Deriving acceptable concentrations for endocrine disrupting chemicals using aquatic toxicity data

Chief Scientist's Group report

January 2026

We are the Environment Agency. We protect and improve the environment.

We help people and wildlife adapt to climate change and reduce its impacts, including flooding, drought, sea level rise and coastal erosion.

We improve the quality of our water, land and air by tackling pollution. We work with businesses to help them comply with environmental regulations. A healthy and diverse environment enhances people's lives and contributes to economic growth.

We can’t do this alone. We work as part of the Defra group (Department for Environment, Food & Rural Affairs), with the rest of government, local councils, businesses, civil society groups and local communities to create a better place for people and wildlife.

Published by:

Environment Agency
Horizon House, Deanery Road,
Bristol BS1 5AH

www.gov.uk/environment-agency

Further copies of this report are available from our publications catalogue: www.gov.uk/government/publications or our National Customer Contact Centre: 03708 506 506

Email: research@environment-agency.gov.uk

Author(s):
Rebecca J Brown
Dean Leverett
Grace H Panter
Kate Roylance
Olivia Tran

Keywords:
nonylphenol, octylphenol, 17α-ethinylestradiol, EE2, endocrine, predicted no effect concentrations, species sensitivity distributions

Research contractor:
wca Environment Limited
Brunel House, Volunteer Way
Faringdon SN7 7YR
01367 246026

Environment Agency’s Project Managers:
Ian Doyle and Morné Van Der Mescht

Citation:
Environment Agency (2026) Deriving acceptable concentrations for endocrine disrupting chemicals using aquatic toxicity data. Environment Agency, Bristol.

Research at the Environment Agency

Scientific research and analysis underpins everything the Environment Agency does. It helps us to understand and manage the environment effectively. Our own experts work with leading scientific organisations, universities and other parts of the Defra group to bring the best knowledge to bear on the environmental problems that we face now and in the future. Our scientific work is published as summaries and reports, freely available to all.

This report is the result of research commissioned by the Environment Agency's Chief Scientist's Group.

You can find out more about our current science programmes at https://www.gov.uk/government/organisations/environment-agency/about/research

If you have any comments or questions about this report or the Environment Agency's other scientific work, please contact research@environment-agency.gov.uk.

Dr Robert Bradburne
Chief Scientist

Contents

Executive summary

Several Substances of Very High Concern (SVHC) are listed in Annex 14 of the UK Registration, Evaluation, Authorisation and restriction of CHemicals (UK REACH) regulation based on their endocrine disruption (ED) potential in aquatic wildlife. Their continuing use is subject to authorisation. Authorisation can be granted in two ways: either where "adequate control" has been demonstrated (i.e. exposure concentrations are below an acceptable threshold), or due to socioeconomic need. To date, authorisations for environmental endocrine disrupting chemicals (EDCs) have only been sought on socioeconomic grounds. No guidance is available to support applicants who might want to make "adequate control" arguments. Furthermore, decision makers have asked the regulator to provide an opinion on potential risks for applications made under the socioeconomic route.

A SETAC Pellston Workshop® on Ecotoxicological Hazard and Risk Assessment Approaches for Endocrine-Active Substances held in 2016 concluded that a predicted no effect concentration (PNEC) can be derived for EDCs provided that "sufficient" data are available (Matthiessen *et al*., 2017). This implies that a risk assessment can be performed in principle.

This report explores how PNECs for environmental EDCs might be derived using the type of ecotoxicity data that are available to identify them as SVHCs in the first place. It reviews three data-rich EDCs to establish what PNECs might be considered sufficiently protective of ED effects in the aquatic environment: the pharmaceutical ethinyloestradiol (EE2), and the Annex 14 listed substances nonylphenol (NP) and octylphenol (OP). Ecotoxicity datasets were compiled using reliable studies at Levels 4 and 5 of the Organisation for Economic Co-operation and Development's (OECD) conceptual framework (CF) for EDCs. The suitability of each dataset was compared to the SETAC Pellston Workshop® criteria and mechanistic results were reviewed to provide supporting context.

The initial review suggested that suitable PNECs could be derived by applying assessment factors (AFs) to the lowest no-observed effect concentrations (NOECs) from relevant ecotoxicity studies. The size of the AF reflects the amount of uncertainty in the dataset, as indicated by its level within the OECD CF. To test the concept, the initial findings from the three case studies were applied to eight further EDCs (industrial chemicals and pesticides) that have smaller datasets. A comparison was also made with PNECs derived following standard approaches for chemicals with non-specific mechanisms of action (MoA), together with an additional default AF of 5 to account for ED effects. The ED data considered for these test cases came from screening studies such as the fish short term reproduction assay (FSTRA), the more sophisticated fish sexual development test (FSDT), and the highest tier fish lifecycle toxicity test (FLCTT). These included mechanistic data, so it was possible to consider whether the derived PNECs offered a degree of protection against biological changes that may not in themselves be considered adverse but are a signal of endocrine system perturbation.

The report shows that it is possible to derive threshold-based values for an EDC based on NOECs from an OECD CF Level 4 or 5 test for oestrogen, androgen, and steroidogenesis (EAS) modalities. A Level 5 test assesses the whole lifecycle (providing greater coverage compared to a Level 4 test), includes relevant mechanistic data and is generally the most sensitive study. Therefore, a higher AF for Level 4 studies reflects the uncertainty that arises from testing only part of the lifecycle. Whilst a Level 3 FSTRA and a Level 4 FELS test are not usually acceptable for the identification of an EDC as a SVHC, an AF of 1 000 and 100, respectively, appears to be appropriate for the derivation of an indicative PNEC covering potential ED impacts for risk assessment purposes. The general approach is summarised below.

Recommended assessment factors for data from different levels of the OECD CF

Available data	AF
Lowest NOEC/EC_{10}[a] from a Level 3 FSTRA study	1 000
One NOEC/EC_{10} from a Level 4 FELS test or an FSDT	100
One NOEC/EC_{10} from a Level 5 FLCTT covering the sensitive window for sexual differentiation and reproduction	10

Note: A small additional AF might also be considered if the PNEC is not protective of mechanistic effects, etc.
[a] 10% Effect Concentration

The approach is an improvement compared with the application of an additional default AF to account for the ED properties, because the AF reflects the fish toxicity data available to provide a more relevant threshold value.

Some uncertainties remain for multigenerational effects on fish species with alternative reproductive strategies (e.g. live bearers), species that are long lived or marine, and for amphibians. Whilst it is assumed that these species are adequately protected based on the proposed AFs, further evidence from such species would be helpful to provide reassurance on this point. Insufficient data were available for thyroid-related effects, which would need to be considered in future.

This report could provide the basis for guidance to applicants for authorisation and allow the regulator to judge when the aquatic toxicity database is sufficient for PNEC derivation in other circumstances (such as restriction and evaluation decisions).

Acknowledgements

The Environment Agency gratefully acknowledges the peer review of the report provided by members of the Hazardous Substances Advisory Committee (HSAC), and HSE's UK REACH Independent Scientific Expert Panel (RISEP).

1. Introduction

1.1 Endocrine disruption and chemical regulation

Hormones are signalling molecules produced by glands in multicellular organisms that regulate many essential biological processes (collectively known as the “endocrine system”). Examples include the sex hormones oestradiol and testosterone and the growth regulating hormone thyroxine. Endocrine-disrupting chemicals (EDCs) mimic, block or interfere with the production, metabolism or action of hormones. They can therefore alter the normal functioning of the endocrine systems of biota, leading to a range of adverse effects, for example on behaviour, growth and/or reproductive fitness (Gore *et al.,* 2014). Concern about their impact has been growing since the 1970s, when researchers first began to notice changes in the reproductive health of several wildlife groups (Jun *et al.,* 1972; Schug *et al.,* 2016; Sifakis *et al.,* 2017; Godfray *et al.,* 2019). Classic examples in a UK context are the widespread feminisation of male fish in rivers caused by oestrogenic substances and reductions in reproductive output and survival of marine mollusc species associated with tributyl tin exposure (WHO/IPCS, 2002; EA, 2002; 2004; Godfray *et al.,* 2019).

Endocrine disruption is not a toxicological endpoint in itself, but rather a mode of action that may lead to adverse effects. The World Health Organisation’s (WHO) EDC definition (WHO/IPCS, 2002) is:

- An endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub) populations; and
- A potential endocrine disruptor is an exogenous substance or mixture that possesses properties that might be expected to lead to endocrine disruption in an intact organism, or its progeny, or (sub) populations.

Various international bodies have identified EDCs from a range of chemical groups. For example, polybrominated diphenyl ethers, phthalate esters, 4-alkylphenols, bisphenols, polychlorinated biphenyls and dioxins (Diamanti-Kandarakis *et al.,* 2009). Naturally occurring substances may also be endocrine active – for example, plant isoflavones can have hormone-like activity in mammals and have been used in human food supplements for this purpose (NIEHS, 2024). Many pharmaceuticals are also designed to interact with the endocrine system, the most obvious example being the contraceptive pill.

The WHO definition has been adapted for regulatory purposes, most recently as part of legislation for classification, labelling and packaging (CLP) of substances and mixtures in the EU (EC, 2023), which states that a substance shall be identified as having endocrine disrupting properties for the environment if it meets the following criteria:

1. It causes an adverse effect in an intact organism or its offspring or future generations relevant at the population level;

2. It has endocrine activity; and

3. There is a biologically plausible link between the endocrine activity and the adverse effects.

A Weight-of-Evidence (WoE) assessment is usually used to demonstrate the biologically plausible link. This hazard classification category is under discussion at the United Nations and has not yet been adopted within Great Britain.

There are no formal information requirements for ED properties under UK REACH, but they can be considered by regulators under Substance Evaluation with reference to OECD Guidance Document (GD) 150 (OECD, 2018). This provides a Conceptual Framework (CF) to assess different types of information. Levels 1 and 2 concern non-test and *in vitro* data, Level 3 concerns mechanistic information from *in vivo* studies, and Level 4 and 5 concern more complex *in vivo* studies which include apical endpoints to allow adversity to be assessed.

In general, GD 150 focuses on the following endocrine modalities:

- oestrogen mediated (E)
- androgen mediated (A)
- thyroid hormone mediated (T)
- steroidogenesis interference (S)

An EDC may be identified as a Substance of Very High Concern (SVHC) under Article 57f of the Registration, Evaluation, restriction and Authorisation of CHemicals (REACH) Regulations. This is triggered by the regulator, who must provide an argument that the chemical poses an "equivalent level of concern" to substances that meet the criteria listed in Article 57a-e of the legislation. For environmental EDCs, this usually relies on *in vivo* tests performed using fish when investigating oestrogen, androgen and steroidogenic (EAS) modalities. Amphibian testing is currently used to identify thyroid disrupting chemicals in the environment, although methods that include thyroid relevant measurements as part of fish studies are under development.

1.2 Authorisation of environmental EDCs

SVHC identification triggers various obligations for suppliers of the EDC and the SVHC may also be added to Annex 14 of REACH (the "authorisation list"). Authorisation obliges companies that wish to carry on using an SVHC beyond a specified date to submit a package of information to justify why they should be allowed to do so.

Article 60 of REACH states that authorisation of an SVHC can be granted in two ways: either demonstration that exposure concentrations are below an acceptable threshold ("adequate control") or due to socioeconomic need. The European Union (EU) is ambiguous about which route is appropriate for chemicals with an ED MoA. Applicants are advised to assume there is no safe level of exposure unless they can justify otherwise, but there is no guidance about how this justification can be made (EC, 2016). To date, authorisations under both EU and UK REACH for ED SVHCs have only been granted on socioeconomic grounds.

The purpose of this report is therefore to explore whether and how a case could be made to use the adequate control route for SVHCs that are environmental EDCs. This will also be relevant to other REACH activities such as restriction and Substance Evaluation.

1.3 Report outline

Data for three relatively data-rich substances with a common oestrogenic ED MoA were compiled for an initial assessment:

- 4-Nonylphenol (NP, also known as para-NP) and 4-tert-octylphenol (OP) are two well-studied industrial substances with ER agonist activity. They were selected for this project because they are transformation products of derivatives called alkylphenol ethoxylates, which are subject to authorisation under REACH.
- 17α-Ethinyloestradiol (EE2) is a synthetic hormone used in various medical treatments due to its potent ER agonist activity. It therefore provides a benchmark for assessing the properties of other ER agonists.

Following data compilation, the quality of the datasets was assessed against the Society of Environmental Toxicology and Chemistry (SETAC) Pellston Workshop® criteria (Matthiessen *et al*., 2017). Next, there was an exploration of how acceptable threshold levels might be derived for these substances following standard approaches for chemicals with non-specific MoAs to derive Predicted No Effect Concentrations (PNECs), but using existing data relevant for the detection of endocrine effects. Finally, based on the conclusions for NP, OP and EE2 it was considered how the approach might be used for EDCs with more limited datasets.

Section 14 of this report outlines the methodology and approaches taken for deriving the PNECs, and for assessing what a “sufficient dataset” means for an EDC. Sections 24, 41 and 56 detail the derivation of proposed PNECs for NP, OP and EE2, respectively. Section 77 discusses these PNECs and Section 82 considers the derivation of PNECs for substances with ED concerns but limited datasets, by considering the eight additional chemicals as examples. Section 104 discusses what a sufficient dataset is for the purposes of risk assessment of EDCs, while Section 117 provides conclusions and recommendations.

Note:

1. Applicants for authorisation are only required to address impacts arising from the emission of the substance listed on Annex 14. Consideration of potential mixture effects is not legally required, so is not discussed in this report (see EA, 2022; RSC, 2025 for a wider discussion of this issue).

2. This report is about the control of substances that have already been identified as environmental EDCs; it does not discuss the way they are identified, or associated test strategies.

2. Methodology

2.1 Data collation for NP, OP and EE2

2.1.1 Literature collation methodology

Existing regulatory assessments or reviews were used as the starting point for data collection for NP, OP and EE2. As the purpose of the existing data summaries differed between the substances (i.e. regulatory PNEC derivation for NP versus an SVHC assessment for OP) there were differences in the criteria for acceptability of the data cited in the different reports. For example, studies without supporting analytical verification were not considered acceptable for use in deriving a regulatory PNEC for NP (ECHA, 2014a). New literature searches, detailed below, were performed to provide contemporary data published since the regulatory assessments were completed. It was assumed that earlier published literature would have been adequately captured by the respective regulatory assessments.

2.1.1.1 Existing regulatory assessments for NP and OP

The following sources were used to initially identify effects data for NP and OP:

- REACH restriction dossier for nonylphenol ethoxylates (NPEO) in textiles and the accompanying ECHA opinion (ECHA, 2014a; b);
- SVHC dossiers for NP and OP (ECHA, 2011; ECHA, 2012a, b; ECHA, 2013a);
- Environment Agency risk evaluation report for OP (EA, 2005);
- REACH Substance Evaluation conclusion for nonylphenol, branched, ethoxylated (ECHA, 2018), which includes relevant data for NP;
- Relevant Robust Study Summaries published in the ECHA dissemination portal for the registrations of NP and OP; and
- Additional studies cited in the octylphenol ethoxylate (OPEO) REACH registration Chemical Safety Report (data provided by the EA).

2.1.1.2 Literature searches for NP and OP

Literature searches were conducted to identify any new data for NP published between 1st January 2014 and 31st August 2022 and data for OP published between 1st January 2012 and 31st August 2022. The start dates were prior to the publication of the most recent regulatory assessments, to provide some overlap between the literature searches conducted for the previous assessments (although the search methodologies for those assessments were not available so could not be checked for consistency). Searches of published scientific literature were conducted on 31st August 2022 using bibliographic databases PubMed (United States National Library of Medicine) and Web of Science (Clarivate Analytics). The following search strings were used:

Nonylphenol: *("Nonylphenol" OR "Nonyl phenol" OR "2,6-Dimethyl-4-heptylphenol") AND (development OR reproduction OR growth OR chronic OR subchronic OR NOEC OR algae OR fish OR amphibia* OR daphnia OR invertebrate OR crustacea* OR vitellogen* OR imposex OR intersex OR metamorphosis OR ovotest* OR testis-ova OR gonado-somatic)*

Octylphenol: *("Octylphenol" OR "Octyl phenol") AND (development OR reproduction OR growth OR chronic OR subchronic OR NOEC OR algae OR fish OR amphibia* OR daphnia OR invertebrate OR crustacea* OR vitellogen* OR imposex OR intersex OR metamorphosis OR ovotest* OR testis-ova OR gonado-somatic)*

Note: NOEC = no observed effect concentration

The searches retrieved 1 021 unique results for NP and 413 unique results for OP. Titles and abstracts were screened for aquatic ecotoxicological studies with information on ED or adverse effects on apical endpoints following the relevance and reliability assessment approaches detailed in Section 15. Full papers were obtained if information in the title and abstract were insufficient to determine the relevance or reliability.

2.1.1.3 Existing data for EE2

The following sources were used to identify data for potential use in deriving a PNEC for EE2:

- Matthiessen *et al*. (2017): Case Study Summary for EE2 prepared for discussion at a SETAC Pellston Workshop®;
- European Commission report on the evaluation of 12 substances in the context of the ED Priority List of Actions (EC, 2002);
- German Environment Agency Department of Pharmaceuticals Environmental Quality Standard (EQS) report for EE2 (JRC, 2022); and
- Literature search conducted and provided by the Environment Agency covering the period 2016 – 3 June 2021 (this was conducted using the search terms specified in the supplementary information of Matthiessen *et al*. (2017), using Scopus).

2.1.2 Data collection and assessment

2.1.2.1 Relevance assessment

The relevance criteria applied to data for consideration in the PNEC derivation were based on those used in the REACH restriction dossier for NPEO in textiles (and accompanying ECHA opinion) (ECHA, 2014a) as follows:

1. Guidance for standard OECD ecotoxicology tests were used to assess the suitability of the study. Such studies would usually test at least three concentrations with spacing factors between test concentrations of at least 3.2. Non-standard tests

were also included where population-relevant endpoints (survival, growth, development and/or reproduction), were included.

2. Since PNECs for an EDC should be based on chronic rather than acute endpoints, only data from long-term studies were used, typically at Levels 4 and 5 of the OECD CF for EDCs (OECD, 2018) (Appendix 151). Suitable studies for deriving a long-term PNEC are also listed in the EC Technical Guidance for deriving Environmental Quality Standards (EC, 2018).

   It should be noted that:

   - OECD CF Level 3 studies for EAS activity are generally short-term adult exposures intended to provide mechanistic information. They are not sufficiently sensitive for use in PNEC derivation, so are used as supporting information only, unless otherwise justified (i.e. testing of more than three concentrations).

   - Fish growth studies (e.g. based on the fish juvenile growth test, OECD Test Guideline (TG) 215 (OECD, 2000)) are not included in the OECD CF because they are not expected to be sensitive for endocrine active substances as they do not cover sensitive life stages or include reproductive endpoints. These studies have therefore not been used for the PNEC derivation.
   - The Fish Early Life Stage (FELS) test (OECD TG 210 (OECD, 2013)) and Fish Sexual Development Test (FSDT, OECD TG 234, OECD, 2011) are both Level 4 studies and therefore considered suitable for PNEC derivation. There is, however, expected to be a difference in sensitivity between these studies (especially for detecting effects of endocrine active substances) because the FSDT has a longer species-specific test duration (e.g. for Zebrafish *Danio rerio*, 60 days post hatch in the FSDT compared with 30 days post hatch in the FELS test), higher replication (e.g. 120 fish per treatment in the FSDT compared with 80 in the FELS test) and incorporates EAS-mediated endpoints such as sex ratio. The duration of a FELS test with the Rainbow Trout (*Oncorhynchus mykiss*) is 60 days post hatch, which is an equivalent duration to the FSDT conducted with the smaller fish species. Both studies have been included in the present assessment, but study design has been considered as part of the data sufficiency assessment.

3. Exposure should be via water as this is most relevant for the pelagic compartment.

4. Results from studies based on nominal concentrations have been used as supporting information only.

5. The study should derive a bounded NOEC (i.e. not “less than” or “greater than” values). If a NOEC cannot be derived and the lowest observed effect concentration (LOEC) represents between a 10% and 20% effect compared with the control, then

a NOEC calculated from the LOEC/2 can be used (ECHA, 2008). If a NOEC cannot be defined, and no EC_{10} (test concentration for a 10% effect) was reported, the study is considered not relevant for PNEC derivation purposes.

6. Algae and plants were excluded as they do not possess functional oestrogen or androgen hormones.

7. SVHCs identified on the basis of environmental ED concerns usually rely on evidence of endocrine-mediated toxicity in aquatic vertebrates, as they are sensitive to EAS modalities. Nevertheless, invertebrate data were initially included to maximise the NP and OP datasets even though they are not expected to have functional steroid receptors relevant for the oestrogenic MoA (Crane *et al*., 2022). Invertebrates have, however, shown relatively high sensitivity following chronic exposure to alkylphenols (e.g. EA, 2005; ECHA, 2014a). Furthermore, the inclusion of invertebrates increases the number of species for a probabilistic PNEC, arguably increasing its statistical robustness (Posthuma *et al*., 2019). Invertebrates were not included in the EE2 assessment because a large amount of fish data is available.

Studies may also include endpoints which provide mechanistic information about an endocrine MoA, such as effects on secondary sex characteristics (SSC), gonad histopathology, gonadosomatic index (GSI) and vitellogenin (VTG) measurements. These provide evidence for biological changes at the level of the individual, which are expected to occur at lower test concentrations compared with population-level outcomes. They were therefore used as supporting contextual information. For further information on AOPs and MoAs see Appendix K.

2.1.2.2 Data collation

All relevant data were collated in a spreadsheet detailing the following:

- Substance tested
- Species tested and taxonomic group
- Media of relevance (freshwater or marine)
- Test guideline used (where applicable)
- Study type and OECD CF level
- Test endpoint and whether it was population-relevant
- Value of the NOEC or EC_{10}
- Reliability of the study (i.e. Klimisch score, Klimisch *et al.,* 1997)
- A study ID number with associated study reference
- The source of the study (i.e. from a previous regulatory assessment or new literature search).

Effects in the relevant studies were also summarised in tabular format. These tables include population-relevant and mechanistic endpoints.

2.1.2.3 Data reliability assessment

The assessment of reliability for all studies was similar to that used in ECHA (2014a) with standardised test designs being used as a reference for test methodology. Non-standard tests were included assuming that test organisms were maintained under acceptable conditions, the methods and results were adequately reported, and there was supporting analytical verification of test concentrations.

The reliability of new studies identified in the scientific literature for this report were assessed using the Criteria for Reporting and Evaluating ecotoxicity Data (CRED) method developed by Moermond *et al*. (2016). CRED provides a framework for assessing reliability of a study which is aligned with the Klimisch scoring system (Klimisch *et al*., 1997). Only studies assessed as reliable (R1) or reliable with restrictions (R2) were used for PNEC derivation. CRED uses “C1” etc, but for consistency in this report, “R1” etc is used to cover both CRED and Klimisch scoring.

The reliability of data included in existing regulatory reports was assessed as follows:

- The data used for PNEC derivation for NP in the REACH restriction dossier (ECHA, 2014a) and for OP by the Environment Agency (EA, 2005) were used without re-evaluation for reliability. Klimisch scores were used as reported.
- The ECHA (2011) SVHC support document evaluating the endocrine properties of OP included some studies that were potentially suitable for PNEC derivation. The reliability of studies reported in ECHA (2011) but not included in EA (2005) were therefore assessed for reliability as part of the current assessment.
- The data used for the PNEC derivation for EE2 primarily came from the studies considered in Matthiessen *et al*. (2017) and assessments published by the European Commission (EC, 2002) and Joint Research Centre (JRC, 2022). A re-evaluation of their reliability was undertaken to ensure consistency. This is because the approach adopted for this assessment deviated from Matthiessen *et al*. (2017) and JRC (2022), in which the LOEC/2 was used for some studies where the LOEC was > 20% of the NOEC.

Some studies cited in the existing reviews were taken from secondary sources. For example, these included proprietary studies cited from early regulatory assessments or the International Uniform ChemicaL Information Database (IUCLID), which are no longer accessible. However, results from studies which were generated for regulatory purposes, and which have been assessed as reliable in a previous regulatory assessment (for example an ECHA SVHC Support Document), were included in this assessment to maximise the dataset.

2.2 PNEC derivation

2.2.1 PNEC derivation methodology

Both deterministic and probabilistic approaches were taken for PNEC derivation for NP, OP and EE2 following the appropriate technical guidance (ECHA, 2008; EC, 2018).

The standard deterministic approach extrapolates the lowest population-relevant endpoint obtained in (eco)toxicity tests (such as the concentration that does not cause any statistically significant effects on survival, growth, development and reproduction compared to a control group) using an Assessment Factor (AF) to yield a PNEC. The AF is a way of taking account uncertainties related to:

- intra- and inter-laboratory variation of toxicity data;
- intra- and inter-species variations (biological variance);
- short-term to long-term toxicity extrapolation;
- laboratory data to field impact extrapolation.

For freshwaters, the default AF can be 1 000, 100, 50 or 10, depending on the amount of (eco)toxicity test data available (the highest value being used when there are only acute (short-term) toxicity data, and the lowest value when there are chronic (long-term) toxicity data for fish, invertebrates and algae/aquatic plants). It is assumed for regulatory purposes that chemical concentrations below the PNEC will not lead to significant harm to wildlife populations. This is a well-established approach for assessments under REACH and preceding legislation.

Probabilistic PNEC derivation uses the same long-term (eco)toxicity data but applies statistical modelling to derive a concentration – called the HC_5 – that is assumed to be protective of 95% of all species in the relevant environment (e.g. freshwater). An AF is applied to the HC_5 to derive the PNEC; the AF reflects the higher level of certainty in the dataset, so is typically lower than those used for deterministic PNEC derivation.

