren K, Gimlin K, Ferriero DM, Noble-Haesslein LJ (2013) Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. *Prog Neurobiol* 106–107:1–16

- Sheets LP, Li AA, Minnema DJ, Collier RH, Creek MR, Peffer RC (2015) A critical review of neonicotinoid insecticides for developmental neurotoxicity. *Crit Rev Toxicol* 46:153–190
- Spinu N, Bal-Price A, Cronin MTD, Enoch SJ, Madden JC, Worth AP (2019) Development and analysis of an adverse outcome pathway network for human neurotoxicity. *Arch Toxicol* 93:2759–2772
- Stryjek R, Modlińska K, Turlejski K, Pisula W (2013) Circadian rhythm of outside-nest activity in wild (WWCPS), albino and pigmented laboratory rats. *PLoS ONE* 8:e66055
- Stump DG, Beck MJ, Radovsky A, Garman RH, Freshwater LL, Sheets LP, Marty MSJR, Waechter JM, Dimond SS, VAN Miller JP, Shiotsuka RN, Beyer D, Chappelle AH, Hentges SG (2010) Developmental neurotoxicity study of dietary bisphenol A in Sprague-Dawley rats. *Toxicol Sci* 115:167–182
- Suárez M, Fiol De Cuneo M, Vincenti L, Ruiz RD (1996) Changes in corticosterone levels and sperm functional activity by chronic stress in rats. *Arch Physiol Biochem* 104:351–356
- Tyl RW, Crofton K, Moretto A, Moser V, Sheets LP, Sobotka TJ (2008) Identification and interpretation of developmental neurotoxicity effects: a report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. *Neurotoxicol Teratol* 30:349–381
- Van Berlo D, Woutersen M, Muller A, Pronk M, Vriend J, Hakkert B (2022) 10% Body weight (gain) change as criterion for the maximum tolerated dose: a critical analysis. *Regul Toxicol Pharmacol* 134:105235
- Vorhees CV, Williams MT (2015) Reprint of “Value of water mazes for assessing spatial and egocentric learning and memory in rodent basic research and regulatory studies.” *Neurotoxicol Teratol* 52:93–108
- Vorhees CV, Williams MT (2020) Litter effects: comments on Golub and Sobin’s “Statistical modeling of litter as a random effect in mixed models to manage ‘intralitter likeness.’” *Neurotoxicol Teratol* 77:106852
- Vorhees CV, Williams MT (2024) Tests for learning and memory in rodent regulatory studies. *Curr Res Toxicol* 6:100151
- Weiner AMJ, Irijalba I, Gallego MP, Ibarburu I, Sainz L, Goñi-De-Cerio F, Quevedo C, Muriana A (2024) Validation of a zebrafish developmental defects assay as a qualified alternative test for its regulatory use following the ICH S5(R3) guideline. *Reprod Toxicol* 123:108513
- Weinstock M (2008) The long-term behavioural consequences of prenatal stress. *Neurosci Biobehav Rev* 32:1073–1086
- Yang D, Kim KH, Phimister A, Bachstetter AD, Ward TR, Stackman RW, Mervis RF, Wisniewski AB, Klein SL, Kodavanti PR, Anderson KA, Wayman G, Pessah IN, Lein PJ (2009) Developmental exposure to polychlorinated biphenyls interferes with experience-dependent dendritic plasticity and ryanodine receptor expression in weanling rats. *Environ Health Perspect* 117:426–435
- Zaichenko MI, Sidorina VV, Grigoryan GA (2025) Effects of an enriched environment on learning and memory in the Morris water maze in rats with acute and chronic proinflammatory stress. *Neurosci Behav Physiol* 55:451–458

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Pragati S. Coder¹ · Richard Lewis² · Anika Christoffel³ · Phil Botham⁴ · Richard Bartlett⁴ · Megan D'souza⁵ · Colin Ehnes⁶ · Frank Faulhammer⁷ · Susann Fayyaz⁸ · Dilshan S. Harischandra⁹ · Parvinder Kaur¹⁰ · Susanne N. Kolle¹¹ · Abby Li¹² · Jason Manton¹³ · Benoît Nihoul¹⁴ · Heidrun Potschka¹⁵ · Emily Richmond¹⁶ · Martijn Rooseboom¹⁷ · Eddie Slotter¹⁸ · Lorna F. Tempest¹⁹ · Abigail Tolpin²⁰ · Corina van den Heuvel²¹ · Matthew Wolter²²

✉ Megan D'souza
megan.dsouza@ecetoc.org; info@ecetoc.com

¹ Charles River Laboratories, Ashland, OH, USA

² Regulatory Science Associates Largs Yacht Haven, Irvine Road, Largs, Scotland KA30 8EZ, UK

³ BASF SE, Ludwigshafen, Germany

⁴ Syngenta, Jealott's Hill, Bracknell RG42 6EY, UK

⁵ ECETOC, Rue Belliard 40, Brussels 1040, Belgium

⁶ BASF SE, APD/ET, Speyerer Strasse 2, Limburgerhof 67117, Germany

⁷ BASF SE, Global Toxicology & Ecotoxicology, Ludwigshafen, Germany

⁸ Clariant Produkte (Deutschland) GmbH, Industriepark Höchst, Building G 860, Frankfurt am Main, Germany

⁹ Labcorp, Greenfield, IN, USA

¹⁰ RIFCON GmbH, Goldbeckstraße 13, 69,493, Hirschberg, Germany

¹¹ BASF, DE SE, Ludwigshafen, Germany

¹² Exponent Health Sciences, Menlo Park, California, USA

¹³ Toxiqua Ltd, Worcester, UK

¹⁴ Dow Silicone Belgium, Rue Jules Bordet, 7180 Seneffe, Belgium

¹⁵ Institute of Pharmacology, Toxicology, and Pharmacy, Ludwig-Maximilians-Universität, Munich, Germany

¹⁶ Exponent International, Harrogate, UK

¹⁷ Shell Global Solutions International B.V., The Hague, The Netherlands

¹⁸ The Lubrizol Corporation, Wickliffe, OH, US

¹⁹ Huntsman Surface Sciences UK, Cambridge, UK

²⁰ Corteva Agriscience, Indianapolis, IN, USA

²¹ Charles River Laboratories, 's-Hertogenbosch, The Netherlands

²² S.C. Johnson & Son, Inc, Racine, WI, USA