At least 10 (preferably 15) reliable and relevant NOEC / EC_{10} values from different species covering at least eight representative taxonomic groups are generally required. However, since the PNECs considered in this report focus specifically on the results from ED relevant studies at Level 4 and 5 of the OECD CF with invertebrates and vertebrates (OECD, 2018), such an extensive taxonomic coverage is unlikely to be met in most cases for EDCs. Therefore, for this project, a Species Sensitivity Distribution (SSD) was generated based on all the available reliable and relevant data (i.e. from studies from Level 4 or 5 of the CF) with no regard for these taxonomic coverage criteria.

SSDs were initially generated using the ETX version 2.3 computer program for calculating the HC_5 (log-normal model) (Van Vlaardingen *et al*., 2004). However, recent developments in SSD modelling suggests that alternative or consensus (average) models may provide a better fit and more reliable estimate of the HC_5 especially for small datasets (Fox *et al*., 2021). The HC_5 values were therefore derived using alternative and averaged models in ssdtools version 1.0.2 (Thorley and Schwarz, 2018) in a web-based app (Dalgarno, 2018). The minimum dataset for these models is six species. The most appropriate models for inclusion in the assessment were chosen based on goodness-of-fit parameters, using the Anderson-Darling statistic and Akaike's Information Criterion corrected for sample size (AICc), as well as visual inspection of the data. The AICc method is recommended for assessing goodness-of-fit by ssdtools (the lowest value indicating the best fit) based on Burnham and Anderson (2002). The Anderson-Darling statistic (the highest value indicating the best fit) is recommended for the same purpose in the EQS technical guidance (EC, 2018).

The results for all median HC_5 models (and the average model) were reported for the models with a good fit along with lower and upper confidence intervals. The appropriate value for use in PNEC derivation is discussed for each substance.

2.2.2 SETAC Pellston Workshop® criteria

The datasets for PNEC derivation were evaluated against the SETAC Pellston Workshop® criteria for EDCs (Matthiessen *et al*., 2017) for the purposes of data gap analysis. The SETAC criteria are outlined below:

- Have the most appropriate taxa been tested (with relevant endpoints)?
- Have sensitive life stages, or the entire life cycle, been tested (again with relevant endpoints)?
- Have delayed and multigenerational (i.e. latent) effects been considered?
- Do Non-Monotonic Dose Relationships (NMDRs) or other unusual temporal patterns of toxicity affect the ability to predict reliable no-adverse-effect levels (NOAELs, the equivalent term to NOECs in mammalian toxicity studies)?
- Does a threshold for adverse endocrine-mediated effects exist?

Standard default AFs were initially applied to the datasets for NP, OP and EE2 but the AFs were evaluated and amended based on the data quality, the assessment against the criteria outlined above and whether the resulting PNEC was protective for a NOEC based on mechanistic effects, as discussed in Sections 24, 41 and 56.

2.3 Substances with smaller datasets

2.3.1 Selection criteria

To test the initial proposals for deriving PNECs for EDCs based on evidence for NP, OP and EE2, further substances with an EAS MoA were selected to evaluate the applicability of AFs in cases where there were fewer data. The selection criteria were substances with:

1. a well-documented and understood endocrine mechanism, and
2. clear environmental effect data at concentrations below the limit of solubility, and which are not excessively impacted by systemic toxicity.

Three ER agonist substances were chosen along with five substances with different types of endocrine activity, but still predominantly involving EAS pathways. These are listed in Table 1. Data extracted from existing compilations for each substance were used at face value; it was beyond the scope of this project to perform an in-depth assessment of data reliability for them.

Table 1 Initial list of substances used for assessing the sufficiency of an ED dataset

Substance	CAS number	Endocrine mechanism	Data sources
3-Benzylidene camphor (3-BC)	15087-24-8	ER agonist	ECHA, 2016
Methyl paraben	99-76-3	ER agonist	IUCLID, 2022b
Propyl paraben	94-13-3	ER agonist	IUCLID, 2021b
Fenbuconazole	114369-43-6	Thyroid activity, AR antagonist and steroidogenesis	EFSA, 2010a; IUCLID, 2021c
Cyproconazole	94361-06-5, 94361-07-6, 113096-99-4	Steroidogenesis	EFSA, 2010b; INERIS, 2011
Propiconazole	60207-90-1	AR antagonist	EFSA, 2017; Matthiessen *et al*., 2017; JRC, 2020
Prochloraz	67747-09-5	Aromatase inhibitor	OECD, 2006; EFSA, 2011; JRC, 2020
Vinclozolin	50471-44-8	AR antagonist	Matthiessen *et al*., 2017 and various scientific publications

Three substances initially selected as validation substances did not meet the criteria as follows:

- **Bis(2-ethylhexyl) phthalate (DEHP), CAS no. 117-81-7:** DEHP has been identified as an SVHC under REACH based on its ED properties, acting mainly as an ER agonist (ECHA, 2014c). Ecotoxicity testing is difficult because it has a high adsorption potential to surfaces and organic matter (the n-octanol/water partition coefficient (log K_{OW}) is 7.5) and low water solubility (3 µg/L at 20°C) (ECHA, 2014c). The SVHC dossier discusses difficulties in interpreting the available ecotoxicology dataset for DEHP, because studies were either conducted at test concentrations above the water solubility limit or do not have supporting analytical chemistry. Since it is beyond the scope of this project to review the individual studies for DEHP, and because it is difficult to determine an exact NOEC / EC_{10} value for the ED effects for PNEC derivation due to the limitations of the available studies, it is not considered a suitable validation substance.

- **Tributyl tin (TBT) compounds, CAS no. 688-73-3 / 36643-28-4**: This is a class of organotin compounds which contain a $(C_4H_9)_3Sn$ group. They are well established as having an EAS mode of action, but the endocrine mechanism for invertebrates appears to be via retinoid X receptor (RXR) and peroxisome proliferator-activated receptor (PPAR) pathways (Matthiessen *et al*., 2017 – see supplemental data S4). These endocrine mechanisms are not well understood compared with the EAS pathways which are covered by OECD GD 150 and European guidance on how to identify an EDC (EFSA/ECHA, 2018). It was therefore decided to remove this substance from scope.

- **Etridiazole, CAS no. 2593-15-9**: The main ED concerns for this substance are for effects via the thyroid, mainly based on mammalian toxicity data (ECHA, 2013b; EA, 2023). It was therefore considered not to be a relevant for an approach focused on EAS modalities because amphibian data are considered necessary for evaluating the hazards of thyroid active substances.

2.3.2 Endocrine activity of the selected substances using ToxCast

To assess whether the potency based on endocrine activity could be used as supporting information when deriving a PNEC, data on the EAS activity of all the substances was derived from the United States Environmental Protection Agency's (US EPA) CompTox Dashboard. The dashboard was checked on 24 January 2023. The CompTox Dashboard includes ToxCast data from the US EPA Endocrine Disruptor Screening Programme. These are *in vitro* assays based on mammalian test systems. The following information on relevant for EAS modalities was extracted:

- Bioactivity model scores. The bioactivity models use the results from multiple receptor-mediated *in vitro* assays (based on mammalian cell lines) to make an overall prediction for ER (Browne *et al*., 2015) or AR (Kleinstreuer *et al*., 2017) agonist or antagonist activity. The models generate area under the curve (AUC) scores typically between 0 and 1.0. Area under the curve scores below 0.001 indicate a substance to be inactive, AUC scores above 0.1 indicate activity. Scores between these values are inconclusive.

- Steroidogenesis based on the result (AC_{50}) for aromatase inhibition (TOX21_Aromatase_Inhibition) as well as cell viability (TOX21_Aromatase_Inhibition_viability) in the associated assay.

- Steroidogenesis effects based on the result for hormone responses in H295R cell line (up or down). The hormones typically measured are cortisol, 11-deoxycortisol, deoxycorticosterone, corticosterone, androstenedione, 17α-hydroxypregnenolone, progesterone, 17α-hydroxyprogesterone, oestradiol, oestrone, and testosterone.

Where an AC_{50} for cytotoxicity was less than or within 10 μM of the result from the endocrine specific assay, the endocrine assay was considered to have been affected by cytotoxicity and therefore was assessed as not valid. The data are provided in Appendix 218. Section 79 discusses the ToxCast data for NP, OP and EE2, while Section 82 discusses the data for the individual validation substances.

2.3.3 Other evidence of *in vitro* endocrine activity

Increasing evidence that the endocrine system including androgen and oestrogen receptors are conserved between mammalian and fish species (Brown *et al.*, 2023; Panter *et al*., 2023). Therefore *in vitro* mechanistic data can provide useful screening level information for endocrine activity in fish. For some substances *in vitro* data based on effects in fish specific receptors (including medaka ERα) have been generated as part of the Ministry of Environment of Japan’s program on endocrine disrupting effects of chemical substances. Data for *in vitro* medaka ERα activity of validation substances reported in Onishi *et al*., (2021) and Kawashima *et al*., (2021) were extracted as further evidence for endocrine activity.

3. PNEC derivation: Nonylphenol

3.1 Long-term aquatic toxicity data

A previous regulatory assessment for NP (ECHA, 2014a) was used as the basis for this assessment. This has been updated with new data which were published in the scientific literature since the 2014 assessment. Detailed information on the substance identity and physicochemical properties of NP is given in Appendix 146.

3.1.1 Fish

OECD CF Level 3 fish and amphibian studies (used as supporting information) are summarised in Table 43, and Level 4 and 5 studies are summarised in Table 44 and Table 45, respectively in Appendix 155. Only studies falling under Levels 4 or 5 were considered suitable for derivation of the PNEC.

Rice fish / *Adrianichthyidae* (*Oryzias latipes*)

One of the most sensitive species to NP is the Japanese medaka (*Oryzias latipes*). There are three FSTRA studies all with similar results (NOECs based on fertility or fecundity of 16.5 µg/L (Ishibashi *et al*., 2006), 18.8 µg/L (Kawashima *et al*., 2021), and 50.9 µg/L (Kang *et al*., 2003). From these studies VTG in male fish was the most sensitive biomarker endpoint, being significantly induced compared with the control at 5.4 – 50.9 µg/L (Ishibashi *et al*., 2006, Kawashima *et al*., 2021, Kang *et al*., 2003). There were also three Level 4 studies based on an FSDT. In one of these studies, growth was the most sensitive population-relevant endpoint with a NOEC of 11.6 µg/L based on reduced body weight (Seki *et al*., 2003), in the second sex reversal was affected at 31 µg/L (Horie *et al*., 2021), and in the third there were no effects on population-relevant endpoints at the highest tested concentration of 29 µg/L (Balch and Metcalfe, 2006). In all three FSDT studies testis-ova were observed at concentrations between 11.6 and 31 µg/L, SSC were affected between 3.0 and 44.7 µg/L and in one study (Seki *et al*., 2003) there was induction of VTG at 11.6 µg/L. The most sensitive FSDT study with *O. latipes* and nonylphenol was Seki *et al*., (2003). This was a non-standard test because the exposure from fertilised eggs until 60 days post hatch was then followed by a 30 day depuration phase. Since this depuration phase did not appear to affect the sensitivity of the study it is included for consideration for PNEC derivation (it was also assessed as reliable according to ECHA, 2014). As there were more sensitive Level 5 studies with *O. latipes* this study does not affect the derivation of the final PNEC.

There are also two Level 5 studies with *O. latipes*. The first is a life cycle test which exposed embryos until adult and then followed F1 fish until sexual differentiation (Yokota *et al*., 2001). This study was included in the ECHA (2014a) restriction dossier and assessed as reliable with restriction (R2). The NOEC from the study was 8.2 µg/L based on reduced survival post the swim-up phase in the F0 generation at 17.7 µg/L (Yokota *et al*., 2001). At 17.7 µg/L the sex ratio was female biased (9 males:19 females compared with 28:31 in the control and 26:28 in the solvent control), with only three reproductive pairs, whose fertility was reduced to 76% of that in the controls (although the effect on fertility was not statistically significant due to the small sample size). Mean fertility was > 90% in the controls and at test concentrations < 17.7 µg/L. There were significant effects on testis-ova, and SSC at 17.7 µg/L.

The experimental design of this study differs from the Medaka Extended One Generation Test (MEOGRT) which is initiated using mating pairs. Also, the reproductive phase in this study was based on six mating pairs in the control and in each treatment, whereas the MEOGRT uses twelve mating pairs in the control to increase the statistical robustness of the test. This lower resolution may explain the difference in sensitivity in this study compared with a subsequently conducted MEOGRT.

The second study is a MEOGRT which is the recommended Level 5 CF fish test according to the OECD. The study (Watanabe *et al*., 2017) is an early study relative to the publication of the OECD TG 240 (2015) and was considered to be reliable with restriction (R2). The study is also considered to be reliable in the ECHA registration dossier. There were however significant effects on growth and reproduction at the lowest test concentration of 1.27 µg/L, which means that the study is unbounded and difficult to interpret for use in PNEC derivation. A full summary and interpretation of the study is given in Appendix 179. The NOEC for reproduction is < 1.27 µg/L based on a concentration-dependent reduction in the number of eggs per female per day and the number of fertile eggs per female per day. Reduced growth at the lowest concentration of 1.27 µg/L was not considered to be biologically relevant.

No clear NOEC is derived for fertility and fecundity (unbounded study), however, according to ECHA (2008) if the LOEC is > 10 and < 20% effect, the NOEC can be calculated as LOEC/2. Since the effects on egg production are 12% reduction (based on eggs per female per day) and there is a 13% reduction (based on fertilised eggs per female per day) an estimated NOEC of 0.64 µg/L can be derived for this study. The lack of a clear NOEC for NP is a potential issue because of the ED MoA. However, given the lack of effects on reproductive endpoints at 8.2 µg/L in a second life cycle study with the same species (Yokota *et al*., 2001) and that in the MEOGRT key biomarkers for oestrogenic activity such as effects on male VTG induction in F0 and male SSC in F1 sub adults all followed a concentration response with clear NOEC values ≥ 2.95 µg/L the approach can be justified.

Carps and minnows / *Cyprinidae* (*Pimephales promelas*)

Several studies are available for the Fathead Minnow (*Pimephales promelas*). In a fish growth study, there was significantly reduced growth and survival of *P. promelas* with a NOEC of 38 µg/L and 77.5 µg/L, respectively (Brooke, 1993b). There were also three Level 3 reproduction studies with effects on biomarker endpoints for induction of plasma VTG at 71 µg/L (Harries *et al*., 2000) and testis histopathology at 1.6 µg/L (Miles-Richardson *et al*., 1999).

There was one valid Level 4 study with *P. promelas* which was a FELS test with a NOEC of 7.4 µg/L based on survival (Ward and Boeri, 1991a). This effect was assessed as suitable for use in PNEC derivation according to ECHA (2014a). It is noted that the study did not include mechanistic endpoints.

Danio rerio

There was one Level 4 study with Zebrafish considered in ECHA (2014), which was of similar design to an FSDT (60 days) (Lin and Janz, 2006) but the study included no analytical verification and therefore is used as supporting information only. There were effects on sex ratio at the lowest (nominal) test concentration of 10 µg/L, and therefore no NOEC was derived for the study. VTG induction and gonad histology were also affected at 10 µg/L (nominal). Two Level 5 multigeneration studies (up to four generations) with *D. rerio* were also not considered reliable and relevant for PNEC derivation due to the wide spacing of test concentrations, limited analytical verification and lack of transparency of reporting (so that validity criteria could be checked) (Sun *et al*., 2017; 2021). The results from these studies are generally supportive of data for other fish species with effects on hatching, growth, sex ratio and gonads at 20.3 µg/L (albeit not following a concentration response for sex ratio). There was limited evidence that the effects of NP were worsening in subsequent generations. The effect on sex ratio was observed in the F1 generation but not the F2 generation (Sun *et al*., 2017). Effects on hatching and fertility ratio were observed at similar test concentrations across multiple generations (Sun *et al*., 2021). There was a reduction in male body length at a lower test concentration in the F2 (20.3 µg/L) compared with F1 generation (182.6 µg/L) in one of the studies (Sun *et al*., 2021), but this was not observed in the other study (Sun *et al*., 2017).

Salmonids / *Salmonidae* (*Oncorhynchus mykiss*)

There were two Level 3 reproduction screening assays with Rainbow Trout (*Oncorhynchus mykiss*) with effects on mechanistic endpoints only. In one study, VTG was significantly induced at 20.3 µg/L, and GSI was affected at 54.3 µg/L (Jobling *et al*., 1996), while in the second study VTG was induced at 8.3 µg/L, and there were effects on GSI and increased plasma oestradiol at 85.6 µg/L (Harris *et al*., 2001).

There are also two Level 4 studies with *On. mykiss*, a FELS test (Spehar *et al*., 2010) and two studies similar to an FSDT (Ashfield *et al*., 1998; Ackermann *et al*., 2002). A NOEC of 6 µg/L was reported in the FELS study based on reduced body weight (- 29%) and length (- 14%) after 91 days' exposure. This result was the lowest NOEC used in the ECHA (2014) restriction dossier. There were no biomarker endpoints included in the study. There was also an effect on growth at 10 µg/L in one of the FSDT-style studies, but the study does not meet the relevance criteria for PNEC derivation in the present assessment because of the lack of supporting chemical analysis (Ashfield *et al*., 1998). The third FSDT style study tested up to a maximum concentration of 10.17 µg/L without adverse effects on population-relevant endpoints (Ackermann *et al*., 2002). For biomarker endpoints Ackermann *et al*. (2002) determined effects on hepatic VTG at 1.05 µg/L and on zona radiata protein at 10.17 µg/L, while in the Ashfield *et al*. (1998) study there were effects on the ova-somatic index at 30 µg/L (nominal).

There is also a Level 5 study with *On. mykiss* (Schwaiger *et al*., 2002) which is used as supporting information only in the ECHA (2014a) restriction dossier (intermittent dosing on 10 days each month for four months, only two wide spaced test concentrations of 1 and 10 µg/L, limited replication, F1 generation were not exposed to NP) and also in the present assessment. However, the study indicates effects on hatching at 1 µg/L and confirms effects on biomarker endpoints for induction of male VTG (at both 1 and 10 µg/L) in the F0 generation. There were also some effects on biomarkers in the F1 generation (despite no direct exposure to NP). These were increased VTG in females, increased testosterone in females and increased oestradiol in males (all at 10 µg/L).

Other fish studies

The available dataset includes fish growth studies with three species. In a study with the Bluegill (*Lepomis macrochirus*) the NOEC was 59.5 µg/L based on mortality (Brooke *et al*., 1993b). In a study with Brown Trout *Salmo trutta* the NOEC was > 100 µg/L based on growth (Shirdel *et al*. 2016) and with the Copper Redhorse (*Moxostoma hubbsi*) the NOEC was > 50 µg/L based on adult mortality (IUCLID, 2014). There is also a Level 3 FSTRA-style study with the cyprinid Chinese Rare Minnow (*Gobiocypris rarus*) but without adverse population level effects at the highest test concentration of 20 µg/L (Zha *et al*., 2008). There were however testis-ova at 18.53 µg/L and induction of male plasma VTG at 4.52 µg/L, which is similar to the results observed with other fish species. These are not OECD CF Level 4 or 5 studies. The high NOEC values (> 50 µg/L for population-relevant endpoints) confirms that they have low sensitivity compared with Level 4 and 5 fish studies, where the NOECs range from 0.64 – 31 µg/L.

It is worth noting that there is a reliable 96 hour LC_{50} (exposure concentration lethal to half of the test animals) of 17 µg/L for 2 day old larvae of Winter Flounder (*Pleuronectes americanus*) (Lussier *et al*., 2000). This suggests very high sensitivity for a marine fish species. An estimate of long-term toxicity could be explored using an acute-to-chronic ratio approach using data for other species. However, a NOEC derived this way would not be positively associated with ED effects. It would add to the conservatism of the assessment, but there would be significant uncertainty about its relevance for this project. It has therefore not been considered further.

3.1.2 Amphibians

An Amphibian Metamorphosis Assay (AMA) with *Xenopus laevis* is available (Kamel *et al*., 2022). This is a Level 3 study, most typically used for identifying endocrine activity via the thyroid, however it is the only amphibian study available for NP. There were no effects on thyroid mediated endpoints including thyroid histopathology. There was an increase in bodyweight at the highest test concentration of 48.2 µg/L at day 7 but as there was no effect at day 21, this is not considered population-relevant. At day 21 snout-vent length (SVL) was significantly increased by 11, 21, 23 and 21 % at 1.82, 5.17, 12.5 and 48.2 µg/L. Increased SVL could be indicative of general toxicity and related to delayed development, however there was no corresponding effect on body weight or on development stage. An effect on SVL alone is not considered thyroid-mediated and the population relevance of the effect is unclear. The relevant NOEC from this study is therefore either 0.91 µg/L (LOEC/2) or 48.2 µg/L. As this is a Level 3 study it was not used for PNEC derivation.

3.1.3 Invertebrates

The long-term studies on invertebrates exposed to NP are given in Table 46 of Appendix 155. These are studies whose design would be expected to fall under Levels 4 or 5 of the OECD CF.

There was one study with the nematode *Caenorhabditis elegans* (Höss *et al*., 2002) which was considered reliable and relevant according to ECHA (2014a). The endpoints from this study were increased growth at 65.6 µg/L and increased reproduction at 40.2 µg/L (the lowest concentration tested). ECHA (2014a) chose to use a NOEC of 40.2 µg/L, without explaining why this was acceptable, given that increased reproduction was observed at the lowest test concentration, and because the increase in reproduction/growth in this study does not appear to be adverse. A clear concentration response relationship is observed with increased reproduction and associated growth up to a plateau of about 200 µg/L (i.e. the larger worms are showing enhanced reproduction). In ECHA (2014a) it is stated that the reason for the stimulating effects is not known. Recent attempts to standardise the *C. elegans* assay have found that the live food source (bacteria *Escherichia coli* at 10^9 cells/mL in the NP study) may affect the study result either through metabolism of the test substance or indirectly if the substance has antibiotic activity (Beydoun *et al*., 2021; van der Voet *et al*., 2021). NP concentrations were not maintained within ± 20% in the Höss *et al*. (2002) study (mean measured concentrations being 65.6 - 80% of nominal over the 72 hour exposure period) and therefore it is plausible that the NP concentrations were reduced in the study due to bacteria utilising NP as a food source. Since it is not clear whether the observed effects on *C. elegans* are an adverse effect of the test item or are an artifact of the test design, the study is not considered suitable for PNEC derivation.

There is one valid study with an insect reported in the ECHA (2014a) restriction dossier. The result was from a 56 day water exposure study with the midge (*Chironomous tentans*) with a NOEC of 42 µg/L based on mortality of the larvae (Kahl *et al*., 1997). No other studies with insects were identified.

There were several studies with the water flea, *Daphnia magna*, which are summarised in the ECHA (2014) restriction dossier. These were Scholz *et al*. (1992), Brooke (1993a), Comber *et al*. (1993), Fliedner (1993), Baldwin *et al*. (1997), Kopf (1997), Sun and Gu (2005) and Brennan *et al*. (2006). Of these studies three results were considered valid and these generated NOECs of 3.45 µg/L (Fliedner, 1993), 24 µg/L (Comber *et al*., 1993) and 116 µg/L (Brooke, 1993a), all based on reproduction. A fourth study was considered reliable but there was no effect on reproduction at the highest test concentration of 100 µg/L (Scholz *et al*., 1992). The value used for the ECHA (2014a) SSD was a geometric mean of 59.5 µg/L using the results from Comber *et al*. (1993) and Brooke (1993a). Fliedner (1993) is not used for deriving a PNEC in the ECHA (2014a) restriction dossier, presumably because they could not access the source data. However, according to ECHA (2012) this is a GLP study assessed as reliable without restriction (R1). Including this result in the present assessment means a geometric mean of 21.3 µg/L can be derived based on reduced reproduction from three *D. magna* reproduction studies (Fliedner (1993), Comber *et al*. (1993), Brooke (1993a)).

Furthermore, a new multigeneration study with *D. magna* was identified as part of the updated literature search (Campos *et al*., 2016). This study covers four generations, the parental generation (F0) and three subsequent generations (F1, F2 and F3). The experimental design for each generation is based on OECD TG 211. There are some concerns regarding the reporting of this study with no information on the actual physicochemical parameters or measured test concentrations (reported as being within 10% nominal), and no data are presented for reproductive endpoints (which were not affected by exposure). It was therefore not possible to confirm whether test validity criteria were met, except for survival which was 100% in the controls in each generation. The main affected endpoint was survival, which was reduced at 20 and 40 µg/L in the F0 and F3 generations, at 20 µg/L in the F2 generation (but not 40 µg/L) and at 40 µg/L only in the F1 generation. The NOEC based on survival was therefore 10 µg/L. There was a reduction (- 5.7%) in initial body length at 10 µg/L in the F3 generation only but as this effect was not evident in the final body length measurement it was not considered a population-relevant effect. The overall NOEC from this study was therefore 10 µg/L based on juvenile survival.

The ECHA (2014a) restriction dossier considered that a NOEC of 88.7 µg/L based on effects on reproduction of *Ceriodaphnia dubia* in a study by England (1995) was reliable for PNEC derivation. A new study with *Ceriodaphnia silvestrii* was also identified in the published literature (Spadoto *et al*., 2018). This was an 8 day reproduction study according to the Brazilian National Standard (NBR 13373), with the test being repeated three times to obtain the result. The LOEC was 19 µg/L and NOEC was 15 µg/L based on number of offspring per female.

There is a valid GLP study reported in the EU REACH registration dossier with the freshwater mollusc (*Potamopyrgus antipodarum*), conducted according to OECD TG 242, using semi-static exposure to 4-NP (IUCLID 2020). There was a significant concentration dependent decrease in reproduction (mean number of embryos per female) at 14.7, 24.7 and 66.7 µg/L compared with the control after 28 days' exposure. The NOEC from this study was 6.01 µg/L.

A study investigating the effects of 4-NP on the growth of two freshwater bivalves (rainbow mussel, *Villosa iris* and fatmucket, *Lampsilis siliquoidea*) using a flow-through test is also available (Ivey *et al*., 2018). Growth (dry weight and biomass) was reduced at the highest test concentration for both species (nominal 200 µg/L). However, the measured test concentration at this level was 30 µg/L, and this discrepancy could not be explained.

Three marine invertebrate studies were considered valid for use in an SSD according to ECHA (2014a). This included two studies with the mysid *Americamysis bahia*, one with a NOEC of 3.9 µg/L (based on growth) (Ward and Boeri, 1991b) and the second with a NOEC of 9.5 µg/L based on survival and reproduction (Kuhn *et al*., 2001). The value of

3.9 µg/L was used for the SSD by ECHA (2014a). There is also a valid NOEC of 20 µg/L from a reproduction study with the marine copepod *Tisbe battagliai* (Bechmann *et al*., 1999).

A study with another marine copepod (*Tigriopus japonicas*) reported a NOEC of 0.01 µg/L (nominal) based on *nauplii* development in the F1 generation (Marcial *et al*., 2003); however, there is wide spacing of the test concentrations, and no supporting chemical analysis. It has therefore not been considered further.

3.2 PNEC derivation with consideration of ED endpoints

In this section aquatic PNECs for NP have been derived using only data generated in OECD CF Level 4 and 5 studies with invertebrates and vertebrates.

There were reliable and relevant data for 10 species, seven invertebrates (including two marine crustaceans) and three fish. Most of the results were from studies from Level 4 of the CF. These included FELS tests with two of the fish species. There were three results from Level 5 studies, only one of which was for fish – a MEOGRT with *O. latipes* (Watanabe *et al*., 2017). The other two were for invertebrates: a multigeneration test with *D. magna* (Campos *et al*., 2016), and a life cycle test with *C. tentans* (Kahl *et al*., 1997).

3.2.1 Suitability of the dataset for derivation of a PNEC for an EDC and data gap analysis

This section assesses the reliable and relevant dataset collated in the preceding sections against the SETAC Pellston Workshop® criteria for deriving a PNEC with an EDC (Matthiessen *et al*., 2017).

1. *Have the most appropriate taxa been tested (with relevant endpoints)?*

Based on the MoA as an ER agonist, vertebrates would be expected to be the most sensitive taxonomic group. There are reliable NOECs for three standard OECD test species of fish (*O. latipes*, *On. mykiss* and *P. promelas*), representing three different families. A fish (*O. latipes*) was the most sensitive species tested, based on the result from a MEOGRT that included multiple endocrine-relevant measurements (Watanabe *et al*., 2017). However, the NOECs for the other two species (*On. mykiss* and *P. promelas*) were both from FELS tests which did not include relevant endocrine mediated endpoints. There are no marine fish species in the dataset, nor any species with alternative reproductive strategies (e.g. live bearers).

Invertebrates were also sensitive to NP with reliable NOEC values between 3.9 and 88.7 µg/L for seven species (Table 2). The NOEC of 3.9 µg/L for reproduction of *A. bahia* was the second most sensitive in the dataset. However, the apparent sensitivity may be because there are more invertebrates than fish in the dataset, and the available fish studies may not have measured the most sensitive endpoints. This is partly because studies with other fish species did not meet the relevance criteria for inclusion in the assessment.

2. *Have sensitive life stages, or the entire life cycle, been tested (with relevant endpoints)?*

For fish there was one relevant and reliable NOEC from a Level 5 multigeneration study for use in PNEC derivation. A NOEC of 0.64 µg/L was derived from the MEOGRT with *O. latipes* and was the most sensitive result in the dataset. Effects were seen at the lowest concentration tested so the NOEC is an extrapolation, which introduces some uncertainty. Nevertheless, confidence that this result is protective for this species is high because results from four other Level 4 or 5 studies with the same species all had much higher NOECs (8.2 to 31 µg/L) (Yokota *et al*., 2001; Seki *et al*., 2003; Balch and Metcalfe, 2006; Horie *et al*., 2021). Experimental design may explain the difference between the sensitivity of *O. latipes* in the two Level 5 studies, because the MEOGRT has been optimised for increased control replication to increase the statistical power of the test for detecting changes.

Where there are Level 5 studies of lower reliability, these can be used as supporting information for considering differing species sensitivities. A Level 5 study with *On. mykiss* using intermittent dosing indicates a LOEC of 1 µg/L based on mortality before the eyed egg stage in the F1 generation (10.1% at 1 µg/L compared with 1.7% in the control) (Schwaiger *et al*., 2002). The NOEC from a Level 4 FSDT-type study was ≥ 10.17 µg/L (Ackermann *et al*., 2002), whereas the NOEC from a Level 4 FELS study for this species was 6 µg/L (Spehar *et al*., 2010). It is therefore possible that a reliable Level 5 study may give a more sensitive result than the NOEC selected for this analysis. Similarly, multigeneration studies with a cyprinid species (*D. rerio*) suggest a NOEC in the region of 2 µg/L (Sun *et al*., 2017; 2021), which is lower than the NOEC of 7.4 µg/L for another cyprinid species (*P. promelas*) from a Level 4 FELS study (Ward and Boeri, 1991a).

The NOECs from three FSDT with *O. latipes* range from 11.6 to 31 µg/L, which are all higher than the NOECs from FELS tests with *P. promelas* and *On. mykiss* (of around 6 to 7 µg/L). This is assumed to be a species difference since the FSDT study design would usually be expected to be more sensitive compared with the FELS test, due to the relatively (on a species basis) longer test duration and inclusion of additional endpoints, although the test duration of the *On. mykiss* FELS test and the *O. latipes* FSDT are similar. Nevertheless, each NOEC is explicitly linked to the test design and dose spacing used in individual experiments, which complicates comparisons.

Taken together, the available data for fish imply that *O. latipes* might not be the most sensitive species if comparable Level 5 data were available for other commonly tested species.

The dataset also includes multigeneration studies with invertebrates including *D. magna* (Campos *et al*., 2016), and a life cycle test with *C. tentans* (Kahl *et al*., 1997). It is rare to find substances which have multigeneration studies with both invertebrates and fish.

3. *Have delayed and multigenerational (i.e. latent) effects been considered?*

Multigeneration tests such as the MEOGRT (Watanabe *et al*., 2017) should be sufficient to address any concern for delayed or latent effects of NP. There is minimal evidence for toxicity of NP increasing over multiple generations. In Watanabe *et al*. (2017) the F2 hatching was less sensitive compared with the F1 fertility endpoint, and in the multigeneration test with Zebrafish, effects were similar across three generations (Sun *et al*., 2017; 2021). In the *Daphnia* multigeneration study the F0 generation was the most sensitive (Campos *et al*., 2016).

4. *Do NMDRs or other unusual temporal patterns of toxicity affect the ability to predict reliable no-adverse-effect levels?*

There was no convincing evidence for NMDRs in studies with NP based on either population-relevant or mechanistic endpoints. In the MEOGRT study there is some evidence for increased body weight of F1 adult male and female fish at higher test concentrations, but it is unclear if this is a chemical effect or a feature of the test design where a cohort of fish are selected for breeding (Watanabe *et al*., 2017, discussed in Appendix 179). The biomarkers for ER agonist activity (VTG, SSC, gonad phenotype and presence of testis-ova) all followed a clear concentration response pattern in this study.

An inverted concentration response based on increased growth and reproduction was reported for the nematode *C. elegans* (Höss *et al*., 2002). Based on the discussion in Section 45, it is inferred that this effect could potentially be a consequence of feeding live bacteria which can utilise NP as a food source. For this reason, it is not considered to be either endocrine related or an adverse effect and therefore does not constitute an NMDR of concern.

5. *Does a threshold for adverse endocrine-mediated effects exist?*

Figure 1 shows the range of NOEC values for endocrine population-relevant (reproduction and sex ratio) and “biomarker” endpoints (VTG induction, gonad histopathology, SSC, steroid hormone levels and presence of testis-ova) from 27 reliable OECD CF Level 3, 4 or 5 fish studies (note that most of these studies were not used for PNEC derivation, either because they did not meet relevance criteria, i.e. Level 3 studies were excluded, or because there was a more sensitive result with the same species). Most NOEC values fall between 1 and 100 µg/L regardless of study type. The lowest “biomarker” NOEC was 0.4 µg/L based on testis histology in an extended FSTRA style study with *P. promelas* (Miles-Richardson *et al*., 1999). These data would appear to demonstrate a threshold level for endocrine effects in fish with NOECs being above 0.1 µg/L.

For population-relevant endpoints the most sensitive effects (at approximately 1 µg/L) were from Level 5 fish tests (Schwaiger *et al*., 2002 (supporting information only), Watanabe *et al*., 2017). This supports the assumption that multigeneration studies are necessary to fully define an Environmental Threshold of No Concern (ETNC) for an oestrogenic substance. It should also be noted that true NOECs could not be obtained for either of these tests.

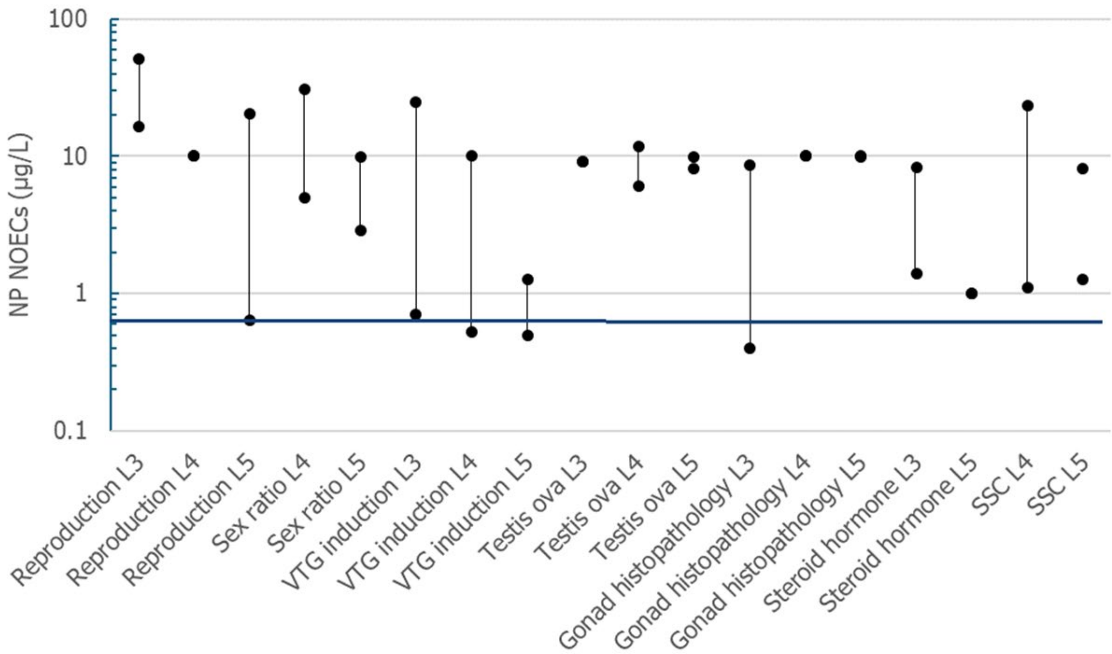


Figure 1 Range of NOEC values from all reliable OECD CF Level 3, 4 and 5 fish studies with NP. The blue line indicates the lowest population level NOEC of 0.64 µg/L from a MEOGRT (Watanabe *et al.*, 2017).

Overall, the number of relevant and reliable studies with fish (representing the group of organisms which would be expected to have greatest sensitivity based on the ER MoA) for use in PNEC derivation is low, comprising NOECs from only three species. Only one of these studies included EAS-mediated endpoints (the Level 5 multigeneration study) with the others being from Level 4 FELS tests. For an ED-specific PNEC, based on concern for ER agonist activity, uncertainty could be reduced if data were available for more fish species, preferably from Level 5 multigeneration studies or a Level 4 test incorporating an endocrine mediated endpoint such as sex ratio (e.g. the FSDT) and for marine species and/or species with alternative reproduction strategies. Nevertheless, there are studies with endocrine biomarker endpoints as summarised in Table 2 which can be used as supporting information, including a Level 5 test with *On. mykiss* (Schwaiger *et al*., 2002). Furthermore, multigeneration studies with *D. rerio* (not used for PNEC derivation due to lack of supporting analytical data or reporting of validity criteria) do not report NOEC values < 1 µg/L (Sun *et al*., 2017; 2021). It is therefore considered that there are sufficient data for NP to satisfy the SETAC Pellston Workshop® criteria for vertebrates.

3.2.2 Deterministic PNEC assessment for NP

The reliable dataset for fish includes only one multigeneration fish study incorporating endocrine mediated endpoints covering the ER agonist MoA. This is also the most sensitive result in the reliable dataset by almost an order of magnitude. It is therefore recommended that a deterministic approach is used for deriving a PNEC for nonylphenol.

Using standard approaches the PNEC for NP is derived deterministically by applying an AF of 10 to the NOEC of 0.64 µg/L from a MEOGRT (Watanabe *et al.*, 2017), as follows:

NP deterministic PNEC = 0.64 /10 = 0.064 µg/L

A case could be made for the application of additional AF based on the substance having a confirmed endocrine MoA. However, the endocrine mechanism is already accounted for because a multigeneration fish study is included in the dataset. The PNEC is also considered protective as it is below the lowest “biomarker” NOEC of 0.4 µg/L based on testis histology in an extended FSTRA style study with *P. promelas* (Miles-Richardson *et al.*, 1999).

Further discussion of the AFs is given in Section 78.

3.2.3 Probabilistic assessment for NP

As there are only three reliable NOECs for vertebrates it is not possible to conduct a probabilistic assessment with vertebrate data only. There is, however, enough data for a probabilistic assessment if the reliable and relevant OECD CF Level 4 and 5 invertebrate data are included. Whilst the effects are not necessarily associated with an ED MoA, invertebrates are relatively sensitive to NP (the NOEC of 3.9 µg/L for effects on growth of *A. bahia* being the second most sensitive result in the dataset). Their inclusion may therefore provide an additional layer of precaution. The total dataset of reliable values is comprised of 10 species (seven invertebrates and three fish) and summarised in Table 2.

Table 2 Values selected for constructing an SSD for NP for this report

Taxon	Species	Endpoint (NOEC)	Value (µg/L)	Value selected for SSD	Reference
Insects	*Chironomus tentans*	Larval mortality	42	42	Kahl *et al.*, 2002
Gastropod molluscs	*Potamopyrgus antipodarum*	Reproduction	6.01	6.01	IUCLID, 2020
Crustaceans	*Ceriodaphnia dubia*	Reproduction	88.7	88.7	England, (1995)
	Ceriodaphnia silvestrii	Reproduction	15	15	Spadoto *et al.*, 2018
	Daphnia magna	Survival	10	10	Campos *et al.*, 2016

Taxon	Species	Endpoint (NOEC)	Value (µg/L)	Value selected for SSD	Reference
		Reproduction[1]	24		Comber *et al.*, 1993
		Reproduction[1]	3.45		Fliedner, 1993
		Reproduction[1]	116		Brooke, 1993a
	Americamysis bahia	Growth	3.9	3.9	Ward and Boeri, 1991b
		Reproduction	9.1		Kuhn *et al.*, 2001
	Tisbe battagliai	Mortality	20	20	Bechmann *et al.*, 1999
Fish	*Oryzias latipes*	F0 post swim up mortality (sex ratio and fertility)	8.2	0.64	Yokota *et al.*, 2001
		Sex reversal	31		Horie *et al.*, 2021
		Growth	11.6		Seki *et al.*, 2003
		Fertility	0.64		Watanabe et al., 2017
	Pimephales promelas	Survival	7.4	7.4	Ward and Boeri, 1991a
	Oncorhynchus mykiss	Growth	6	6	Spehar et al., 2010

[1] Geometric mean of three *Daphnia* reproduction studies is 21.3 µg/L

SSDs for NP were generated using the log normal method based on ETX version 2.3 or the weighted average method using ssdtools.

The SSDs for the reliable and relevant NP dataset derived using ssdtools are shown in Figure 2. The following models showed the best fits to the data: the gamma distribution, the log logistic distribution (llogis), the lgumbel model, the log normal model (lnorm) and the Weibull model.

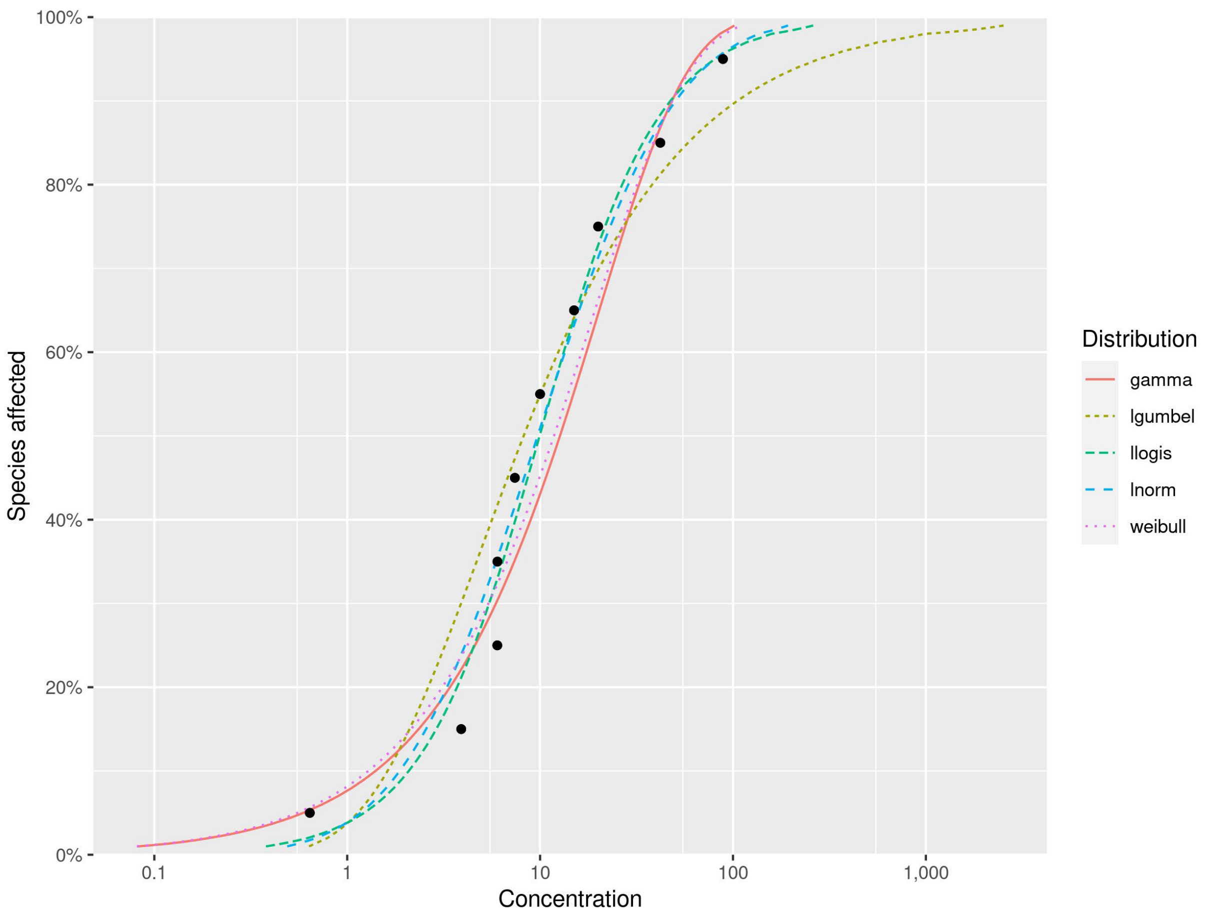


Figure 2 SSDs models from ssdtools for deriving an aquatic HC_5 for NP

The HC_5 values derived using the different distributions and using the ETX software are given in Table 3.

Table 3 HC_5 values derived using the different distribution models

Distribution	HC5 (µg/L)	Confidence Intervals (µg/L)	Anderson-Darling statistic	AICc[1]	Weight	Software
Log normal	0.97	0.19 – 2.45	Accepted	-	-	ETX v. 2.3
Average	0.953	0.245 – 4.61	-	-	-	ssdtools v. 1.0.2
Gamma	0.591	0.0417 – 5.33	0.365	-4.39	0.181	
Lgumbel	0.82	0.521-3.5	0.494	-2.68	0.077	
Llogis	1.23	0.322 – 4.61	0.198	-5.24	0.277	
Inorm[2]	1.17	0.376 – 4.82	0.242	-5.13	0.262	
Weibull	0.552	0.0875 – 4.64	0.309	-4.63	0.204	

[1] Akaike's Information Criterion corrected for sample size (AICc)
[2] Note that a small differences in model results (less than a factor of 2) for different estimation packages is expected due to different estimation strategies (Fox *et al*., 2021)

Using ssdtools it is possible to demonstrate how the different distributions, all of which have an acceptable fit to the dataset, lead to HC_5 values between 0.552 and 1.23 µg/L. Using the model average estimate recommended by Fox *et al*. (2021) the HC_5 is 0.953 µg/L.

For NP, this HC_5 value would not appear to be protective for effects in fish since it has an oestrogenic MoA and available data from fish multigeneration tests suggest adverse effects can occur below 1 µg/L. The available dataset is based on studies with invertebrates which are potentially less sensitive to NP, or from fish studies which do not include sufficient endocrine sensitive endpoints. Multigeneration fish studies are not well represented. The probabilistic approach is therefore not appropriate for NP, and potentially for many other endocrine active substances, since it is not practical or ethically justified to conduct multiple multigeneration fish studies for a single substance risk assessment.

3.2.4 Comparisons with other NP PNECs

PNECs for NP derived in other assessments are given in Table 4 and compared with the deterministic PNEC of 0.064 µg/L derived in Section 33 of the present report. The PNEC of 0.42 µg/L from ECHA (2014a) is based on a probabilistic HC_5 of 2.12 µg/L with an AF of 5 applied. The standard approach outlined in ECHA (2008) was used to derive the SSD which includes both algae and invertebrates and also pre-dates the MEOGRT study

(Watanabe *et al*., 2017). The ECHA (2014a) assessment therefore does not account for the endocrine MoA of NP and would not necessarily be protective for fish. An additional AF of 5 was applied to the initial PNEC of 0.42 µg/L to account for this uncertainty.

Table 4 Comparison of PNECs derived for NP

Type of PNEC	PNEC (µg/L)	AF	Reference
Deterministic[1]	0.064	10	This report
SSD (non-fish)	0.42	5	ECHA, 2014
SSD (fish)[2]	0.08	5	ECHA, 2014
SSD	0.0365	5	EC, 2022

[1] Dataset includes a multigeneration study with fish incorporating endocrine relevant endpoints

[2] ECHA proposed an additional AF of 5 to be applied to the initially derived probabilistic PNEC of 0.42 µg/L to be protective of fish however they did not provide an opinion on whether the PNEC would be protective for the ED properties.

More recently, the EU derived a probabilistic PNEC of 0.037 µg/L for NP for use as an EQS (EC, 2022), generally following the relevant technical guidance (EC, 2018). The EQS has been approved by the EC Scientific Committee on Health and Environmental Risks (SCHEER, 2022). The EQS is lower than the PNEC derived in the present assessment following a deterministic approach. For EQS derivation a probabilistic approach was used, including only data for invertebrates and vertebrates (similar to the approach described in Section 34). However, the dataset was more extensive compared with the present assessment, with NOEC values being used from studies without supporting analytical verification (see Table 5). Use of the NOECs of ≤ 1 µg/L (nominal) from the studies with *Gambusia holbrooki* (Drèze *et al*., 2000) and *C. dubia* (Isidori *et al*., 2006) serve to lower the EU EQS, compared to the PNEC derived in the present assessment. These studies both have significant methodological issues (aside from lack of supporting analytical verification) that preclude their use in probabilistic risk assessment. The study with *G. holbrooki* did not show a clear concentration response including for endocrine endpoints for sex ratio and gonad histopathology where data from the mid-nominal test concentration (5 µg/L) was pooled with data from the control to assess effects at the other test concentrations (0.5 and 50 µg/L nominal). There are reliable data for *C. dubia* reproduction based on a regulatory style study used in the present assessment (England, 1995), rather than the Isidori *et al*. (2006) study which used cladocerans obtained from MicroBioTests™ rather than continuous laboratory culture. In addition, the inclusion of a low sensitivity NOEC for *Mytilus edulis* based on a 72-hour exposure duration (Gramno *et al*., 1989) may have impacted the slope of the SSD curve leading to a lower HC_5 value.

Table 5 Data included for the NP EQS assessment (EC, 2022) but not ECHA (2014)

Species	NOEC (µg/L)	Analytical verification	Reference
Gambusia holbrooki	0.5	N	Drèze *et al.*, 2000
Gobiocypris rarus	10	N	Zha *et al.*, 2007
Rana nigromaculata	20	N	Yang *et al.*, 2005
Mytilus edulis	200	N	Gramno *et al.*, 1989
Ceriodaphnia dubia	1	N	Isidori *et al.*, 2006

The dataset used for EU EQS derivation does not include any data post-2015. This means that it does not include the multigeneration study with *D. magna* (Campos *et al.*, 2016), the mollusc reproduction study with *P. antipodarum* (IUCLID, 2020) or more notably the FSDT (Horie *et al.*, 2021) and MEOGRT (Watanabe *et al.*, 2017) with *O. latipes*. Therefore, while the deterministic PNEC of 0.064 µg/L from the present assessment is in the same order of magnitude as the EU EQS value of 0.037 µg/L, the dataset on which it is based contains higher quality and more relevant data.

The HC_5 outcomes from SSDs depend both on the fitted distribution and the data chosen for inclusion in the SSD. These discrepancies were the topic of a European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) workshop (ECETOC, 2014). Recent developments suggest the emergence of two alternative approaches to deriving probabilistic SSDs which are discussed in Fox *et al*. (2021). The first broadly follows the approach taken for NP in the present assessment focusing on the use of relevant (e.g. sensitive species and endpoints) and reliable data. A problem with this approach is that the dataset is often limited, and the SSD is therefore based on a small dataset. This has led to a focus on the how to best use mathematical models for SSD curve fitting and for deriving reliable HC_5 values, for example by using the weighted average approach (Fox *et al*., 2021).

Posthuma *et al*. (2019), however, argue that the application of strict data criteria excludes NOEC values leading to SSDs which are less statistically robust due to small datasets. They suggest an approach which includes all available data in the SSD with specific investigation only where data points appear to be implausible. A quality score is applied to the SSD considering the following aspects: derivation of the SSD slope, taxa representation, data origin and whether data bridging was used to derive the SSD. Even using this approach, which does not discount use of quantitative structure-activity relationship (QSAR) or read across data, priority is given to NOEC values based on measured concentrations of the test substance. The lack of data quality checks has allowed SSDs to be derived for thousands of chemicals (Posthuma *et al*., 2019).

Nevertheless, these SSDs have not been rigorously assessed for plausibility and therefore are only suitable as a screening level assessment. There is already recognition that these SSDs can be refined, for example by focusing on sensitive taxonomic groups (Oginah *et al*., 2023).

3.3 Summary and conclusion

The dataset of relevant and reliable OECD CF Level 4 and 5 studies for use in deriving an aquatic PNEC for NP has been updated. The most sensitive result in the updated dataset is the NOEC (derived from the LOEC/2) of 0.64 µg/L based on reduced fecundity and fertility in a MEOGRT study (Watanabe *et al*., 2017).

The appropriateness of this dataset to derive a threshold level for an EDC has been evaluated against the SETAC Pellston Workshop® criteria. While the dataset contains multigeneration studies with both fish and invertebrates covering sensitive life stages, there was only a single reliable NOEC from a Level 5 fish study for NP. The other two NOECs for fish were from Level 4 FELS studies, and these species might be expected to be more sensitive to NP in longer studies which include endpoints for either reproduction and or sex ratio. Nevertheless, the available data (including supporting multigeneration fish studies) do not flag significant concern for NMDR, or delayed effects, and so there is no specific reason why an ETNC cannot be derived for NP. The key biomarkers for ER agonist activity (VTG induction, SSC, gonad histopathology) typically followed a concentration response with the lowest NOEC for these endpoints being 0.4 µg/L.

A probabilistic approach to deriving a PNEC for NP was considered inappropriate because the HC_5 value of 0.953 µg/L, which included NOECs from three species of fish and seven invertebrate species, was not protective for effects in the MEOGRT. It is therefore assumed that an SSD relevant for the ED properties would need to include more (or exclusively) data from multigeneration fish studies. This is not practically or ethically justified and therefore a deterministic PNEC of 0.064 µg/L was derived by applying a default AF of 10 to the lowest NOEC value of 0.64 µg/L, following standard risk assessment approaches. An AF of 10 is justified because the dataset includes multigeneration fish studies and because the result would also be protective for “biomarker” responses which are usually expected to occur at test concentrations lower than population-relevant effects.

As this project is concerned with defining sufficient datasets for PNECs for EDCs, the suitability of the AF of 10 is reconsidered based on the available datasets for other EDCs in Section 77, after an evaluation of the datasets for OP and EE2.

4 PNEC derivation: Octylphenol

4.1 Long-term aquatic toxicity data

A regulatory assessment undertaken by ECHA (2011), conducted to assess whether OP should be identified as a Substance of Very High Concern (SVHC), was used as the basis for this assessment. Studies considered valid as part of ECHA (2011) were generally accepted as appropriate for PNEC derivation. However, since the purpose of the ECHA (2011) assessment was not to derive a PNEC, there has been a re-evaluation of the reliability of some of the studies where they may impact the outcome of the present assessment (i.e. those which might result in the lowest NOEC for a particular species). Some of the relevant studies are also discussed in EA (2005). Studies published in the scientific literature since the previous assessment have also been included. Detailed information on the substance identity and physicochemical properties of OP is given in Appendix 149.

4.1.1 Fish

The OECD CF Level 4 and 5 fish studies which include endpoints suitable for PNEC derivation are summarised in Table 49 and Table 50 of Appendix E, respectively. Level 3 fish and amphibian studies are summarised in Table 48 of Appendix E. There are reproduction screening assays with nine different fish species, typically with VTG as the main endpoint. The LOECs for VTG effects occur in the range 10 to 269 µg/L. Effects on other endpoints, including reproduction and gonad histopathology, occur within the same range. Mechanistic endpoints are only used as supporting information.

Adrianichthyidae (*Oryzias latipes*)

There are six FSDTs with *O. latipes*, including three conducted as part of the ring test validation work for OECD TG 234 (Seki *et al*., 2003; Gray *et al*., 1999a; Gray *et al*., 1999b; OECD, 2012). Sex ratio was affected at 12.1 to > 50 µg/L, VTG at 11.4 to 104.7 µg/L and testis ova induction at 23.7 to > 100 µg/L. Sex ratio (increased number of females) was the most sensitive population-relevant endpoint. The lowest valid NOECs for sex ratio are 23.5 µg/L (OECD, 2012; Lab 9) and 23.7 µg/L (Seki *et al*., 2003), The geometric mean value for sex ratio from these two studies is 23.6 µg/L. However, there is also a LOEC of 12.1 µg/L (unbounded result) from a reliable study (OECD, 2012; Lab 5).

There are four OECD CF Level 5 fish tests with OP. Two of these studies are from fish life cycle toxicity tests (FLCTT) (which exposed fish from embryos and these are summarised in ECHA (2011). One of these studies (Ministry of the Environment of Japan, 2002) reporting a NOEC of 30.4 µg/L, based on reduced fertility and fecundity, is from a secondary source and therefore limited details are available. The second study, which was considered reliable according to ECHA (2011), reported significant effects on offspring mortality, and altered sex ratio at the lowest test concentration of 2 µg/L (unbounded result) although these effects did not follow a clear concentration response (Knörr and

Braunbeck, 2002). A re-assessment of the reliability of this study indicates that it is not suitable for PNEC derivation due to the low-test replication and other uncertainties in the test design and statistical analysis. The experimental design of these two studies differs from the MEOGRT which is initiated with mating pairs of fish and includes more control replicates to improve the statistical robustness.

OP was one of the test substances included in the validation of the OECD 240 test guideline for the MEOGRT. Two MEOGRT studies were conducted, one by the United States Environmental Protection Agency (US EPA) and one by Japan’s National Institute for Environmental Studies (NIES), and both are summarised by Flynn *et al*. (2017). The study conducted at the US EPA is considered reliable with restriction (R2), although it would not meet the control performance criteria for the published OECD 240 test guideline for fecundity in the F1 and F2 generations (17.4 and 19.5 eggs/pair/day respectively compared with recommended ≥ 20 eggs/pair/day) and fertility (78%, which is less than the required 80%) in the F2 generation. Overall reproduction rates (i.e. combined mean fecundity and percent fertility for the F0, F1 and F2 generations would meet MEOGRT test guideline requirements).

The study conducted at NIES did not have analytical verification and therefore has been used in the present assessment as supporting information only. In addition, the LOEC values reported in a summary table in the paper were difficult to cross check with the more detailed summary of data included as supplementary information, because statistically significant differences between the control and treatments were not reported in the supplementary information data summary. This study would have met the test validity criteria for fecundity and fertility in both the F1 and F2 generations (F0 data not considered).

The most sensitive endpoint in the US EPA study according to the summary table in Flynn *et al*. (2017) is a LOEC of 12.5 µg/L based on a 10% increase in length. Fertility was also significantly reduced at 51.1 µg/L in the F2 generation, while hatching success was significantly reduced at 102 µg/L. Since increased body size could be related to the ER agonist MoA the population-relevant NOEC from the US EPA study was therefore 6.2 µg/L based on increased body length.

There was an effect on hatching at nominal 25 µg/L in the NIES study, although this did not follow a clear concentration response, and a significant increase in body weight was observed at 25 µg/L (NOEC = nominal 12.5 µg/L). The LOEC for SSC (number of anal fin papillae) and VTG induction (based on both studies) was 6.25 µg/L in the NIES study. The lowest NOEC from the NIES MEOGRT was 12.5 µg/L (nominal), from the US EPA MEOGRT was 6.2 µg/L (Flynn *et al*., 2017), whereas the NOEC from the Ministry of the Environment of Japan (2002) FLCTT was 30.4 µg/L. The most sensitive (and precautionary) population-relevant endpoint from a reliable Level 5 study with OP and *O. latipes* was the NOEC of 6.2 µg/L based increased body length in the US EPA study (Flynn *et al*., 2017).

Cyprinidae (*Danio rerio*)

There are results for *D. rerio* from three FSDT (OECD, 2012) and one FLCTT (Wenzel *et al*., 2001). The three FSDT were conducted as part of the OECD validation of the test method (OECD, 2012). All three studies showed a consistent effect with VTG being induced in the range 26 – 42.5 µg/L and a sex ratio shift to females at 13.8 - 26 µg/L. The NOECs for sex ratio from these three studies were 5.7, 9.5 and < 13.8 µg/L. Using a geometric mean value (excluding the unbounded result) gives a NOEC of 7.4 µg/L. The FLCTT (Wenzel *et al*., 2001) is a reliable test according to ECHA (2011). There were effects on growth and reproduction (eggs per female, time to first spawn and percent fertilisation capacity) of the F0 fish exposed from embryos at 35 µg/L, although there were no effects on sex ratio of the F0 fish, or on survival, hatching or growth of F1 embryos. The NOEC from this study was 12 µg/L.

Stickleback / Gasterosteidae (*Gasterostreus aculeatus*)

Two FSDT studies with the stickleback, *Gasterostreus aculeatus* were conducted as part of the OECD TG validation (OECD, 2012). The test outcomes were based on mean measured concentrations because test solutions were not maintained within ± 20% nominal. Neither study met the current guideline criteria for hatching success of ≥ 80%. The most sensitive endpoint from these two studies was reduced hatching success at 66.9 µg/L from the test conducted by laboratory 6. Since hatching in both the control and solvent control in this study was close to the guideline criterion (77%) and post hatch survival was good (90 and 91%, respectively) the NOEC of 22.2 µg/L from this study is considered reliable for use in PNEC derivation.

Salmonids / Salmonidae (*Oncorhynchus mykiss*)

There were two studies with *On. mykiss*. A FELS test indicated a LOEC of 11 µg/L based on larval growth. This was a GLP study according to standardised guidelines that had previously been cited on IUCLID (EA, 2005; ECHA, 2011). However, this is from a secondary source, and it is not possible to check the details (i.e. it is not summarised on the ECHA portal, checked on the 20/12/2022). The NOEC from this study is 6.1 µg/L. In a second study, without supporting chemistry, an effect on growth at the lowest test concentration of nominal 1 µg/L did not follow a concentration response (Ashfield *et al*., 1998). Following the relevance criteria in Section 15 this result is not considered suitable for PNEC derivation.

Other fish species

A reliability assessment undertaken as part of the current project for a study initiated with pregnant female eelpout, *Zoarces viviparus* (Rasmussen *et al*., 2002), concluded that it was not reliable for use in PNEC derivation, because the fish were wild caught with minimal acclimation (seven days) and were not fed during the 35 day exposure. There was a reduction in embryo growth at the lowest test concentration of 14 µg/L. Two studies with the guppy, *Poecilia reticulata*, (Kinnberg *et al*., 2003; Toft and Baatrup, 2003) are only relevant as supporting information due to issues with the test design: in the Kinnberg *et al*.

(2003) study there was only a single test concentration, while the study by Toft and Baatrup (2003) reported results based on two separate experiments combined for analysis without reporting the effects from each individual study. Gonad histopathology was affected in the guppy at 26 µg/L in Kinnberg *et al*. (2003), whereas reduced body colouration was observed at 200 µg/L in Toft and Baatrup (2003).

4.1.2 Amphibians

For amphibians the available OP dataset comprises Level 3 screening studies with five species (Table 38). These are mainly development studies which are usually used to investigate mechanistic effects via the thyroid. There is one short term reproduction study with the Chinese brown frog, *Rana chensinensis* (Li *et al*. 2021). These studies were assessed for relevance of an alternative MoA via the thyroid but are not relevant for PNEC derivation because they are Level 3 mechanistic studies.

There is some limited evidence for effects via the thyroid in studies with *Xenopus laevis*. In a valid AMA, Kamel *et al*. (2022) reported accelerated development at 39.8 µg/L and significantly increased hind limb length (HLL) and normalised HLL at all test concentrations (≥ 5.43 µg/L), but with no effect on thyroid histopathology. The HLL effects did not follow a clear concentration response. In addition, for normalised HLL all developmental stages were grouped, rather than separating Nieuwkoop and Faber (NF) stage < 60 and NF stage > 60, which could have impacted the result. There was no effect on thyroid histopathology.

In a similar study with the streamside salamander (*Ambystoma barbouri*) (Rohr *et al*., 2003) an effect on SVL and hatching was only observed at the highest test concentration of 500 µg/L which also caused systemic toxicity in the form of reduced survival. Two studies investigating tadpole development of the Northern leopard frog (*Rana pipiens*) or the wood frog (*Rana sylvatica*) over 10 days or two weeks found no decrease in body weight at nominal concentrations > 200 µg/L (Crump *et al*., 2002, Hogan *et al*. 2006). In one exposure with *R. pipiens*, growth stimulation was observed at the lowest test concentration of 50 µg/L, when exposed from Gosner stage 36 but not Gosner stage 26 (Hogan *et al*., 2006). The biological and population relevance of the increased growth is unclear.

There is an OECD CF Level 3 short term reproduction style study with the Chinese brown frog (*Rana chensinensis*) with male only exposure (Li *et al*., 2021). There was reduced spawning and hatching at the highest test concentration of 206 µg/L (nominal).

The possible thyroid activity observed in the AMA study (Kamel *et al*., 2022) are not repeated in valid Level 4 larval amphibian growth and development assay (LAGDA) with *X. laevis* conducted as part of US EPA test guideline validation (Haselman *et al*., 2015; ABC labs, 2012; US EPA, 2013). These studies are summarised in Table 49. There were no effects of OP on apical endpoints for growth, development, or sex reversal at the highest test concentrations of 54.2 µg/L (Haselman *et al*., 2015) or 38 µg/L (ABC Labs, 2012). In the study by Haselman *et al*. (2015), there was reduced serum thyroxine and minor histopathological alterations of the thyroid in the form of increased incidence and severity of thyroid follicular cell hypertrophy at 25 and 54.2 µg/L, and increased incidence of moderate thyroid follicular cell hyperplasia at 54.2 µg/L. However, effects on gonad histopathology were much more sensitive (NOEC < 6.3 µg/L) suggesting that ER agonist activity is the more relevant endocrine effect. However, histopathological endpoints are generally unsuitable for use in PNEC derivation because the changes are difficult to qualify in terms of severity and direct link to a population-relevant outcome. In the second LAGDA, there were no effects on plasma VTG or thyroxine at 38 µg/L. The study report did not include effects on organ histopathology (ABC labs, 2012). The LAGDA validation (US EPA, 2013) included a third study with OP but there is insufficient reporting of data to assess the quality of the study, however the results from all three tests are similar with no adverse effects on growth or development.

There is also a result for OP in a 151 day LAGDA style study with *X. tropicalis* (Porter *et al*., 2011). In this study there were also no effects on growth, development, or the phenotypic sex ratio up to the highest test concentration of 36 µg/L. Sex ratio based on gross examination of the gonad tissue indicated significantly more females at 11 µg/L compared with the control, however the biological relevance of this observation is unclear because it did not follow a concentration response (no effect at 36 µg/L). Mechanistic endpoints for ER agonist activity (VTG induction and for oviduct development in males) were affected at 36 µg/L.

There is evidence for ER agonist activity with induction of VTG observed in the three LAGDA studies from 6.3 – 54.2 µg/L, and effects on the gonads and oviducts within the same range (OECD, 2013). Evidence for thyroid activity in amphibians is weak; reduced serum thyroxine and minor histopathological alterations of the thyroid at ≥ 25 µg/L in one of three LAGDA (Haselman *et al*., 2015) and accelerated development in an AMA at 39.8 µg/L but without histological correlate (Kamel *et al*., 2022).

It is difficult to determine the valid NOEC values for amphibians, due to the lack of population-relevant effects on growth and reproduction at the highest test concentrations in the relevant OECD CF Level 4 studies, which were a NOEC > 38 µg/L for *X. laevis* or > 36 µg/L for *X. tropicalis*. These studies did not, however include a reproductive phase which may have resulted in an effect at a lower test concentration based on the mechanistic evidence for ER agonist activity. Usually, the result from a Level 3 study would not be used for PNEC derivation but the result from the AMA is considered valid for use because there were four test concentrations, acceptable spacing between the test concentrations, and the test concentration range was higher than in the LAGDA style studies (up to measured 103 µg/L). Significantly increased development rate at 39.8 and 103 µg/L was concentration dependent and considered to be population-relevant. The NOEC from this study was 13.5 µg/L.

4.1.3 Invertebrates

The available data for invertebrates are given in Table 51. There were two reproduction studies with *Daphnia magna* with supporting analytical chemistry and following a test design similar to OECD TG 211 (IUCLID, 2000; Hüls, 1992). The studies are both from secondary sources: The IUCLID (2000) study is cited as reliable in EA (2005), while the Hüls (1992) study is considered reliable according to the ECHA registration dossier. The studies both report inhibition of reproduction, with LOECs of 120 µg/L (IUCLID, 2000), and 100 µg/L (Hüls, 1992), and NOECs of 62 and 30 µg/L, respectively.

There were also reliable studies with the mollusc *Potampyrgus antipodarum*. Six studies were conducted to assess the potential for OP to be used as a validation substance for stimulation of reproduction (UBA, 2017). The measured concentrations were similar between studies and therefore the arithmetic mean of the six studies is used for reporting of the data. Stimulation in reproduction was observed in three of six studies with a LOEC of 1.7 µg/L and a NOEC of 0.5 µg/L. Stimulation of reproduction was not generally observed at the higher test concentrations of 21 and 72 µg/L (in one of the six studies only). Inhibition of reproduction was observed in two of six studies at the highest test concentration of 72 µg/L with a NOEC of 21 µg/L. In studies where reproduction was higher (typically during summer months) it was not possible to detect stimulation in reproduction. It is not possible to attribute a stimulation in reproduction to an endocrine MoA (see Crane *et al*., 2022) or even a toxicological mechanism (stimulation in reproduction may be a secondary effect due to the test item causing microbial growth and increasing food availability). The effect however has been reported in other studies with OP (Jobling *et al.,* 2003) although this study was assessed as not reliable according to ECHA (2011). The overall conclusion in the UBA report was that OP was not suitable for use as a validation substance for stimulation of reproduction, and the OECD *Potamopyrgus antipodarum* Reproduction Test was not validated for effects via stimulation of reproduction (OECD, 2016). Stimulation of reproduction is not considered a population-relevant endpoint in the absence of other information to indicate adversity (e.g. smaller offspring, or reduced survival of parents or offspring). A NOEC of 21 µg/L based on inhibition of reproduction is proposed for use in the present assessment. This is considered worst case as the effect was only observed in two of the six studies.

The only other long-term study with invertebrates which was considered reliable according to ECHA (2011) is a reproduction study with the marine copepod *Tigriopus japonicas* (Marcial *et al*., 2003), however this study did not include analytical verification of the test solutions and therefore is not considered relevant for use in the present assessment.

4.2 PNEC derivation with consideration of ED endpoints

In this section aquatic PNECs for OP have been derived using reliable and relevant data in OECD CF Level 4 and 5 studies undertaken with invertebrates and fish, and a Level 3 study with an amphibian. The result from a Level 3 amphibian metamorphosis assay (AMA) was included because the test design includes sensitive juvenile life stages.

The dataset for OP covers eight species (two invertebrates, four fish and two amphibians) with more reliable data for vertebrate species compared with NP.

4.2.1 Suitability of the dataset for derivation of a PNEC for an EDC and data gap analysis

The reliable and relevant dataset has been assessed against the SETAC Pellston Workshop® criteria for deriving a PNEC with an EDC (Matthiessen *et al*., 2017).

1. *Have the most appropriate taxa been tested (with relevant endpoints)?*

Based on the MoA as an ER agonist, vertebrates would be expected to be the most sensitive species to OP. There are six vertebrate species included in the dataset, four fish and two amphibians. The fish are all standard OECD test guideline species but represent different families. There are no marine species or species with alternative reproductive strategies (e.g. live bearers). Two of the vertebrate species are amphibians which were both less sensitive compared with fish (although all vertebrate NOECs were within a factor of 6). This is probably because the amphibian tests focus on developmental effects via the thyroid rather than reproductive endpoints which are more relevant for effects via EAS modalities. The LAGDA studies do however cover development and sex ratio and therefore could be considered equivalent to the FSDT. There are no validated studies covering reproductive or multigeneration effects in amphibians.

Nevertheless, the availability of reliable and relevant data from six valid OECD CF Level 4 or 5 vertebrate studies is considered to be a relatively robust dataset for PNEC derivation given the focus on long-term vertebrate testing.

2. *Have sensitive life stages, or the entire life cycle, been tested (with relevant endpoints)?*

There are relevant and reliable Level 4 FSDTs with three fish species (OECD, 2012). The FSDT has an endpoint for effects on sex ratio, with an increase in the number of females being diagnostic for ER agonist activity. A shift in sex ratio towards females was observed for *D. rerio*, *O. latipes* and *G. aculeatus*. There were also reliable and relevant Level 5 studies with two fish species (*D. rerio*, and *O. latipes*), studies which include endpoints for development and reproduction (Wenzel *et al*., 2001; Ministry of the Environment of Japan, 2002; Flynn *et al*., 2017).

The most sensitive result in the dataset was a NOEC of 6.1 µg/L from a Level 4 FELS test with *On. mykiss*. There is therefore a possible concern that the result from a Level 5 test with this species may result in a lower NOEC.

In addition, there are two Level 4 LAGDA style studies with two amphibian species (Porter *et al*., 2011; US EPA, 2013) which also cover relevant developmental endpoints.

3. *Have delayed and multigenerational (i.e. latent) effects been considered?*

For OP, the results from FSDT studies with *D. rerio* were more sensitive for identifying adverse effects (increase in females), compared to a life cycle test which measured effects in an F1 generation. Equally the LOEC for sex ratio for *O. latipes* in an FSDT was more sensitive than the effects reported in a MEOGRT or FLCTT with this species, although the statistical power in the MEOGRT test design to detect small changes in sex ratio may be lower (OECD, 2015). Nevertheless, there is no suggestion that toxicity to OP increases over multiple generations.

4. *Do NMDRs or other unusual temporal patterns of toxicity affect the ability to predict reliable no-adverse-effect levels?*

For OP there is no convincing evidence for an NMDR relationship for the key indicators for endocrine activity (e.g. VTG induction, presence or testis-ova, SSC or gonad histopathology) or effects on EAS mediated endpoints (sex ratio).

An increase in body length was observed in a reliable MEOGRT (US EPA) with OP with a NOEC of 6.2 µg/L (Flynn *et al*., 2017), and an increase in body weight in the second MEOGRT (NIES) based on a nominal NOEC of 12.5 µg/L. It cannot be ruled out that these effects may be related to the endocrine mechanism (e.g. related to feminisation) and therefore they were considered relevant. However, these effects were accounted for alongside more traditional endpoints for adversity (i.e. reduced growth) and were generally not the most sensitive endpoint.

NMDR have also been observed in some studies with molluscs, most notably with *P. antipodarum* (Jobling *et al*., 2003; UBA, 2017). The mechanistic basis for this effect is unknown as there is a lack of functional role for steroid hormones in invertebrates (Crane *et al*., 2022). It is also not known whether the effect is adverse or potentially a secondary response to stimulation of microbial growth and an increase in food availability. Nevertheless, a NOEC of 0.5 µg/L for a stimulation in reproduction in molluscs (UBA, 2017) is lower compared with oestrogenic effects in fish and therefore is considered as supporting information.

5. *Does a threshold for adverse endocrine-mediated effects exist?*

Figure 3 shows the range of NOEC values for endocrine population-relevant (reproduction, hatching and sex ratio) and “biomarker” endpoints (VTG induction, gonad histopathology, SSC, and presence of testis-ova) from 25 reliable OECD CF Level 3, 4 or 5 fish studies (note that most of these studies were not used for PNEC derivation, either because they did not meet relevance criteria, i.e. Level 3 fish studies were excluded, or because there were multiple studies with the same species). All NOEC values fall between 1 and 100 µg/L regardless of study type. The lowest “biomarker” NOEC was 1 µg/L based on male VTG induction in a FSTRA style study with *On. mykiss* (Routledge *et al*., 1998). These data would appear to demonstrate a threshold level for endocrine effects in fish (as the most sensitive taxonomic group) with NOECs being above 1 µg/L.

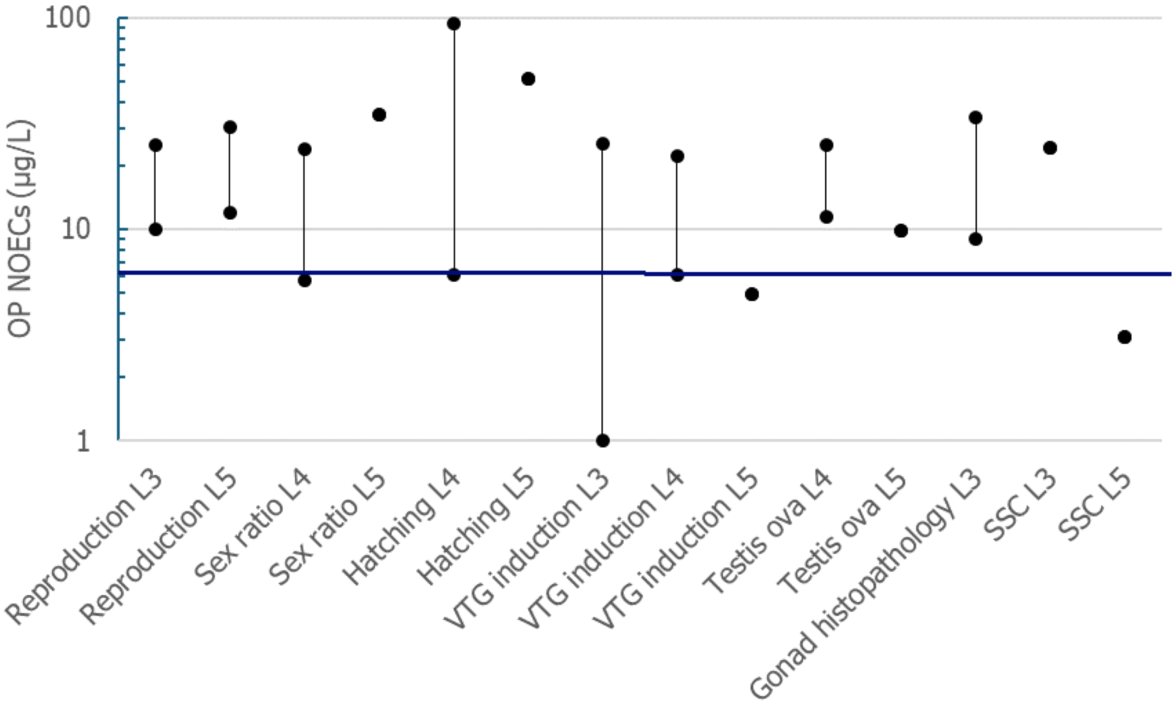


Figure 3 Range of NOEC values from all reliable OECD CF Level 3, 4 and 5 fish studies with OP. The blue line indicates the lowest population level NOEC of 6.1 µg/L from a FELS study for *On. mykiss* (ABC labs, 1984, cited in EA, 2005 and ECHA, 2011).

The most sensitive population-relevant NOEC for OP is 6.1 µg/L for growth of *On. mykiss* (ABC labs, 1984, cited in EA, 2005; ECHA, 2011)). The repeatability of the response of fish to OP has been demonstrated in validation studies for the FSTRA (US EPA, 2007), FSDT (OECD, 2012), LAGDA (US EPA, 2013), and MEOGRT (Flynn *et al.*, 2017).

Overall, the dataset for OP is considered to satisfy the SETAC Pellston Workshop® criteria for vertebrates.

4.2.2 Deterministic PNEC assessment for OP

The reliable dataset for fish and amphibians includes multiple studies designed to investigate the oestrogenic MoA. Nevertheless, the most sensitive result is a NOEC of 6.1 µg/L from a FELS study with *On. mykiss* (ABC labs, 1984, cited in EA, 2005; ECHA, 2011). An AF of 10 has been applied to this NOEC to derive a deterministic PNEC following standard approaches as follows:

OP deterministic PNEC = 6.1/10 = 0.61 µg/L

A case could be made for the application of additional AF based on the substance having a confirmed endocrine MoA. However, the endocrine mechanism is potentially already accounted because the dataset includes multigeneration fish studies (although the lowest NOEC is from a Level 4 FELS test). The PNEC of 0.61 µg/L is also protective of the lowest “biomarker” NOEC of 1 µg/L for male VTG induction in *On. mykiss* (Routledge *et al.*, 1998). In addition, it would be protective of the LOEC of 1.7 µg/L for stimulation of reproduction in *P. antipodarum* (UBA, 2017) that has been omitted as a key study due to the uncertainties described above.

4.2.3 Probabilistic assessment for OP

A probabilistic assessment has been conducted using the reliable and relevant Level 4 and 5 CF data for OP, comprising eight species (two invertebrate, four fish and two amphibian). In addition, the probabilistic assessment was conducted with or without the invertebrate data.

Table 6 summarises the reliable and relevant NOECs for use in constructing an SSD for OP. The NOEC values range from 6.1 µg/L for *On. mykiss* in a FELS test, up to 36 µg/L for *X. tropicalis* in a LAGDA.

Table 6 Values for selected for deriving an HC_5 for OP

Taxon	Species	Endpoint (NOEC)	Value (µg/L)	Value selected for SSD	References
Crustaceans	*Daphnia magna*	Reproduction	30	30	Hüls, 1992
Mollusc	*Potampyrgus antipodarum*	Reproduction	21	21	UBA, 2017
Amphibians	*Xenopus laevis*	Accelerated development	13.5	13.5	Kamel *et al.*, 2022
		Growth and development	38		ABC, Labs, 2012
	Xenopus tropicalis	Growth and development	36	36	Porter *et al.*, 2011
Chordata	*Oryzias latipes*	Sex ratio	23.6	6.2	OECD, 2012, Lab 9 / Seki *et al.*, 2003
		Increased growth	6.2		Flynn *et al.*, (2017)

Taxon	Species	Endpoint (NOEC)	Value (µg/L)	Value selected for SSD	References
	Danio rerio	Reproduction	12	7.4	Wenzel *et al.*, 2001
		Sex ratio	7.4		OECD, 2012
	Gasterostreus aculeatus	Hatching rate	22.2	22.2	OECD, 2012
	Oncorhynchus mykiss	Growth	6.1	6.1	ABC labs, 1984, cited in EA, 2005 and ECHA 2011

The SSD for the reliable and relevant OP dataset is shown in Figure 4.

HC_5 values were derived using ETX version 2.3 or the weighted average method using ssdtools with and without invertebrate data. In both cases the HC_5 without invertebrates was more sensitive and therefore used for the analysis (Table 7).

The SSDs for the reliable and relevant OP dataset derived using ssdtools are shown in Figure 4. The following models showed the best fits to the data: gamma, llogis, lgumbel, lnorm and Weibull.

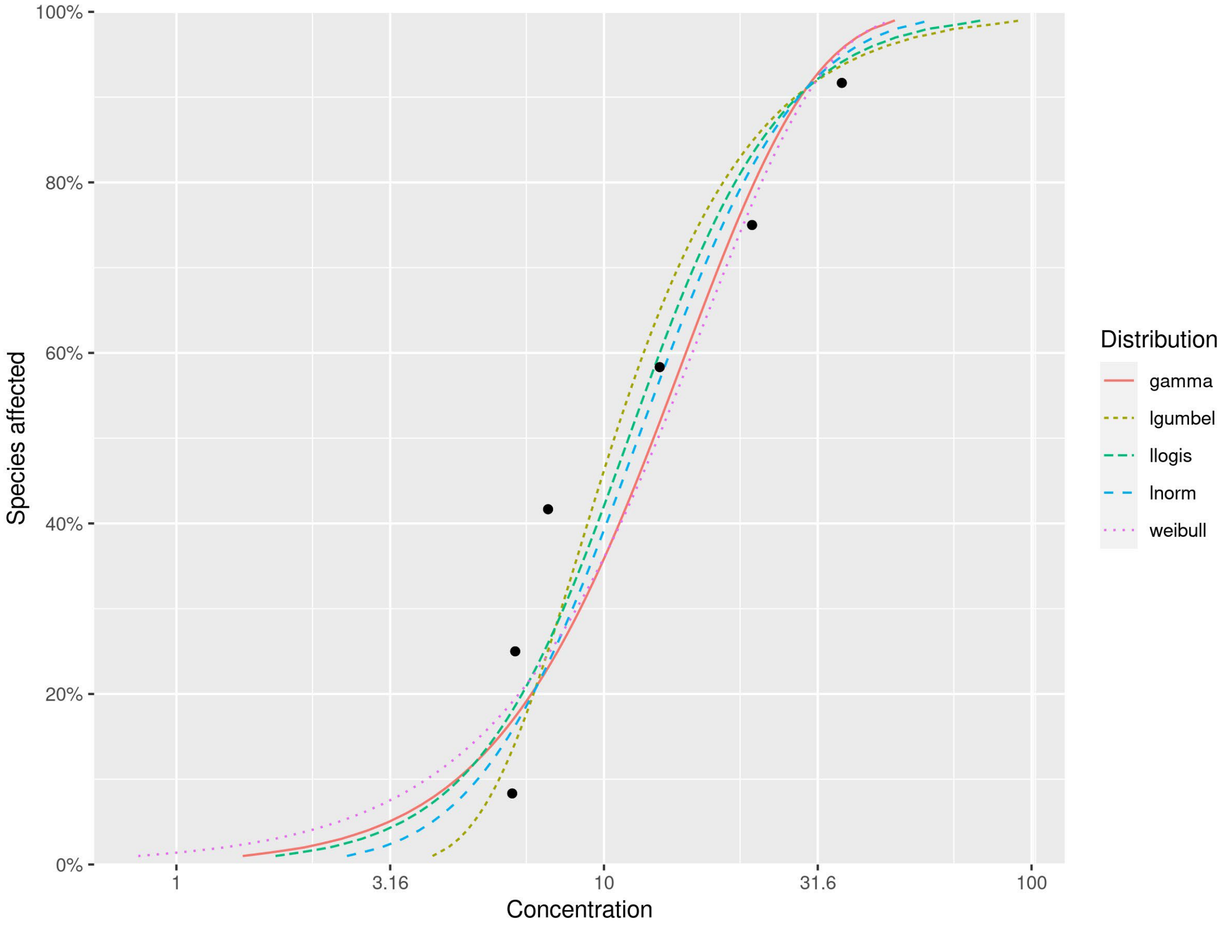


Figure 4 SSDs models from ssdtools for deriving an aquatic HC_5 for OP

Table 7 HC_5 values derived using the different distribution models

Distribution	Inverts	HC5 (µg/L)	Confidence Intervals (µg/L)	AICc[1]	Weight	Software
Log normal	Y	4.21	1.47 – 7.27	-	-	ETX v. 2.3
	N	3.30	0.78 – 6.3	-	-	
Average	Y	4.42	2.10 – 10.6	-	-	ssdtools v. 1.0.2
	N	**3.87**	**2.04 – 6.56**	-	-	
gamma	N	3.14	0.817 – 10.4	7.83	0.155	
lgumbel	N	4.96	3.48 – 9.41	6.3	0.334	
llogis	N	3.38	1.24 – 8.65	7.74	0.163	
lnorm[2]	N	3.97	2.14 – 9.67	7.13	0.22	
weibull	N	2.39	0.614 – 0.614	8.22	0.128	

Value highlighted in bold used for PNEC derivation
Andersen Darling statistic was not calculated in ssdtools due to the small dataset. It was accepted for the ETX calculations.
[1] Akaike's Information Criterion corrected for sample size (AICc)
[2] Note that a small differences in model results (less than a factor of 2) for different estimation packages is expected due to different estimation strategies (Fox *et al.*, 2021)

All the HC_5 values for OP were relatively similar due to the relatively small range of NOEC values for OP (Table 7). Due to the small sample size not all the goodness-of-fit statistics, including the Andersen Darling statistic, could be applied to the models. The weighted average HC_5 from ssdtools was used for the SSD. Unlike the probabilistic HC_5 value for NP, the probabilistic HC_5 for OP would be protective based on the NOECs in the OP dataset and therefore the approach appears to be acceptable. A default AF of 5 was applied to the data based on the dataset generally meeting the SETAC Pellston Workshop® criteria.

Probabilistic OP PNEC = 3.87/5 = 0.77 µg/L

The probabilistic PNEC is less conservative compared with the deterministic PNEC of 0.61 µg/L but would still be protective of biomarker effects and stimulation of reproduction in molluscs. There are still, however, the following concerns:

- The SSD dataset is small, with only four fish species; and

- The lowest NOEC was not from a multigeneration study and therefore the 'actual' NOEC for *On. mykiss* may be less than 6.1 µg/L.

The deterministic PNEC of 0.61 µg/L is similar to a PNEC of 0.63 µg/L used in the EU REACH registration dossier for OP (ECHA, 2023). Further discussion of the suitability of using SSDs for EDCs and the application of an additional AF is discussed in Section 77.

4.2.4 Combining the OP and NP datasets

Due to the limited species coverage, it may be illustrative to consider combining information from both NP and OP into a single data set. Even though their potencies differ slightly, they have a similar MoA as ER agonists and are structural analogues. The main difference between the OP dataset and the combined dataset is the inclusion of a NOEC of 7.4 µg/L for *P. promelas* (Ward and Boeri, 1991a) and for *O. latipes* the use of the NOEC of 0.64 µg/L from the MEOGRT with NP (Watanabe *et al*., 2017) rather than at NOEC of 6.2 µg/L from an MEOGRT with OP (Flynn *et al*., 2017).

The HC_5 from the combined dataset is 1.05 µg/L using the weighted average method from ssdtools. This HC_5 value is not suitable for use for deriving a PNEC for NP as it is not protective of the effects of NP in the MEOGRT. It is also not justified for use for OP as the HC_5 value for the combined dataset is influenced by the result for NP with *O. latipes* in a MEOGRT when the OP dataset already includes OP-specific NOECs for *O. latipes* from both FSDT and MEOGRTs which are more relevant. The combined dataset is therefore not considered further in the present assessment.

4.3 Summary and conclusion

A dataset of relevant and reliable studies suitable for use in deriving a PNEC for OP has been compiled. There were valid OECD CF Level 4 and 5 NOECs available for eight species (two invertebrate, two amphibians, and four fish). The most sensitive NOEC in the dataset was reduced growth of *On. mykiss* at 6.1 µg/L (ABC labs, 1984, cited in EA, 2005; ECHA, 2011).

Although there are a relatively small number of species, the dataset met the SETAC Pellston Workshop® criteria.

- Fish were the most sensitive taxa based on data from a range of test methods (FSDT, FLCTT, and MEOGRT) designed to cover sensitive life stages and endpoints known to be affected by the ER agonist MoA (sex ratio and reproduction). There were also LAGDA studies with amphibians which cover developmental endpoints.
- A MEOGRT study is available to cover multigeneration effects.
- The main evidence for NMDR was for an increase in growth in fish. This effect in fish is probably due to the endocrine mechanism and is accounted for as an adverse population level effect in the assessment. Increased reproduction in molluscs is not considered relevant for the ER agonist MoA because based on

present biological knowledge there is no known functional role for the ER in invertebrates. The endpoint for increased reproduction in molluscs is therefore used as supporting information only. Key indicators for endocrine activity (e.g. VTG induction, presence or testis-ova, SSC or gonad histopathology) or effects on EAS mediated endpoints (sex ratio) were typically concentration related.

- There appears to be an ETNC for OP with effects on fish (mechanistic or population-relevant) typically observed above 5 µg/L. The lowest NOEC for a biomarker endpoint was 1 µg/L based on male VTG induction in *On. mykiss* (Routledge *et al*., 1998).

The most sensitive NOEC for OP was 6.1 µg/L from an OECD CF Level 4 study (e.g. reduced growth in a FELS test with *On. mykiss*) rather than a Level 5 test incorporating a second generation, although the MEOGRT showed similar sensitivity to the FELS result (NOEC of 6.2 µg/L based on increase in body length). OP has been used in validation studies of OECD fish and amphibian tests and therefore the repeatability of the results has been demonstrated, improving confidence that an ETNC can be determined. There are no validated studies with amphibians that cover reproduction endpoints or multigeneration effects, but the available data based on effects on sex ratio in the LAGDA do not suggest that amphibians would be more sensitive to fish to OP.

Most of the fish data was for ‘standard’ small test species, which are rapid breeders kept in warm water. The exception was the lowest NOEC in the dataset which was for *On. mykiss* based on a FELS test. Therefore, there is a data gap for EAS mediated effects in longer-lived cold-water species. Other data gaps are for effects on marine species and for fish with alternative reproductive strategies (e.g. live bearers). There were only two invertebrate species in the reliable and relevant OP dataset, but invertebrates were less sensitive to OP compared with fish, as expected based on the oestrogenic MoA, and were therefore not used for the probabilistic risk assessment.

The deterministic PNEC for OP was 0.61 µg/L derived by applying an AF of 10 to the lowest population-relevant NOEC of 6.1 µg/L. The probabilistic PNEC was 0.77 µg/L, derived by applying an AF of 5 to a weighted average HC_5 of 3.83 µg/L using the dataset for vertebrates only.

Consideration was given to combining the NP and OP datasets to provide a PNEC based on a larger amount of data. This was ruled out because although the main benefit was to increase the dataset for NP, fish are more sensitive to NP compared with OP meaning the result was not sufficiently precautionary.

5 PNEC derivation: EE2

5.1 Long-term aquatic toxicity data

As part of the SETAC Pellston Workshop® on "Ecotoxicological Hazard and Risk Assessment Approaches for Endocrine-Active Substances" Matthiessen *et al*. (2017) derived an HC_5 of 0.047 ng/L for EE2, using only relevant data for fish. The authors concluded that there were no data gaps for EE2 because of the extensive dataset on effects in fish, and its well-established MoA. To ensure that no other potentially relevant studies (not considered by Matthiessen *et al*., 2017) were available for the current assessment, additional EE2 assessments published by the European Commission (EC, 2002) and the JRC (as part of EQS derivation) (JRC, 2022) were also reviewed. In JRC (2022), an SSD was generated using aquatic taxa expected to be most sensitive to EE2, namely fish *and* amphibians.

In some cases, a re-evaluation of the reliability of studies used in Matthiessen *et al*. (2017) was undertaken where the result was considered to potentially impact the outcome of the assessment (i.e. those which might result in a change in the lowest NOEC for a particular species), in addition to the reliability assessments carried out for published data not included in Matthiessen *et al*. (2017). This ensured consistency with the approach taken for NP and OP. For studies where a clear NOEC was not derived (i.e. unbounded, with effects reported at the lowest exposure concentration), and which therefore did not meet the relevance criteria as defined in Section 15 of the current assessment, a LOEC/2 approach was applied if the effect was between 10 and 20% of the control, in accordance with guidance given in ECHA (2008). However, in Matthiessen *et al*. (2017) and in JRC (2022) this approach (LOEC/2) was also used for some studies where the LOEC was > 20% of the control result.

Detailed information on the substance identity and physicochemical properties of EE2 is given in Appendix 150.

5.1.1 Updates to the dataset

The fish and amphibian data used to construct an SSD in JRC (2022) and Matthiessen *et al*. (2017), and the further data identified in the present assessment, are described in the sections below. Data from invertebrates are not included in this assessment for EE2 given the amount of data for, and its known high potency in, aquatic vertebrates.

A literature search was conducted by the EA covering the period from 2016 to 3 June 2021 to identify new published data for EE2. Twenty-three papers containing potentially relevant data were identified. Four papers were not considered further, as two involved exposures to bivalves and two were acute exposures. Two were review papers (Bosker *et al*., 2017; Rutherford *et al*., 2020) which were used to cross check that no potentially relevant studies had been overlooked. The remaining 17 papers were assessed for reliability and relevance, according to the criteria and approaches detailed in Section 15.

Only one new study using the Ten-spotted Livebearer *Cnesterodon decemmaculatus* (Young *et al*., 2020) was considered relevant and reliable for PNEC derivation.

5.1.2 Amphibians

JRC (2022) included NOECs from amphibian studies with *Xenopus tropicalis* and *Rana temporaria* (Pettersson and Berg, 2007). Amphibian data were not considered by Matthiessen *et al*. (2017). The Pettersson and Berg (2007) study was reviewed in the present assessment and the *X. tropicalis* data considered to be reliable and relevant for inclusion in the assessment. *X. tropicalis* tadpoles were exposed for approximately 32 days (from Nieuwkoop and Faber stage 47 to completion of metamorphosis). Endpoints included survival, sex ratio and time to metamorphosis. The most sensitive endpoint was sex ratio and time to metamorphosis with a bounded NOEC of 1.77 ng/L. The *R. temporaria* dataset was not considered relevant or reliable, as egg masses were collected from the wild and may have been previously exposed to the test substance (or affected by viruses or parasites, etc.), and only two concentrations were tested.

5.1.3 Fish

Cyprinidae (*Danio rerio)*

A modified OECD CF Level 4 study (FSDT) with the zebrafish, *D. rerio* (Maack and Segner, 2004), is available. Female bias in the sex ratio was determined to be the most sensitive endpoint with a statistically significant change reported at the lowest exposure concentration of 1.67 ng/L (NOEC < 1.67 ng/L). Five life cycle studies were also available (Wenzel *et al*., 2001; Nash *et al*., 2004; Schäfers *et al*., 2007; Soares *et al*., 2009; Larsen *et al*., 2008). A sixth life cycle study by Segner *et al*. (2003) was deemed to be the same data as that reported in Wenzel *et al*. (2001) and these studies have been considered together in Table 54 (referred to as Wenzel *et al*. (2001) for the remainder of this report). In three of the life cycle studies, bounded NOECs could be derived; Larsen *et al*. (2008) based on sex ratio (0.05 ng/L); Wenzel *et al*. (2001) based on F1 growth (0.1 ng/L); and Schäfers *et al*. (2007) based on F0 growth/reproduction/fertilisation success (0.31 ng/L). In two of the life cycle studies, which were both considered reliable, the NOEC was unbounded (Nash *et al*., 2004; Soares *et al*., 2009), as effects on F1 fertility and F1 growth/mortality occurred at the lowest exposure concentrations of 0.5 and 0.19 ng/L, respectively. It was not possible to apply the LOEC/2 approach to estimate the NOEC as the data did not meet the relevant ECHA (2008) criteria for such an estimation. The lowest bounded NOEC for *D. rerio* is 0.05 ng/L, based on sex ratio (Larsen *et al*., 2008). In this study there were 69% males and 31% females in the control, with a significant reduction to 59% male at 0.5 ng/L and only 5% male at 5 ng/L.

Matthiessen *et al*. (2017) used a geometric mean of 0.21 ng/L from the NOECs for fertility in Wenzel *et al*. (2001), Schäfers *et al*. (2007) and Larsen *et al*. (2008). JRC (2022) used the NOEC of 0.3 ng/L from Wenzel *et al*. (2001), based on reproduction (the Larsen *et al*., 2008 study is not mentioned in JRC, 2022). The key biomarkers for oestrogenic activity in these studies all followed a concentration response with effects on male and female VTG induction in F0 adults with LOECs of 0.5 and 4.5 ng/L, respectively (Nash *et al*., 2004) and a decrease in male SSC with a LOEC of 0.05 ng/L (Larsen *et al*., 2008). There was no threshold value for effects on this biomarker, with effects generally reported at the lowest exposure concentrations.

Cyprinidae (*Pimephales promelas*)

Two reliable Level 5 multigenerational studies with the fathead minnow *P. promelas* are available (Parrott and Blunt, 2005; Länge *et al*., 2001). They are of similar study design, with the lowest (bounded) NOEC from Länge *et al*. (2001) of 0.76 ng/L based on F0 growth. In Parrott and Blunt (2005) a clear NOEC could not be determined, as a significant decrease in fertilisation success and an increase in female bias in the sex ratio was reported at the lowest nominal concentration of 0.32 ng/L. In Matthiessen *et al*. (2017) a geomean value of 0.86 ng/L was used in their SSD, based on the overall NOEC of 0.76 ng/L (Länge *et al*., 2001) and a value of 0.96 ng/L from Parrott and Blunt, 2005 (although this was not the lowest NOEC from the study). Since the NOECs represented different endpoints it was not considered viable to use a geometric mean in the present assessment. It should also be noted that in Parrott and Blunt (2005) the key biomarker for oestrogenic activity, male SSC (LOEC 0.96 ng/L) and female ovipositor index (LOEC 3.54 ng/L) demonstrated a concentration response with a decrease and increase observed, respectively. These LOEC values are greater than the NOEC of 0.76 ng/L used in our assessment. In Länge *et al*. (2001) there was a concentration response increase in the presence of testis-ova (LOEC 2.8 ng/L), except at the highest concentration where none were found, however at this concentration almost all the fish (96%) were female.

Cyprinidae (*Gobiocypris rarus*)

One Level 5 study was available with the Chinese Rare Minnow, *G. rarus* (Zha *et al*., 2008). This multigenerational study was initiated with embryos and exposure was over two generations. Exposure of a third generation was planned but this was not undertaken due to complete spawning failure in the F2 generation at all EE2 exposure concentrations. A clear NOEC could not be determined in this study as effects were reported at the lowest exposure concentration of 0.18 ng/L (i.e. decrease in F0 and F1 reproduction, fertility and an increase in female bias in the sex ratio). Although the percentage changes were > 20% (i.e. 42% and 35% for F0 reproduction and fertilisation rates, respectively), the LOEC/2 approach was used in Matthiessen *et al*. (2017) and JRC (2022) to estimate a NOEC of 0.09 ng/L for this study. Biomarkers for oestrogenic activity all followed a concentration response with effects on male and female VTG induction in F0 and F1 adults observed at exposure concentrations of ≥ 0.18 ng/L in males and ≥ 0.961 ng/L in females and a LOEC for induction of testis-ova at 0.18 ng/L.

Cyprinidae (*Rutilus rutilus*)

The one Level 4 study available with the Roach, *R. rutilus* (Lange *et al*., 2008) included in the JRC (2022) assessment was considered unreliable for the present assessment. The results were based on only two test concentrations, so it was not possible to establish a concentration response. Measured concentrations were only 30 and 40% of nominal. This study was not considered in Matthiessen *et al*. (2017). It should also be noted that for one of the key biomarkers for oestrogenic activity, VTG, an increase was only reported at the highest concentration (LOEC = 4 ng/L).

Salmonidae (*Oncorhynchus mykiss*)

Three Level 4 studies were available for rainbow trout, *On. mykiss* (Schultz *et al*., 2003; Brown *et al*., 2007; Depiereux *et al*., 2014). The study by Brown *et al*. (2007) was the most sensitive and was initiated with adult males exposed for 56 days before semen was collected and used to fertilise unexposed eggs. Survival in the resulting F1 embryos was determined at 19 days post fertilisation. A clear NOEC could not be determined in this study as a significant decrease in F1 survival was reported at the lowest exposure concentration of 0.8 ng/L. However, the percentage change was approximately 20%, so the data met the ECHA (2008) requirements for use of the LOEC/2 approach to estimate a NOEC of 0.4 ng/L. This approach to derive a NOEC was also undertaken in the Matthiessen *et al*. (2017) assessment. JRC (2022) used a higher NOEC of 8 ng/L, calculated from the LOEC/2 (15.6 ng/L) based on reproduction (Schultz *et al*., 2003), however these data did not meet the ECHA (2008) requirements for this approach. A higher (unbounded) NOEC of 10 ng/L was reported by Depiereux *et al*. (2014) based on sex ratio (sex reversal). In Brown *et al*. (2007) plasma VTG concentrations, a key biomarker for oestrogenic activity, in male fish followed a concentration response with an increase observed at all measured concentrations (LOEC = 0.8 ng/L). In Depiereux *et al*. (2014) there was a concentration response for the feminisation of the gonads (i.e. presence of intersex gonads and an oviduct-like structure) with a LOEC of ≥ 10 ng/L and at the highest two exposure concentrations (1 620 and 9 880 ng/L) all fish were identified as female.

Cyprinodontidae (*Cyprinodon variegatus*)

One Level 5 study was available with the sheepshead minnow, *C. variegatus*, (Zillioux *et al*., 2001). This partial multigenerational study was initiated with sub-adults and exposure was continued over the early life-stages (approximately seven days) of the F1 generation. The NOEC from this study was 18.1 ng/L based on a decrease in reproduction and hatching success. The NOEC of 18.1 ng/L was also the NOEC that was used in JRC (2022), but this study was not considered in Matthiessen *et al*. (2017). The key biomarker for oestrogenic activity, induction of testis-ova, was reported in males at all exposure concentrations, LOEC ≥ 18.1 ng/L.

Poeciliidae (*Poecilia reticulata*)

One Level 5 study was available with the guppy, *P. reticulata*, with different endpoints reported in separate papers (Nielsen and Baartrup, 2006; Kristensen *et al*., 2005, respectively). Both describe a study initiated with newborn fish, with exposure continued until adulthood. The NOEC is the same from both studies, 44 ng/L, based on an increase in female bias in the sex ratio, an increase in growth (male body weight) and biomarker endpoints of decreasing body colouration intensity and alteration in sperm quality. The NOEC of 44 ng/L was used by both Matthiessen *et al*. (2017) and JRC (2022). There was also a concentration response regarding colouration (male SSC), a biomarker for oestrogenic activity, with a LOEC of 112 ng/L (Nielsen and Baartrup, 2006; Kristensen *et al*., 2005).

Poeciliidae (*Cnesterodon decemmaculatus*)

A Level 4 modified FSDT study with the Ten-spotted Livebearer, *C. decemmaculatus* was identified in the EA literature search (Young *et al*., 2020). Newborn *C. decemmaculatus*, were exposed for 90 days to three measured concentrations of EE2 (32.3, 99.1 and 296.2 ng/L). Endpoints included growth, survival, sex ratio and SSC (gonopodium development). A clear NOEC (bounded) could not be determined in this study as an increase in female bias in the sex ratio was reported at the lowest exposure concentration of 32.3 ng/L. The percentage change (42.9%) for sex ratio, compared to the control, did not meet the ECHA (2008) guidance criteria for estimating the NOEC from the LOEC/2. Sex ratio changes at the two higher concentrations were considered to have been impacted by systemic toxicity. The mechanistic endpoint for SSC (delay in gonopodium development) for oestrogenic activity followed a concentration response and was observed at all concentrations tested, LOEC ≥ 32.3 ng/L.

Adrianichthyidae (*Oryzias latipes*)

One Level 5 study with the Japanese medaka, *O. latipes* (Balch *et al*., 2004), included in the Matthiessen *et al*. (2017), assessment was considered unreliable for PNEC derivation following an evaluation undertaken for the present assessment, since the results were based on nominal concentrations with the exposure media only being renewed on three days each week. Matthiessen *et al*. (2017) also considered this study to be borderline reliable (Klimisch 2 or 3). This study was listed in the JRC (2022) but was not included in their SSD, and a NOEC from a Level 4 study (Scholz and Gutzeit, 2000) was included instead. Unfortunately, the Scholz and Gutzeit (2000) study did not have supporting analytical verification and therefore is not included in the present assessment. A key biomarker for oestrogenic activity, induction of testis-ova in males followed a concentration response with a LOEC of 2.0 ng/L (Balch *et al*., 2004). In Scholz and Gutzeit (2000), female GSI followed a concentration response decrease (LOEC 10 ng/L) and at 100 ng/L all XY-males developed into females.

Gobiidae (*Pomatoschistus minutus*)

The one Level 5 study available with the Sand Goby *P. minutus* (Robinson *et al*., 2003), was considered by Matthiessen *et al*. (2017) and the present assessment to be unreliable as results were based on a single test concentration, which was applied as a positive control in the testing of other substances. However, Matthiessen *et al*. (2017) included this value in their SSD, with the caveat that this study was considered unreliable (Klimisch 3). This study is also listed in JRC (2022) but not included in their SSD. Regarding the biomarkers for oestrogenic activity, a reduction in GSI and change in SSC were also reported at the single concentration tested (LOEC 6 ng/L).

Fundulidae (*Fundulus heteroclitus*)

The one Level 4 study available with the Mummichog *F. heteroclitus* (Peters *et al*., 2010) was also considered unreliable in the current assessment. It was, however, considered reliable by Matthiessen *et al*. (2017) and included in their SSD. Exposure concentrations were not measured in the tanks during the exposure as insufficient water volumes were collected but were measured in tanks used in a subsequent exposure (control and 100 ng/L). At 100 ng/L a loss of 88% was recorded over 12 hours, indicating that the dosing system was unable to maintain stable exposure conditions at the highest exposure concentration. As the exposure concentrations at the lower concentrations were not measured, it remains unknown if the technical issues affecting the highest concentration extended to the lower doses. Regarding the biomarkers for oestrogenic activity there was a significant increase in liver VTG in females, in female SSC for juveniles and in male and female gonad histopathology at 100 ng/L (LOECs = 100 ng/L (nominal)). However, these findings are equivocal since they are only based on only two to four fish per sex, per concentration.

5.2 PNEC derivation

5.2.1 Dataset evaluation

Applying the relevance criteria given in Section 15, the results from studies at OECD CF Level 3 were not considered further for PNEC derivation, because they are designed as screening assays for mechanistic effects and typically have wide spacing of test concentrations and relatively short test durations.

While over 50 Level 4 and 5 fish studies (covering a range of species) and one amphibian study are available for EE2, most comprised only one or two exposure concentrations without analytical verification, and so failed to comply with the relevancy criteria applied in the current assessment (defined in Section 15). In addition, many studies displayed significant effects at the lowest exposure concentration (unbounded NOEC) and did not meet the ECHA (2008) criteria for estimating a NOEC from the LOEC, and as such also failed to meet the relevancy criteria applied here. Finally, for some studies only an abstract was available which was not considered to contain sufficient detail to determine the reliability and relevance of a study (Reliability of 4 – not assignable).

While the unreliable and not relevant studies were excluded from the present PNEC assessment, some have nevertheless been discussed, because they were included by JRC (2022) or Matthiessen *et al*. (2017) to determine their HC_5 values.

Following assessment of reliability and relevance, only nine EE2 studies were considered reliable *and* relevant for PNEC derivation; eight fish studies and one amphibian study (studies in black in Table 8). Of the fish studies, three were Level 4 studies and five were Level 5 studies. These studies, and those evaluated by Matthiessen *et al*. (2017) and JRC (2022) and used in their HC_5 determinations, are detailed in Appendix 203, Table 52, Table 53 and Table 54.

Table 8 Reliable and relevant chronic vertebrate data for EE2

Species (taxonomic group)	Endpoint (test design)	NOEC (ng/L)	LOEC (ng/L)	Reference
Xenopus tropicalis	Sex ratio (exposure from NF stages 47 – 66)	**1.77**	17.7	Pettersson *et al.*, 2007
Danio rerio	F1 fertility (multigeneration initiated with spawning fish exposed for 40 days, F1 exposure until 210 dpf and F2 until 100 hpf)	< 0.5	0.5	Nash *et al.*, 2004
	F1 growth, mortality (FLCTT initiated with embryos exposed for 8 months and F1 until 80 hpf)	< 0.19	0.19	Soares *et al.*, 2009
	Sex ratio (partial FLCTT initiated with fry exposed from 43-71 dph, then held in clean water until 190 dpf)	< 1.67	1.67	Maack and Segner, 2004
	Sex ratio (FLCTT initiated with embryos exposed for 4 months and F1 for 7 days)	**0.05**	0.5	Larsen *et al.*, 2008
	F1 growth (Multigeneration initiated with embryos exposed for 155 days and F1 until 36 dph)	0.1	0.3	Wenzel *et al.*, 2001
	F0 growth/reproduction/fertility (FLCTT initiated with embryos exposed for 177 days)	0.31	1.1	Schäfers *et al.*, 2007

Species (taxonomic group)	Endpoint (test design)	NOEC (ng/L)	LOEC (ng/L)	Reference
Pimephales promelas	Sex ratio/fertility (FLCTT initiated with embryos exposed to 150 dph)	< 0.32	0.32	Parrott and Blunt, 2005
	F1 growth (multigeneration initiated with embryos exposed to 301 dph and F1 until 28 dph)	**0.76**	2.8	Länge *et al.*, 2001
Gobiocypris rarus	Reproduction/fertilisation success (multigeneration study initiated with embryos. Complete exposure of F0 and F1 generations. F2 not initiated as complete spawning failure in F1 generation)	< 0.18	0.18	Zha *et al.*, 2008
Rutilus rutilus	Sex ratio (partial FLCTT initiated with embryos exposed until 112 dph)	< 0.3	0.3	Lange *et al.*, 2008
Oncorhynchus mykiss	F1 survival (587 dpf male fish exposed for 56 days, semen collected to fertilise unexposed eggs, F1 exposed until 19 dpf)	**< 0.8 or 0.4** (LOEC/2 est.)	0.8	Brown *et al.*, 2007
	Sex ratio (60 dpf male fish exposed for 76 days)	10	80	Depiereux *et al.*, 2014
	Reproduction (adult male fish exposed for 62 days, semen collected to fertilise unexposed eggs, F1 were held in clean water until 28 dpf)	< 15.6	15.6	Schultz *et al.*, 2003

Species (taxonomic group)	Endpoint (test design)	NOEC (ng/L)	LOEC (ng/L)	Reference
Cyprinodon variegatus	Reproduction/hatching success (partial FLCTT initiated with 6 week old fry exposed for 59 days and F1 exposed until 7 dph)	**18.1**	117	Zillioux *et al.*, 2001
Poecilia reticulata	Sex ratio/male growth (newborns exposed 108 days; exposed and unexposed males mated with unexposed females and F1 exposed 72 d)	**44**	112	Nielsen and Baartrup, 2006; Kristensen *et al.*, 2005
Cnesterodon decemmaculatus	Sex ratio (newborns exposed for 90 days, after completion of sexual maturity)	< 32.2	32.2	Young *et al.*, 2020
Oryzias latipes	Reproduction (hatched fry exposed for 60 days)	1.0	10	Scholz and Gutzeit, 2000
Pomatoschistus minutus	Reproduction/fertility (initiated with immature fish exposed for 7 months to assess reproduction)	< 6.0	6.0	Robinson *et al.*, 2003
Fundulus heteroclitus	Reproduction/hatching success (spawning fish exposed for 28 days and F1 exposed for 61 weeks)	10	100	Peters *et al.*, 2010

LOEC/2 est: estimated from LOEC/2 as data meet ECHA (2008) guidance

dph: Days Post Hatch

dpf: Days Post Fertilisation

hpf: Hours Post Fertilisation

NOEC values in **bold** are the lowest bounded NOEC for that species from studies considered reliable and relevant in the current report.

Studies highlighted in red are for supporting use only, as no NOEC could be derived or studies were not considered reliable and relevant in this report.

From the amphibian study (Pettersson and Berg, 2007) it was possible to derive bounded NOEC and LOEC values. It was only possible to derive a bounded NOEC from studies with four fish species (*D. rerio*, *P. promelas*, *C. variegatus* and *P. reticulata*; data highlighted in bold in the third column of Table 8). NOEC values ranged from 0.05 to 44 ng/L. In the majority of cases where effects were reported at the lowest exposure concentration, these effects were > 20% of the NOEC (control), and therefore it was not possible to use the LOEC/2 approach to estimate the NOEC. This approach was only possible, for a fifth fish species with *On. mykiss* (Brown *et al*., 2007).

5.2.2 Suitability of the EE2 dataset for derivation of a PNEC and data gap analysis

Despite a large number of available chronic aquatic vertebrate studies on EE2, reliable and relevant data are only available for six species. This much reduced EE2 dataset has been assessed against the SETAC Pellston Workshop® criteria for deriving a PNEC for an EDC (Matthiessen *et al*., 2017).

1. *Have the most appropriate taxa been tested (with relevant endpoints)?*

The ED MoA of EE2 is an ER agonist and the dataset is based on the most sensitive species (fish and amphibians) known to possess ERs with changes in population-relevant endpoints (i.e. development and reproduction endpoints and sex ratio). The available fish dataset comprises five species within four families (*Cyprinidae, Cyprinodonitadae, Poeciliidae* and *Salmonidae*), including standard OECD test species, freshwater and an estuarine species, and both egg and livebearers. In addition, one amphibian study is included. There are no marine species. The most sensitive endpoints were derived from standard regulatory test species (i.e. *D. rerio* and *P. promelas*).

2. *Have sensitive life stages, or the entire life cycle, been tested (with relevant endpoints)?*

There are reliable and relevant OECD CF Level 4 and 5 tests available, including those similar in study design to a modified LAGDA, FLCTT and multigenerational tests. These tests cover the entire-life cycle and sensitive life stages (embryos and juveniles). These studies also include relevant endpoints sensitive to ER agonist activity, including fecundity and fertilisation success and a female bias in the sex ratio. The majority of the dataset is derived from life cycle studies (Level 5) which included sensitive endpoints such as sex ratio and reproduction. Due to the lack of reliable Level 4 fish studies, it is not possible to determine, based on the available data, if these studies would have been as sensitive as the Level 5 studies. However, the following Level 4/5 studies with *P. reticulata* (Nielsen and Baartrup, 2006; Kristensen *et al*., 2005), *C. variegatus* (Zillioux *et al*., 2001) and *On. mykiss* (Depiereux *et al*., 2014) were initiated with newborns, fry, or sub-adult fish, respectively rather than with embryos. Furthermore, only the study with *C. variegatus* included a second generation and then the second-generation fry were terminated at

7 days old. These fish studies are much less sensitive (NOECs from 10 – 44 ng/L) compared with the studies incorporating an exposure from embryos to sexual differentiation (NOECs 0.05 - 0.76 ng/L).

3. *Have delayed and multigenerational (i.e. latent) effects been considered?*

The majority of the dataset comprises life cycle studies, three life cycles with the *D. rerio*, one with *P. promelas* and one with both *C. variegatus* and *P. reticulata*. Of these, those with *D. rerio* (Wenzel *et al*., 2001) and *P. promelas* (Länge *et al*., 2001) were exposed over two generations. These multigenerational tests should be sufficient to address any concerns for delayed or latent effects of EE2.

4. *Do Non-Monotonic Dose Relationships (NMDRs) or other unusual temporal patterns of toxicity affect the ability to predict reliable no-adverse-effect levels?*

There was no evidence for an NMDR relationship for the key indicators for endocrine activity (i.e. VTG induction, induction of testis-ova, SSC, or gonad histopathology) or effects on EAS mediated endpoints (i.e. sex ratio).

There was limited evidence for NMDRs in some of the multigenerational studies based on either population-relevant or mechanistic endpoints. These included an increase in male GSI at the lowest exposure concentration (Kristensen *et al*., 2005) and in Larsen *et al*. (2008) there was a significant decrease in male weight at 0.05 ng/L (lowest exposure concentration) and an increase at 5.58 ng/L (highest exposure concentration). Although there was a significant increase in GSI at the lowest measured concentration there was a clear non-significant tendency toward a decrease in GSI at the other two measured concentrations. No explanation was provided by the authors for the increase (Kristensen *et al*., 2005). No explanation was provided by the authors for the changes in male weight at the lowest exposure concentration; however, the increase at 5.58 ng/L could be associated with feminisation (Larsen *et al*., 2008).

5. *Does a threshold for adverse endocrine-mediated effects exist?*

Figure 5 shows the range of NOEC values for endocrine population-relevant (reproduction and sex ratio) and “biomarker” endpoints (VTG induction, gonad histopathology, SSC, steroid hormone levels and presence of testis-ova) from sixteen reliable OECD CF Level 4 or 5 fish studies (note that most of these studies were not used for PNEC derivation, either because they did not meet relevance criteria or because there was a more sensitive result with the same species). NOEC values generally cover the range 0.01 - 100 ng/L) excluding some insensitive results for gonad histopathology and steroid hormone measurements. The lowest “biomarker” NOEC (LOEC/2) was 0.025 ng/L based on SSC in the same *D. rerio* life cycle study as the lowest population level NOEC of 0.05 ng/L (Larsen *et al*., 2008). These data would appear to demonstrate a threshold level for endocrine effects in fish with NOECs all being above 0.01 ng/L.

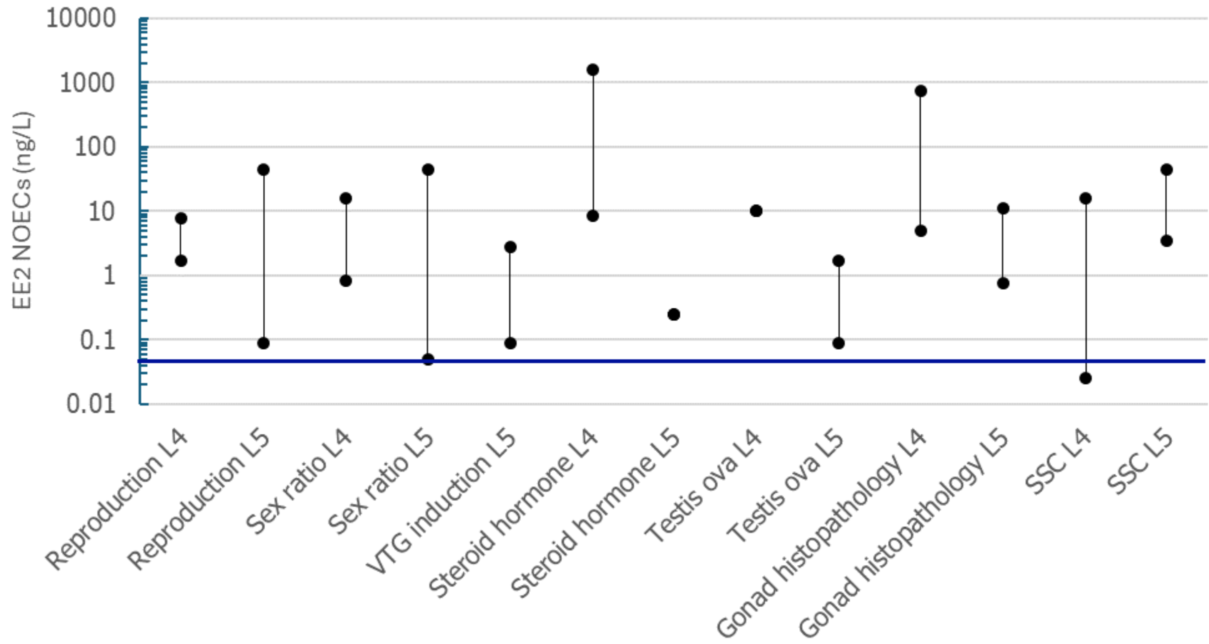


Figure 5 Range of NOEC values from all reliable OECD CF Level 4 and 5 fish studies with EE2. The blue line indicates the lowest population level NOEC of 0.05 ng/L based on sex ratio in a *D. rerio* life cycle study (Larsen *et al*., 2008).

Some caution should be applied since several studies were unbounded (including for biomarker endpoints), although this is partly explained by the lack of standardised test guideline studies. Many of the studies included only three test concentrations and wide spacing of test concentrations (e.g. a factor of 10 in the *D. rerio* life cycle study on which the driving NOEC for the assessment is based (Larsen *et al*., 2008)).

Overall, the reliable and relevant dataset for EE2 generally meets the SETAC Pellston Workshop® criteria for deriving a PNEC for an EDC.

5.2.3 Deterministic PNEC assessment for EE2

Only vertebrates were included in the assessment based on the known ER agonist MoA (this approach is also supported by the NP and OP analysis). The reliable dataset for fish includes OECD CF Level 4 and 5 studies with 5 fish and 1 amphibian incorporating relevant endocrine mediated endpoints. The most sensitive NOEC is 0.05 ng/L based on a shift in sex ratio in a *D. rerio* life cycle study (Larsen *et al*., 2008).

Using standard approaches the PNEC for EE2 is derived deterministically by applying an AF of 10 to the NOEC of 0.05 ng/L as follows:

EE2 deterministic PNEC = 0.05/10 = 0.005 ng/L

A case could be made for the application of additional AF based on the substance having a confirmed endocrine MoA. However, the endocrine mechanism is already accounted because several multigeneration fish studies are included in the dataset. Despite the unbounded NOEC values, it would also be protective of the lowest "biomarker" NOEC (or LOEC/2) of 0.025 ng/L for effects on SSC in the same study (Larsen *et al*., 2008).

Further discussion of AF is given in Section 78.

5.2.4 Probabilistic assessment for EE2

Table 9 summarises the six NOEC values that have been applied in deriving an HC_5 value for EE2. The dataset includes relevant and reliable Level 4 or Level 5 CF fish and amphibian studies. The fish taxa include both estuarine and freshwater species.

Table 9 Values for selected for deriving a HC_5 for EE2

Species	Endpoint (NOEC)	Value selected for HC5 (ng/L)	Reference
Xenopus tropicalis	Sex ratio	1.77	Pettersson and Berg, 2007
Danio rerio	Sex ratio	0.05	Larsen *et al*., 2008
Pimephales promelas	Reproduction	0.76	Länge *et al*., 2001
Oncorhynchus mykiss	Juvenile mortality	0.4 (est.)	Brown *et al*., 2007
Cyprinodon variegatus	Reproduction	18.1	Zilliouz *et al*., 2001
Poecilia reticulata	Reproduction/sex ratio	44	Nielsen & Baartrup, 2006; Kristensen *et al*., 2005

est.: estimated from LOEC/2

SSDs for EE2 were generated using the log normal method based on ETX version 2.3 or the weighted average method using ssdtools.

The SSDs for the reliable and relevant EE2 dataset derived using ssdtools are shown in Figure 6.

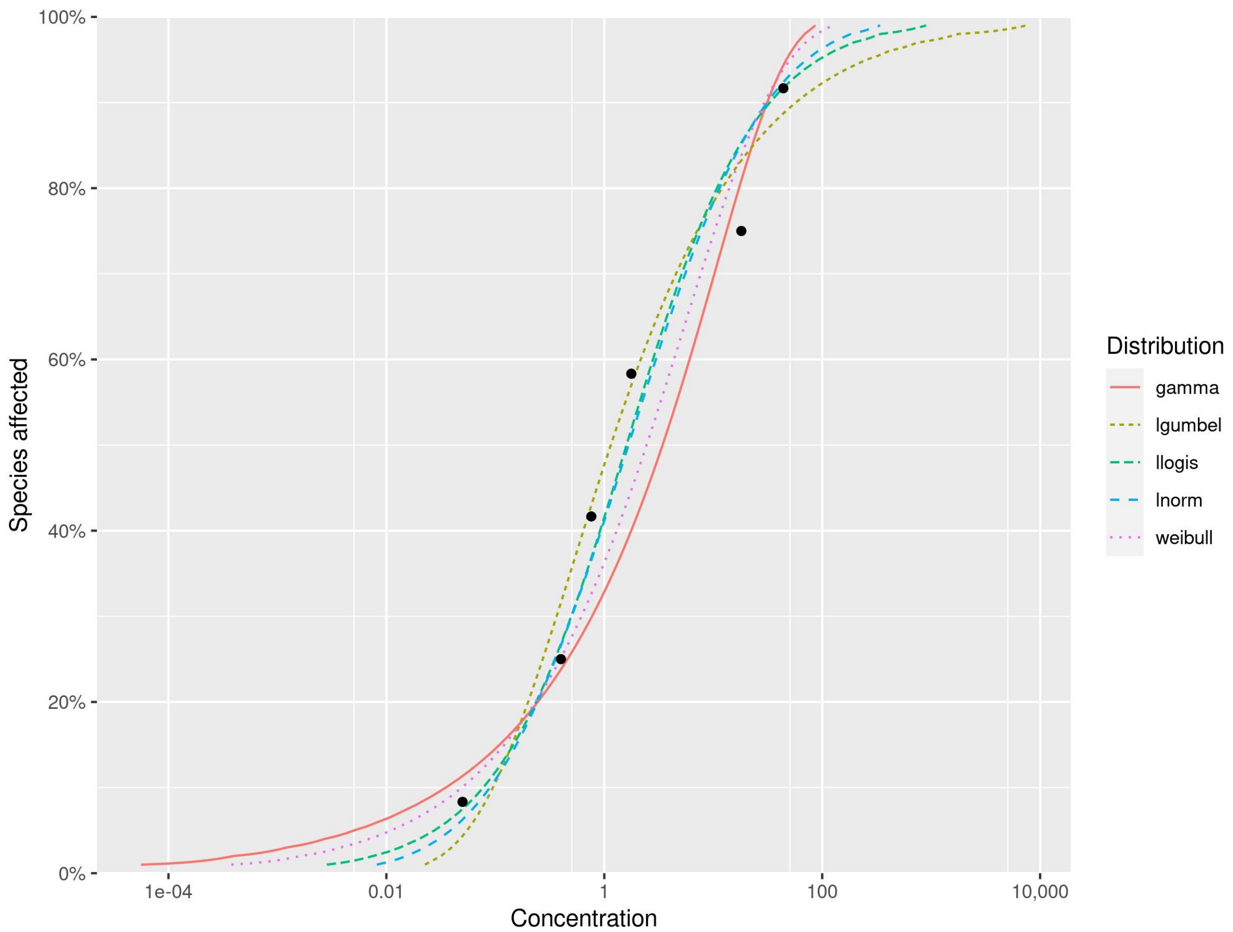


Figure 6 SSDs models from ssdtools for deriving an aquatic HC_5 for EE2

The HC_5 values derived using the different distributions and using the ETX software are given in Table 10.

Table 10 HC_5 values derived using the different distribution models

Distribution	HC5 (ng/L)	Confidence Intervals (ng/L)	AICc[1]	Weight	Software
Log normal	0.021	0.000155 - 0.186	-	-	ETX v. 2.3
Average	0.0288	0.00372 – 0.974	-	-	ssdtools v. 1.0.2
Gamma	0.00506	0.00000380 – 1.73	- 3.59	0.168	
Lgumbel	0.055	0.0152 – 0.813	- 3.96	0.202	
Llogis	0.0276	0.00126 – 0.608	- 3.83	0.189	
Lnorm[2]	0.0388	0.00349 – 1.05	- 4.33	0.243	
Weibull	0.0111	0.0000657 – 0.969	- 3.92	0.198	

Andersen Darling statistic was not calculated in ssdtools due to small dataset
[1] Akaike's Information Criterion corrected for sample size (AICc)
[2] Note that a small differences in model results (less than a factor of 2) for different estimation packages is expected due to different estimation strategies (Fox *et al.*, 2021)

The model average HC_5 for EE2 is 0.0288 ng/L, with confidence limits of 0.00372 – 0.974 ng/L. Using ssdtools it is possible to demonstrate how the different distributions, all of which have an acceptable fit to the dataset, lead to HC_5 values which vary by an order of magnitude (between 0.055 and 0.005 ng/L). The challenge in fitting a curve to the data adds to the uncertainty. It is however worth noting that the initial log normal HC_5 derived using ETX v. 2.3 of 0.021 ng/L is similar to the model average HC_5 value.

The wide confidence limits around the HC_5 values reflect the small dataset and the spread of the values used to generate the SSD and indicates that the estimate is uncertain. Nevertheless, a PNEC has been derived by applying the default AF of 5 to the model average HC_5 value as follows:

Probabilistic EE2 PNEC = 0.0288/5 = 0.00576 ng/L

The outcome is similar to the deterministic PNEC of 0.005 ng/L.

The reasons for the lack of resolution regarding effects on EE2 at low concentrations may reflect that most of the studies were conducted for research rather than regulatory purposes, where the primary aim was not necessarily to ensure a 'no-effect' concentration was produced. This includes studies with wide test concentration ranges (e.g. the driving NOEC for the assessment is from a study (Larsen *et al.*, 2008) with a factor of 10 spacing between the treatments). There have also been improvements in analytical detection techniques over time, so it would now be possible to analytically measure the lower exposure concentrations used in some of the studies excluded for this reason.

Another factor is related to experimental design. The studies with *C. variegatus* and *P. reticulata* (which both show much lower sensitivity to EE2 compared with the other species) do not include exposure of early embryo stages (exposures initiated with 6-week-old fry or newborns, respectively). This represents a limitation of the SSD approach for EE2, where the small dataset is potentially biased by the inclusion of NOEC values from studies which do not follow an experimental design which is sensitive to the ER agonist MoA. Excluding the insensitive studies would lead to a dataset based on four species which is insufficient for an SSD.

Although not a standard approach, for comparative purposes the HC_5 for EE2 has been derived using all nine reliable NOEC values (including those with the same species) to determine if this improves confidence in the HC_5 (see Appendix 215). The alternative model average HC_5 was very similar (0.029 ng/L), with confidence limits of 0.00698 ng/L to 0.375 ng/L. The model average HC_5 was then calculated removing the two insensitive studies of lower relevance to give a value of 0.032 ng/L (0.00861-0.253 ng/L), further improving confidence in the HC_5. However, this result is based on just four species and removes the live bearer and estuarine species from the dataset.

Nevertheless, it provides some confidence that the model average HC_5 would be acceptable for use for PNEC purposes. However, given the similarities between the deterministic and probabilistic PNECs the use of the deterministic PNEC is potentially more justifiable.

For context, the PNEC derived in the present assessment is about two orders of magnitude lower than the concentration (5 to 6 ng/L) which led to the collapse of *P. promelas* populations in a Canadian Lake exposed over three years during the summer months (Kidd *et al*., 2014).

5.2.5 Comparisons with other EE2 PNECs

The HC_5 of 0.0228 ng/L derived in the present assessment is lower compared with HC_5 values of 0.047 ng/L (Matthiessen *et al*., 2017) or 0.0869 ng/L (JRC, 2022) (Table 11). While some variation could be due to potential differences in the statistical approaches used to derive the HC_5s, the main differences appear to be:

- The more stringent data inclusion criteria applied to the present dataset, and
- Use of the NOEC of 0.05 ng/L from the Larsen *et al*. (2008) study.

Table 11 Comparison of HC_5/PNEC values for EE2

Dataset	Lowest NOEC or HC5 (ng/L)	HC5 confidence limits (ng/L)	AF[1]	PNEC	Model used
Deterministic	0.05	-	10	0.005	NA
Present dataset (five fish and one amphibian species)	0.0288	0.000155 – 0.186	5	0.00576	Model average, ssdtools v. 1.0.2
Matthiessen *et al.* 2017 (eight fish species)	0.047	Not reported	5	0.0094	Not reported
JRC 2022 (eight fish and two amphibian species)	0.0869	0.0191 – 0.5491	5	0.017	Log-normal. Bootstrap methods. R version: 3.2.3.; fitdistrplus Package version 1.0.14

[1] Default AF applied

The studies that were considered reliable and relevant in the present assessment, and those considered reliable by JRC (2022) and Matthiessen *et al*. (2017) for use in an SSD, are compared in the Table 12. The differences between the NOECs chosen for inclusion in the SSD for each assessment are discussed in Section 56.

Table 12 Comparison of NOEC values (ng/L) selected for inclusion in an SSD for EE2

Species	NOEC values selected for SSD (Endpoint) (Reference)		
	This report	JRC, 2022	Matthiessen *et al.*, 2017
X. tropicalis	1.77 (sex ratio) (Pettersson and Berg, 2007)		No studies included
Danio rerio	0.05 (sex ratio) (Larsen *et al.*, 2008)	0.3 (reproduction) (Wenzel *et al.*, 2001)	0.21*(fertility) (Signer *et al.*, 2003; Schäfers *et al.*, 2007; Larsen *et al.*, 2008)
Pimephales promelas	0.76 (reproduction) (Länge *et al.*, 2001)		0.86* (reproduction) (Länge *et al.*, 2001; Parrott and Blunt, 2005)
Gobiocypris rarus	No studies included	0.09 est (fertility/sex ratio) (Zha *et al.*, 2008)	
Rutilus rutilus	No studies included	0.3 (sex reversal) (Lange *et al.*, 2008)	No studies included
Oncorhynchus mykiss	0.4 estimated (juvenile mortality) (Brown *et al.*, 2007)	8 (juvenile mortality) (Schultz *et al.*, 2003)	0.4 estimated (juvenile mortality) (Brown *et al.*, 2007)
Cyprinodon variegatus	18.1 (reproduction) (Zillioux *et al.*, 2001)		No studies included
Poecilia reticulata	44 (reproduction/sex ratio) (Kristensen *et al.*, 2005; Nielsen and Baartrup, 2006)		
Oryzias latipes	No studies included	1.0 (reproduction) (Schultz & Gutzeit, 2000)	2.0 (reproduction) (Balch *et al.*, 2004)
Pomatoschistus minutus	No studies included	No studies included	3.0 est (reproduction/fertility) (Robinson *et al.*, 2003)

Species	NOEC values selected for SSD (Endpoint) (Reference)		
	This report	**JRC, 2022**	**Matthiessen *et al*., 2017**
Fundulus heteroclitus	No studies included	No studies included	10 (fertility/hatch/sex ratio) (Peters *et al*., 2010)

est: estimated from LOEC/2

* geometric mean value

Overall, despite the relatively large vertebrate dataset for EE2, the NOECs from only six species were considered suitable for use in generating an SSD for the present assessment. Other assessors (Matthiessen *et al*., 2017; JRC, 2022) have attempted to maximise the dataset by being less strict with the data assessment (e.g. inclusion of LOEC values and results from studies based on nominal, or fewer than three test concentrations). While this larger dataset generates more statistically robust outcomes from the probabilistic assessment (e.g. narrower confidence limits around the HC_5), the equivocal reliability and relevance of the data used in the assessment (particularly studies with significant effects at the lowest test concentration) could be argued to undermine the outcomes to an equal extent. It is also unclear why the Larsen *et al.* (2008) NOEC of 0.05 ng/L based on both fertility and sex ratio was not a driving value in either of the previous assessments. However, arguably this NOEC is precautionary since there is a factor of 10 between the treatments in this study. The lowest NOECs for both population-relevant and biomarker endpoints were from this study (See Figure 6).

The main concerns for the probabilistic assessment are:

- the lack of robust data designed for regulatory purposes leading to studies with unbounded NOECs, lack of analytical verification, and wide spacing of concentration ranges,
- the small dataset, and
- the inclusion of data from potentially insensitive studies.

While the conduct of a new multigeneration regulatory standard fish study with EE2 with a different species (e.g. a MEOGRT) could potentially reduce the uncertainties in a probabilistic risk assessment, especially at the low end of the SSD, the overall benefit may be limited because the dataset would still be small, and the other NOEC values would still have the same data quality issues.

It is therefore recommended that for EE2 a deterministic approach is followed. The existing dataset can then be used in a more qualified manner to refine the applied AF.

5.3 Summary and conclusion

The synthetic steroid hormone EE2 is known to be a highly potent ER agonist. Based on its MoA, fish and amphibians were considered to be the most sensitive taxa, because they have functioning ERs and changes in relevant endpoints confirm this endocrine activity (i.e. VTG induction in male fish, female bias in sex ratio). The aim of this element of the assessment was to evaluate the potential for updating a probabilistic HC_5 value for EE2 derived for fish in Matthiessen *et al*. (2017). The reliability in this dataset was re-evaluated to be consistent with the approach taken in the present assessment for NP and OP. Data from a recent literature search, chronic fish and amphibian studies reported in JRC (2022) and EC (2002) were also considered in the present assessment.

Despite a relatively extensive chronic EE2 dataset for fish and amphibians, only nine OECD CF Level 4 or 5 studies were considered reliable and relevant for PNEC derivation. Several studies were considered not relevant for PNEC derivation because a clear (bounded) NOEC could not be established, with effects reported at the lowest exposure concentration, and for which a NOEC could not be reliably estimated from the LOEC. The lowest reliable and relevant NOEC for EE2 was 0.05 ng/L based on sex ratio in a life cycle study with *D. rerio* (Larsen *et al*., 2008).

The appropriateness of this dataset for deriving a PNEC suitable for deriving threshold values for an EDC was evaluated against the SETAC Pellston Workshop® criteria. Although the dataset included multiple multigeneration studies exposing sensitive life stages without convincing evidence of NMDR or delayed effects, there were issues with the dataset because it was only possible to derive NOEC values for six species. For one of these species (*On. mykiss*), the NOEC was estimated. The two least sensitive species in the dataset were the guppy, *P. reticulata*, and the Sheepshead minnow, *C. variegatus*. Although the available studies were life cycle or partial life cycle tests there was no chronic exposure during early embryo life-stages in these tests, which may also have included the sensitive window of sexual differentiation. It is therefore likely that the experimental design of these tests was not sufficiently sensitive to detect the ER agonist mechanism of action.

Although the biomarker endpoints for ER agonist activity (i.e. VTG induction, SSC, gonad histopathology) typically followed a concentration response, effects on these endpoints were usually observed at the lowest or adjacent to the lowest tested concentration (even in studies where a NOEC for population-relevant endpoints was obtained). The lowest NOEC for a biomarker endpoint was 0.025 ng/L (derived from the LOEC/2) for SSC from the same study with *D. rerio* as the lowest NOEC for a population-relevant endpoint.

Despite the number of studies, the dataset for EE2 does not appear to meet all the SETAC Pellston Workshop® criteria for conducting a risk assessment due to unbounded studies making a threshold level difficult to define and potentially insensitive studies for the live bearers and estuarine species. The issues appear to be related more to experimental design (because published studies were not conducted for regulatory purposes) rather than not being able to derive an ETNC. Nevertheless, a threshold value does appear to be plausible as there were no effects of EE2 at concentrations below 0.01 ng/L.

PNECs have been derived using deterministic and probabilistic approaches. A deterministic PNEC of 0.005 ng/L was derived by applying a default AF of 10 to the lowest NOEC value of 0.05 ng/L, following standard risk assessment approaches. An AF of 10 is protective for the NOECs (or LOECs/2) for “biomarker” responses.

A model average HC_5 for fish and amphibians of 0.0228 ng/L, with confidence limits of 0.00372 ng/L to 0.974 ng/L, was derived. There are wide confidence intervals for all HC_5 values because of challenges in fitting a model to the data. Using a default AF of 5, the PNEC of 0.00576 ng/L is similar to the deterministic PNEC.

The PNECs derived in the present assessment are lower compared with the EU EQS of 0.17 ng/L and a PNEC derived using the HC_5 from the SETAC Pellston Workshop® (0.0094 ng/L). Although the strict criteria applied to the present dataset explains some of the differences, the main reason was that the other assessments did not use the NOEC from the fish life cycle study with *D. rerio* (Larsen *et al*., 2008). This NOEC is more precautionary because there was a factor of 10 spacing between the test concentrations.

Given the availability of fish studies with EE2 it may be possible to justify reducing the default AFs proposed above. AFs are discussed further in Section 77.

6. Reflections on findings for NP, OP and EE2

6.1 Data for use in PNEC derivation for ER agonists

Reliable ecotoxicity data have been compiled for NP, OP and EE2. These substances are recognised EDCs and have an ER agonist MoA. SETAC Pellston Workshop® criteria were applied to the datasets to determine whether they were suitable for PNEC derivation based on effects in OECD CF Level 4 and 5 studies (Matthiessen *et al*., 2017).

For NP and OP both vertebrate and invertebrate data were used to maximise the dataset, even though invertebrates are not known to be affected by an ER agonist MoA. For both substances, however, the most sensitive NOECs were from fish studies, and inclusion of invertebrate data in the SSD led to HC_5 values that were arguably not sufficiently precautionary. In addition, for NP there were only three reliable fish NOECs so it was not possible to derive a probabilistic PNEC based on vertebrate data only. The most sensitive NOECs in the NP dataset were from multigeneration studies (Watanabe *et al*., 2017; Schwaiger *et al*., 2002 (the latter as supporting information only)) suggesting that the dataset for deriving a PNEC for an EDC should include an OECD CF Level 5 fish study. Although the lowest NOEC in the OP dataset was from a Level 4 fish study there were multigeneration fish studies in the dataset which improves confidence in the HC_5 value.

The dataset for EE2 only included vertebrate NOECs. However, there were issues with the dataset related to experimental design. NOECs from fish studies initially classified as OECD CF Level 4 or 5 studies were relatively insensitive to the oestrogenic effects of EE2 because the sensitive embryo life stages were not exposed. In the case of the Level 5 test, it was a test where F0 generation embryos were not exposed, but F1 were, although only to 7 dph (this shows that careful attention should be paid to the treatment of non-standard studies and which CF Level they should be assigned to). Also, most of the studies were not generated for regulatory purposes and there were issues with unbounded NOECs and wide spacing of test concentrations which reduces confidence in the experimental NOEC value in some cases. This was reflected in the SSD where wide confidence intervals were reported (Table 13). In contrast there was much more confidence in the SSD for OP which was based on regulatory style tests generated as part of test guideline validations.

Table 13 HC_5 values and confidence limits for OP and EE2

Substance	Number of species	HC_5	Confidence limits	Unit	Difference in confidence limits	Model
OP	6	3.87	2.04 – 6.56	µg/L	3.2-fold	Model average, ssdtools v 1.0.2
EE2	6[a]	0.0288	0.00372 – 0.974	ng/L	262-fold	

[a] Included insensitive NOECs from studies without embryo exposure

It is worth noting that the SSDs for both OP and EE2 were based on a minimum ssdtools dataset of six vertebrate species. For small datasets, ssdtools provides a way of considering different model fits to the data and improved confidence by using a consensus model average fit. These substances have relatively rich datasets, and it is unlikely that there will be enough data available for most other EDCs to generate an SSD. It is also not ethically justified to generate such datasets by increased testing of vertebrates. Therefore, a deterministic approach will be required for most substances. However, the conclusion from this analysis would be that the dataset for deriving a PNEC for an EDC would at a minimum include a fish test with exposure of embryo life stages, and ideally a multigeneration fish study.

6.2 Assessment factors

The PNECs for NP, OP and EE2 are given in Table 14 including the vertebrate dataset on which the PNEC is based.

Table 14 PNECs for NP, OP and EE2

Substance	Type of PNEC	PNEC value	Biomarker NOEC	Unit	AF	Vertebrate dataset
NP	Deterministic	0.064	0.4	µg/L	10	1 MEOGRT 2 FELS
OP	Deterministic	0.61	1	µg/L	10	7 FSDT 1 FELS 3 LAGDA 1 MEOGRT 1 FLCTT
	Probabilistic	**0.77**		µg/L	5	
EE2	Deterministic	**0.005**	0.025	ng/L	10	2 Level 4 with fish[1] 1 Level 4 with Amphibian 5 FLCTT[1]
	Probabilistic	0.006		ng/L	5	

[1] Included insensitive NOECs from studies without embryo exposure
PNEC values in bold text are those selected for use.

For the deterministic assessment a default AF of 10 is applied to the lowest NOEC, whereas for the probabilistic assessment a default AF of 5 is applied to the HC_5 value, due to the more comprehensive dataset. This is consistent with the approach for other chemicals, where the regulatory data set is assumed to be protective of the aquatic environment even if the MoA is not known. The derived PNECs are protective for biomarker effects in all cases. Biomarker effects are often observed at the same or at lower test concentrations compared with population level effects. We suggest that no additional AFs are required for the present datasets to account for ED. In fact, there is a case for potentially reducing the AF for EE2 based on the number of fish studies available and because the NOEC from the driving study is already precautionary because test concentrations are separated by a factor of 10. For the deterministic assessment for EE2 an AF of 5 rather than 10 would give a PNEC of 0.01 ng/L which is more aligned with the EU EQS (JRC, 2022), is protective for biomarker endpoints and is justified by the number of fish studies including an assessment of the dataset against SETAC Pellston Workshop® criteria. Nevertheless, for a PNEC derived for REACH purposes, an AF below 10 is only used for a SSD derived probabilistic PNEC, and not for a deterministic PNEC.

A caveat to the above is that none of the datasets include marine species. Whilst this does not affect the derivation of PNECs for surface water, this would need to be considered further for marine scenarios. Also, most of the fish studies are based on small, rapidly growing fish species that are used in standard test guideline studies. A study using the longer-lived species *On. mykiss* for OP provided the most sensitive NOEC in the dataset despite being the result from a Level 4 FELS test (ABC labs, 1984, cited in EA 2005 and ECHA 2011). There is only one study for EE2 that reports effects in a species with an alternative reproductive strategy (live bearer) (Nielsen and Baartrup, 2006; Kristensen *et al.,* 2005) and it is unclear whether the test design was adequate to capture ER agonist effects as there was no maternal exposure.

Overall, default AFs appear to be acceptable for the ER agonist MoA when applied to the lowest NOEC from a dataset which includes an OECD CF Level 5 fish study. Expert judgement can potentially be used to refine these AFs using mechanistic data and the SETAC Pellston Workshop® criteria to support the arguments.

6.3 Environmental Threshold of No Concern

One of the aims of this project was to consider whether an ETNC is appropriate for the ER agonist MoA. The PNECs derived for NP, OP and EE2 suggest that thresholds can be defined on an individual substance basis, with values of 770, 64 and 0.005 ng/L for OP, NP and EE2, respectively, reflecting the increasing potency of the substances. Information on the endocrine activity of all the substances included in the review (NP, OP, EE2 and the eight validation substances), based on data extracted from the CompTox dashboard, is given in Appendix 218 (Table 56). The synthetic steroid hormone EE2 has very high affinity to the ER, with a potency in the US EPA Bioactivity model of 1, setting the benchmark against which other ER agonist substances are assessed. In *in vitro* assays, EE2 causes effects at concentrations several orders of magnitude lower than NP or OP.

An example (shown in Table 15) is for the *O. latipes* ERα where the EC_{50} for EE2 is 0.00009 µM compared with 0.022 µM and 0.023 µM for NP and OP, respectively (based on data from Onishi *et al*., 2020 and Kawashima *et al*., 2021). Other ER agonist substances such as bisphenol A (BPA) or methyl paraben and propyl paraben show lower activity via the fish ERα than NP or OP. ER agonist activity in the *O. latipes* ERα has been shown to give a similar outcome to the ToxCast bioactivity model result (Brown *et al*., 2023). There also appears to be some relationship between ER agonist activity *in vitro* and effects on VTG in FSTRA studies with *O. latipes* (FSTRA data from Onishi *et al*., 2020 and Kawashima *et al*., 2021), with notable exceptions for BPA and benzophenone-2 which were not as sensitive as might be expected for VTG induction. Other factors such as differences in adsorption, distribution and metabolism or an alternative MoA may affect the *in vivo* response (Ankley *et al*., 2016; Brown *et al.,* 2023).

It is worth noting that the US EPA Bioactivity model outputs and medaka ERα are *in vitro* screening level data only. Consideration of all available data is likely to give a more informed assessment of endocrine activity and relative potency. For example, the data for medaka ERα activity suggests that OP and NP have equivalent potency for the ER, whereas consideration of the Bioactivity model results and *in vivo* VTG data confirms the lower potency of OP.

Table 15 Summary of ER agonist activity in the ToxCast Bioactivity model, in *vitro* in fish receptor assays and *in vivo* in FSTRA with medaka (*O. latipes*)

Substance	Bioactivity model result	Medaka ERα (EC_{50}) µM[a]	LOEC for VTG induction in a FSTRA with *O. latipes* (µg/L)
EE2	1.0	0.00009	0.018
Bisphenol A [b]	0.450	0.086	4 670
NP	0.435	0.023	5.63
OP	0.393	0.022	82.3
Benzophenone-2	0.397	1.6	9 530
4-tert-Pentylphenol	0.282	0.97	58.4
Propyl paraben	0.205	7	311
Pendimethalin	0.066	3.3	100

Substance	Bioactivity model result	Medaka ERα (EC_{50}) μM[a]	LOEC for VTG induction in a FSTRA with *O. latipes* (μg/L)
4-Chloro-3-methylphenol	0.029	61	1 060
Methyl paraben	0.0068	70	1 900

[a] Based on data reported in Onishi *et al*., (2020), Kawashima *et al*., (2021) and/or Brown *et al*., (2023)
[b] No effect in the absence of systemic toxicity

An initial assessment of the data for EE2 suggested that it may not be possible to derive an ETNC for a substance with such high potency for the ER. However, a threshold level does appear to exist, albeit at very low concentrations (Figure 6). The present analysis highlights the importance of well-designed ecotoxicity studies for demonstrating these threshold levels. These conclusions align with an UBA (2014) review of PNEC derivation for EDCs, where they consider that at least for the EAS MoA, the vertebrate tests covered by the OECD CF are adequate for use as part of a tiered testing strategy. They also consider that low dose effects, NMDR and behavioural effects should be adequately captured in the regulatory fish and amphibian tests.

Given the relationship between potency and the ER agonist MoA it is expected that receptor mediated activity *in vitro* could be used as supporting information for PNEC refinement in the following scenarios:

- Where a PNEC is being derived for a substance with a known MoA based on a limited *in vivo* dataset.
- To check the relevance of a PNEC (chemicals acting as ER agonists would generally not be expected to have lower PNECs than the EE2 threshold of 0.005 ng/L).
- As supporting information for using read across in PNEC derivation.

Although the combined datasets for OP and NP did not improve the PNEC assessment for either substance, it did suggest that data for a more potent substance could be used to derive a PNEC for a structurally similar substance with a similar endocrine mechanism but lower potency. *In vitro* data on potency could be used to support this approach.

7. Exploration of PNEC derivation for additional EDCs

7.1 Proposed AFs to be applied to a limited ED dataset

In this section, we use the findings for NP, OP and EE2 to propose potential AFs that could be used to derive deterministic PNECs for suspected EDCs that are data poor. Although the three case study substances are all ER agonists, any proposed approach would be expected to be relevant for other EAS mechanisms since the fish test endpoints covered in the OECD CF are designed to cover all these pathways. For NP, OP and EE2 fish were the most sensitive species in the dataset (when appropriate test designs were included). In all cases application of an AF of 10 to the lowest NOEC from a dataset which includes a multigeneration fish study (OECD CF Level 5) was considered protective of other species (amphibians and invertebrates) as well as biomarker endpoints for the endocrine mechanism of concern. Therefore, this is a useful starting point for development of a proposal for AFs to be applied to a limited dataset. It is emphasised that these AFs are not proposed to apply to thyroid MoAs.

There is a precedent for the application of an AF of 10 to the NOEC from a Level 5 fish study for endocrine active substances in the environmental risk assessment for pharmaceuticals (European Medicines Agency (EMA), 2018).

An additional AF of 10 is then proposed based on available data at progressively lower OECD CF Levels as indicated in Table 16.

Table 16 Proposed AFs for the deterministic assessment of an EDC with limited data availability

Available data	Assessment factor
At least one NOEC/EC_{10} from a Level 3 fish study	1 000 [a]
One NOEC/EC_{10} from a Level 4 fish study	100 [b]
One NOEC/EC_{10} from a Level 5 fish study	10 [c]

[a] The use of an AF of 1 000 on Level 3 fish toxicity data is aimed at being a protective factor based on the low resolution of a study which is designed for screening purposes.
[b] An AF of 100 applies to a single long-term Level 4 result (e.g. EC_{10} or NOEC) for fish only, as the taxonomic group expected to be sensitive to an EAS mediated MoA. The Level 4 fish study must include chronic exposure of embryos (FELS test as a minimum).
[c] It is assumed that a Level 5 fish study will include more comprehensive data on endocrine relevant endpoints as well as covering extensive parts of the life cycle of the organism and therefore a lower AF of 10 can be applied to the NOEC/EC_{10} from this study. The Level 5 study must include chronic embryo exposure.

This approach is consistent with the general principles of the existing REACH technical guidance for deriving PNECs for chemicals with limited ecotoxicity datasets (ECHA, 2008). The guidance recommends the application of high precautionary AFs when the amount of data is restricted to short-term studies with algae, invertebrates and fish. The difference in the case of derivation of PNECs for EDCs is that the available data generally refer to “longer-term” (chronic) endpoints from OECD CF Level 3 or higher studies (most commonly with fish). The expectation is that there will be OECD CF Level 4 or 5 fish data available for deriving a PNEC for an EDC that is identified as an SVHC. Nevertheless, there may be scenarios where a Level 3 FSTRA is the only test available (e.g. if read across arguments are being made as part of a grouping assessment). This study is not usually considered for deriving a PNEC because the study design has low resolution due to wide spacing of test concentrations and a low level of within-treatment replication. However, for the purpose of this investigation, the proposed use of a FSTRA in deriving a precautionary PNEC for EDCs can be considered equivalent to use of acute toxicity data for deriving an aquatic PNEC for substances with a non-specific MoA (ECHA, 2008). This PNEC would be used as part of a screening level approach with the assumption that confirmatory higher tier testing would subsequently be conducted if a risk were identified (unless the risk can be removed through provision of other data, such as environmental degradation half-lives or better release estimates).

Even though a substance may have endocrine activity, it should not be assumed that fish will necessarily always be the most sensitive species in the dataset. For investigating effects via EAS modalities, it is usually recommended to conduct tests with fish (OECD, 2018, EFSA/ECHA, 2018) but equivalent tests in amphibians could be considered equally relevant for PNEC derivation (where already available and if EAS mediated parameters are covered). The LAGDA is an OECD CF Level 4 study which incorporates sex ratio and could be considered equivalent to an FSDT. The available information based on OP and EE2 suggests that the PNEC derived by applying an AF of 100 to the available Level 4 amphibian data would be lower than the lowest NOEC for fish in the respective datasets. Therefore, in theory the suggested approach could be based on “aquatic vertebrates” rather than fish. However, testing with amphibians is usually only recommended for substances acting via the thyroid (OECD, 2018; EFSA/ECHA, 2018) and there are likely to be few substances acting via EAS modalities which have relevant amphibian data to validate the proposed approach for EAS. In addition, the available data for amphibians suggests lower sensitivity compared with fish for the case study substances (possibly related to test design) and therefore fish are preferred for deriving a PNEC for an EDC acting via EAS modalities.

For the NP, OP and EE2 evaluations, most studies were reported based on the NOEC, and therefore NOEC values were generally used in the assessment. This is partly because the data were from older academic literature (few regulatory study reports were reviewed as part of this assessment). Conversion of NOEC values to EC_{10} values was beyond the scope of the present project and for some studies is not possible due to a lack of raw data availability and limited reporting. It is acknowledged that use of EC_{10} values may have reduced some of the uncertainties in the data resulting from unbounded studies or from studies with wide test concentration ranges, and is generally considered best practice for contemporary risk assessment. The proposed approach is therefore based on either a NOEC or EC_{10} value.

Unlike chemicals with non-specific MoAs (e.g. non-polar narcotics), EDCs have a known MoA and supporting biomarker information (based on endpoints such as VTG, SSC or gonad histopathology) is expected to be available for SVHCs (since they will form part of the scientific case for their identification). In this assessment these mechanistic endpoints are used for context as part of PNEC derivation. While the PNEC does not need to be protective for mechanistic effects (as changes in biomarker endpoints do not necessarily lead to a population-relevant outcome), if the proposed PNEC is lower than the NOEC for mechanistic effects then it could be assumed that it will be protective for population level effects in untested species. Sex ratio is an example of a population-relevant endpoint which is also indicative for mechanistic effects (EAS mediated). Where sex ratio is affected at a lower test concentration compared with biomarker endpoints, the NOEC for sex ratio will be used as the NOEC for mechanistic effects. For the reasons outlined above, a sufficient dataset is based on a fish study incorporating relevant mechanistic endpoints.

These proposed AFs were then applied to datasets compiled for eight validation substances with suspected endocrine activity (including three with ER agonist activity). Deterministic PNECs have been derived for the validation substances based on the substance-specific maximum available dataset but also, where possible, considerations have been made for a scenario where less data are available. The outcomes from the data assessments of the validation substances have been evaluated as follows:

1. Proposed deterministic PNECs for addressing the ED properties compared with deterministic PNECs derived following ECHA (2008) (i.e. including algae and invertebrate data). The ECHA (2008) PNEC will use the lowest NOEC value for fish;
2. Proposed PNECs for addressing the ED properties derived with the maximum available dataset compared with potential PNECs that would result if less data were available (i.e. PNECs based on single/or limited OECD CF Level 3, 4 or 5 dataset(s)); and
3. Proposed PNECs for addressing the ED properties compared with the driving ED-relevant NOEC based on mechanistic endpoints.

7.2 Alternative proposals for an additional “ED” AF

A recent paper suggests the use of additional default AFs on top of the standard default AFs to account for ED properties (James *et al*., 2023). The paper’s authors argue that even for substances with an EU EQS, the ED properties have not been considered as part of the assessment. The authors propose the application of an additional AF of at least 5 to all confirmed EDCs or at least 10 for EDCs which do not include relevant ED endpoints in the dataset (the authors specifically refer to ED-mediated endpoints). Slightly lower AFs (up to 4) are proposed for suspected EDCs (equivalent to EU CLP category 2, EC, 2023), but these are also in addition to default AFs. There is a caveat that expert judgement can be used when applying the AFs. A decision tree for applying AFs for EDC PNEC derivation adapted from James *et al*. (2023) is given in Appendix 221.

The definition for “ED mediated effects” is not clearly presented by James *et al*. (2023). According to existing regulatory guidance (ECHA/EFSA, 2018 and OECD, 2018), EAS mediated endpoints are generally (although not always, see Marini *et al*., 2023) indicative for an endocrine MoA. These would be, for example, male SSC, specific gonad histopathology, sex ratio or transcriptional activity of cyp19a1b for EAS modalities in relevant fish tests.

Based on the analysis conducted for this report, some observations can be made on the proposals from James *et al*. (2023). Firstly, the inclusion of EAS mediated endpoints does not necessarily mean that a study is sufficiently sensitive for PNEC derivation (e.g. if embryo life stages have not been exposed), and conversely, exclusion of EAS meditated endpoints does not mean that a study cannot be used for PNEC derivation if multigeneration developmental and reproductive endpoints are available. Furthermore, the datasets for NP, OP and EE2 suggest an additional AF may not be justified where the ED properties have already been accounted for in multigeneration studies.

To understand how the proposals from James *et al*. (2023) would impact the PNECs for EDCs, the decision tree in Appendix 221 has been followed for the eight validation substances and compared with the PNEC derived using the approach proposed in Section 82. For each validation substance there is therefore an additional PNEC addressing the ED properties according to James *et al*. (2023). James *et al*. (2023) also requires that at least two endocrine mechanisms (usually EAS and T modalities) have been fully addressed when applying a minimal AF. While this is not necessarily the case for the validation substances, they were chosen based on a specific ED MoA and therefore for the purpose of this exercise it is assumed that this is the only ED concern. All the validation substances are also assumed to be confirmed rather than suspected ED (although a formal regulatory ED assessment may not have been conducted).

7.3 Validation exercise results

The datasets are taken from existing compilations and were not independently assessed for quality as part of this report.

7.3.1 Propyl paraben

Propyl paraben has ER agonist activity. It shows lower potency *in vitro* compared with NP and OP (Table 15). The available ecotoxicity dataset is shown in Table 17 (with additional details in Appendix 226.).

Table 17 NOEC dataset applied for PNEC derivation for propyl paraben

Species	Study type	NOEC (µg/L)
Algae (*Raphidocelis subcapitata*)	Growth inhibition (OECD TG 201)	2 100
Daphnia magna	Reproduction (OECD TG 211)	250
Fish	QSAR[1]	406
Fish Level 3 (*O. latipes*)	FSTRA (OECD TG 229)	< 311 (male VTG induction) 311 (reduction in no. of fertilised eggs)
Fish Level 4 (*D. rerio*)	FSDT (OECD TG 234)	165 (based on sex ratio and % females) ≥ 518 (% males and female VTG content)

[1] Chronic fish toxicity data were based on a quantitative structure-activity relationship (QSAR) model estimates using Ecological Structure Activity Relationships (ECOSAR) Predictive Model. The model is based on acute to chronic ratios and is not expected to account for an ED MoA.

PNECs were derived to account for ED properties by applying the AFs proposed in Table 16 to the NOECs in Table 17. The PNECs are compared with a deterministic PNEC derived according to ECHA (2008), including with an additional AF following the decision tree in James *et al*. (2023). The PNECs are given in Table 18.

Table 18 PNECs for propyl paraben

Basis for PNEC	Assessment factor	PNECs
Lowest NOEC algae, *Daphnia* and fish (ECHA, 2008)	10	FSDT NOEC = 165/10 = **17 µg/L**
James *et al*. (2023)	10 AND 5	FSDT includes sex ratio (EAS-mediated) 17/5 = **3.4 µg/L**

Basis for PNEC	Assessment factor	PNECs
One NOEC from a Level 3 fish study (Level 3 PNEC)	1 000	FSTRA NOEC = 311/1 000 = **0.31 µg/L**
One NOEC/EC_{10} from a Level 4 fish study (Level 4 PNEC)	100	FSDT NOEC = 165/100 = **1.7 µg/L**
NOEC for mechanistic effects		**165 µg/L** (female biased sex ratio in the FSDT)

The PNECs derived for propyl paraben to account for ED properties are more *protective* compared with the PNEC derived according to ECHA (2008), which was also based on the available fish NOEC from the FSDT. A PNEC derived using a FSTRA study is an order of magnitude lower compared with the PNEC derived using the substance-specific maximum available dataset (i.e. using the result from the FSDT). A PNEC of 1.7 µg/L derived using the most comprehensive fish study would appear reasonable compared with the PNECs of 0.064 µg/L for NP and 0.61 µg/L OP, based on the lower potency of propyl paraben for the ER agonist mechanism. Propyl paraben is less potent than NP and OP by two orders of magnitude, based on *in vitro* fish-specific ER receptor binding affinity, and the NOEC from an FSDT is more than 100 µg/L higher than the NOECs for NP and OP from similar FSDT studies (31 µg/L, Horie *et al*., 2021; and 23.7 µg/L, Seki *et al*., 2003, respectively). All PNECs derived for propyl paraben are protective for mechanistic effects. Overall, the approach would appear acceptable for propyl paraben but does not suggest that the AF applied to a Level 4 test could be reduced. A PNEC derived following the James *et al*. (2023) decision tree is less precautionary than the approach proposed in this report.

Since the FSDT has demonstrated an effect on sex ratio which is linked to the ED mechanism it is considered sufficient for assessing the effects of propyl paraben assuming that the (Level 4) AF applied accounts for other potential ED effects which might be detected in a Level 5 study incorporating effects on reproduction.

7.3.2 Methyl paraben

Methyl paraben has ER agonist activity. It shows lower potency *in vitro* compared with NP and OP and propyl paraben (Table 15). The dataset for PNEC derivation is shown in Table 19 (with further details in Appendix 222).

Table 19 NOEC data for PNEC derivation for methyl paraben

Species	Study type	NOEC (µg/L)
Algae (*R. subcapitata*)	Growth inhibition (OECD TG 201)	20 000
Daphnia magna	Reproduction (OECD TG 211)	200
Fish	QSAR[1]	1 190
Fish Level 3 (*O. latipes*)	FSTRA (OECD TG 229)	1 900 (reduction in number of fertilised eggs) 357 (male VTG induction)
Fish Level 4 (*D. rerio*)	FSDT (OECD TG 234)	24 (based on total survival and wet weight of all fish at 70 days post fertilisation) 7.96 (female VTG induction)

[1] Chronic fish toxicity data were based on quantitative structure-activity relationship (QSAR) model estimates using Ecological Structure Activity Relationships (ECOSAR) Predictive Model. The model is based on acute to chronic ratios and is not expected to account for an ED MoA.

PNECs to address ED properties were derived by applying the AFs proposed in Table 16 to the NOECs in Table 19. These PNECs are compared with a deterministic PNEC derived according to ECHA (2008), including with an additional AF following the decision tree in James *et al*. (2023). The PNECs are given in Table 20.

Table 20 PNECs for methyl paraben

Basis for PNEC	Assessment factor	PNECs
Lowest NOEC algae, Daphnia and fish (ECHA, 2008)	10	FSDT NOEC = 24/10 = **2.4 µg/L**
James *et al*. (2023)	10 AND 5	FSDT includes sex ratio (EAS-mediated) 24/50 = **0.48 µg/L**

Basis for PNEC	Assessment factor	PNECs
One NOEC from a Level 3 fish study (Level 3 PNEC)	1 000	FSTRA NOEC = 1 900/1 000 = **1.9 µg/L**
One NOEC/EC_{10} from a Level 4 fish study (Level 4 PNEC)	100	FSDT NOEC = 24/100 = **0.24 µg/L**
NOEC for mechanistic effects		**7.96 µg/L** (VTG induction in the female)

The PNECs derived for methyl paraben to account for ED properties are all *protective* compared with a PNEC derived according to ECHA (2008). A PNEC derived using the FSTRA study is *not protective* compared with the PNEC derived using the result from the FSDT. This is because the FSDT was more sensitive than might be predicted based on ER agonist activity *in vitro* and effects in the FSTRA. A Level 4 PNEC of 0.24 µg/L derived using the FSDT is lower than PNECs for OP and propyl paraben which both show greater potency for the ER *in vitro* (Table 15). Nevertheless, both PNECs would appear protective for mechanistic effects (increased female VTG in the FSDT). Overall, the approach would appear acceptable for methyl paraben although this example suggests that endocrine activity *in vitro* is not necessarily predictive for toxicity *in vivo*. It is worth caveating that the reliability of the study was not independently verified and that there were issues maintaining the test item in solution due to its rapid biodegradability. It is not known whether the population-relevant effects in the FSDT (reduced hatching and survival) are endocrine related or due to another MoA (e.g. general systemic toxicity). For ER agonist activity, increased VTG in females usually occurs at the same or a higher test concentration compared with VTG induction in male fish (Brown *et al*., 2023). It is unclear whether a Level 5 test would have provided more certainty for concluding on a MoA.

7.3.3 3-Benzylidene camphor

3-Benzylidene camphor (3-BC) has ER agonist activity in mammalian cell lines in the range 0.68 – 310 µM (ECHA, 2016) but there are no *in vitro* data from ToxCast for comparison with the other ER agonist substances. The only *in vivo* data for 3-BC are from a Level 3 FSTRA, with a NOEC for reduced reproduction and effects on male VTG and SSC of 33 µg/L (ECHA, 2016). The effect on induction of male VTG can, however, be used as an indicator for potency. The LOEC values for VTG induction in FSTRA style studies are 4.52 µg/L for NP (Zha *et al*., 2008), 10 µg/L for OP (Routledge *et al*., 1998), 74 µg/L for 3-BC (ECHA, 2016), 311 µg/L for propyl paraben (Kawashima *et al*., 2021) and 1 900 µg/L for methyl paraben (Kawashima *et al*., 2021), which suggests that the ER agonist potency for 3-BC is in the mid-range of the five ER agonist substances. The dataset for 3-BC is shown in Table 21 (with further details in Appendix 228).

Table 21 NOEC data for PNEC derivation for 3-Benzylidene camphor

Species	Study type	NOEC (µg/L)
Algae	QSAR[1]	202
Daphnia	QSAR[1]	33
Fish	QSAR[1]	26
Fish Level 3 (*P. promelas*)	FSTRA (OECD TG 229)	33 (reduced reproduction) 0.5 (testis and ovary histology)

[1] Chronic algae, Daphnia and fish toxicity data were based on quantitative structure-activity relationship (QSAR) model estimates using Ecological Structure Activity Relationships (ECOSAR) Predictive Model. The model is based on acute to chronic ratios and is not expected to account for an ED MoA.

PNECs to address ED properties were derived by applying the AFs proposed in Table 16 to the NOECs in Table 21. These PNECs are compared with a deterministic PNEC derived according to ECHA (2008), including with an additional AF following the decision tree in James *et al*. (2023). The PNECs are given in Table 22.

Table 22 PNECs for 3-Benzylidene camphor

Basis for PNEC	Assessment factor	PNECs
Lowest NOEC algae, *Daphnia* and fish (ECHA, 2008)	10	Fish QSAR NOEC = 26/10 = **2.6 µg/L**
James *et al*. (2023)	10 AND 10	No EAS mediated endpoints 26/100 = **0.26 µg/L**
One NOEC from a Level 3 fish study (Level 3 PNEC)	1 000	FSTRA NOEC = 33/1 000 = **0.033 µg/L**
NOEC for mechanistic effects		**0.5 µg/L** (VTG induction in the female)

A deterministic PNEC derived according to ECHA (2008) was based on quantitative structure-activity relationship (QSAR) estimates.

The PNEC derived for 3-BC to account for ED properties is *protective* compared with a PNEC derived according to ECHA (2008). The ECHA (2008) PNEC is based on QSAR model estimates that will not account for ED effects in fish, and this PNEC would not be protective for the effects on VTG in the FSTRA. The application of a default additional AF of 10 following the decision tree in James *et al*. (2023) would be protective for VTG effects.

This PNEC based on the Level 3 FSTRA is lower than the deterministic PNEC for NP of 0.064 µg/L. Based on the *in vivo* mechanistic data for VTG induction NP would be expected to be a more potent ER agonist compared with 3-BC and therefore the Level 3 PNEC for 3-BC might be considered protective. However, there are considerable uncertainties in the assessment, inferred by the lack of data from a higher tier fish test. This example highlights the difficulties in attempting to predict toxicity *in vivo* from the mechanistic data. A higher tier fish test would normally be expected to be necessary for a sufficient dataset with the assumption that confirmatory higher tier testing would be conducted if a risk is identified (unless the risk can be removed through provision of other data, such as environmental degradation half-lives or release estimates).

7.3.4 Fenbuconazole

In the US EPA ToxCast assays fenbuconazole shows predominantly anti-androgenic activity (bioactivity model result = 0.129) and is positive for aromatase inhibition (AC_{50} = 109.7 µM). The dataset for fenbuconazole is shown in Table 23 (with additional details in Appendix 230).

Table 23 NOEC data for PNEC derivation for fenbuconazole

Species	Study type	NOEC (µg/L)
Algae (*R. subcapitata*)	Growth inhibition (OECD TG 201)	25
Daphnia magna	Reproduction (OECD TG 211)	78
Fish Level 4 (*P. promelas*)	FELS (OECD TG 210)	82 (reduced growth)
Fish Level 5 (*P. promelas*)	FLCTT (OPPTS 850.1500)	23 (reproduction)

PNECs to address ED properties were derived by applying the AFs proposed in Table 16 to the NOECs in Table 23. These PNECs are compared with a deterministic PNEC derived according to ECHA (2008), including with an additional AF following the decision tree in James *et al*. (2023). The PNECs are given in Table 24.

Table 24 PNECs for fenbuconazole

Basis for PNEC	Assessment factor	PNECs
Lowest NOEC algae, *Daphnia* and fish (ECHA, 2008)	10	FLCTT NOEC = 23/10 = **2.3 µg/L**
James *et al*. (2023)	10 AND 10	No EAS mediated endpoints 23/100 = **0.23 µg/L**
One NOEC/EC_{10} from a Level 4 fish study (Level 4 PNEC)	100	FELS NOEC = 82/100 = **0.82 µg/L**
One NOEC/EC_{10} from a Level 5 fish study (Level 5 PNEC)	10	FLCTT NOEC = 23/10 = **2.3 µg/L**

The PNECs derived for fenbuconazole to account for ED properties are *the same* as the deterministic PNEC derived according to ECHA (2008), because the result from a multigeneration test with fish is available. The PNEC derived using the FELS test is *more protective* compared with the PNEC derived using the FLCTT due to the additional AF.

Following the decision tree proposed by James *et al*. (2023) an additional AF of 10 would potentially be applied to the ECHA (2008) PNEC to account for ED properties because no EAS mediated endpoints were included in the FLCTT. This is the lowest proposed PNEC. However, fenbuconazole would not necessarily be considered an ED based on the available data because there is no data available to indicate *in vivo* endocrine activity (i.e. confirm the endocrine MoA). Therefore, a fish study with mechanistic data is necessary for a sufficient dataset to confirm the ED properties.

7.3.5 Cyproconazole

Cyproconazole is positive for aromatase inhibition (AC_{50} = 33.9 µM) and inhibits the production of some steroid hormones including oestradiol, based on ToxCast data. It also has some inconclusive ER agonist activity with a bioactivity model result of 0.0268. The dataset used for cyproconazole is shown in Table 25 (with further details in Appendix 232). The available environmental studies for cyproconazole consist of duplicate similar tests for algae, invertebrates and fish with results differing by approximately an order of magnitude. The lowest of the duplicate NOECs have been used for PNEC derivation and are reported in Table 26.

Table 25 NOEC data for PNEC derivation for cyproconazole

Species	Study type	NOEC (µg/L)
Algae (*Desmodesmussubspicatus*)	Growth inhibition (OECD TG 201)	21
Daphnia magna	Reproduction (OECD TG 211)	23
Fish Level 3 (*P. promelas*)	FSTRA (OECD TG 229)	≥ 2 000
Fish Level 4 (*On. mykiss*)	FELS (OECD TG 210)	160
Fish Level 5 (*P. promelas*)	FLCTT (OPPTS 850.1500)	125 (reproduction) 500 (VTG inhibition)

PNECs to address ED properties were derived for cyproconazole by applying the AFs proposed in Table 16 to the NOECs in Table 25. The PNECs are compared with a deterministic PNEC derived according to ECHA (2008), including with an additional AF following the decision tree in James *et al*. (2023). The PNECs are given in Table 26.

Table 26 PNECs for cyproconazole

Basis for PNEC	Assessment factor	PNECs
Lowest NOEC algae, *Daphnia* and fish (ECHA, 2008)	10	Algae NOEC = 21/10 = **2.1 µg/L**
James *et al*. (2003)	10 AND 10	No EAS mediated endpoints 21/100 = **0.21 µg/L**
One NOEC from a Level 3 fish study (Level 3 PNEC)	1 000	FSTRA NOEC = 2 000/1 000 = **2.0 µg/L**
One NOEC/EC_{10} from a Level 4 fish study (Level 4 PNEC)	100	FELS NOEC = 160/100 = **1.6 µg/L**

Basis for PNEC	Assessment factor	PNECs
One NOEC/EC_{10} from a Level 5 fish study (Level 5 PNEC)	10	FLCTT NOEC = 125/10 = **12.5 µg/L**
NOEC for mechanistic effects		**500 µg/L** (inhibition of female VTG in a FLCTT)

The PNECs derived for cyproconazole to account for ED properties using OECD CF Level 3, 4 or 5 data ranged from 1.6 – 12.5 µg/L. The PNEC based on the result from the Level 5 fish study only was potentially not protective for effects in algae since algae were the most sensitive taxa in the dataset. PNECs based on AF applied to Level 3 or 4 data are below the most sensitive PNEC based on algae data and therefore are precautionary. The additional AF of 10 applied based on James *et al*. (2023) would appear to be overly precautionary in this example where fish are relatively insensitive to the effects of cyproconazole.

Despite some evidence for endocrine activity (reduced VTG in females) and reduced reproduction in a FLCTT, algae (and *Daphnia*) are both more sensitive to cyproconazole than fish. The lack of evidence for ED effects in the FSTRA at the highest test concentration of 2 000 µg/L suggests low potency for an endocrine mechanism. This is supported by the toxicity data where the NOEC of 160 µg/L in the FELS, and 125 µg/L in the FLCTT are both much lower than the NOEC of 500 µg/L for inhibition of VTG (also in the FLCTT). This suggests that toxicity may be due to an alternative MoA.

The lack of mechanistic effects of cyproconazole in the FSTRA compared with the FLCTT is a concern as it might suggest that the screening assay is not sufficiently sensitive for highlighting apparent ED activity, especially for substances with low potency. Inhibition of VTG in female fish is an expected outcome for aromatase inhibitors but there are concerns that variability in the VTG measurement may sometimes affect the outcome of a test (EA, 2020a; Brown *et al.*, 2023). In the FSTRA with cyproconazole a 37% reduction in female VTG at the highest test concentration of 1 000 µg/L is not statistically significant (IUCLID, 2024). Supporting evidence from the FLCTT and evidence for aromatase inhibition *in vitro* suggest that this could be a false negative result. Nevertheless, the PNEC derived based on applying an AF of 1 000 to the NOEC from the FSTRA is precautionary compared with the PNEC based on data from the Level 5 study. Further standardisation of the VTG methodology is recommended in Burden *et al*. (2023) to help improve confidence in the VTG result for use as supporting mechanistic data.

It is difficult to define a sufficient dataset for cyproconazole probably because the main MoA may not be as an ED. The mechanistic endpoints reported in the FLCTT are important as context for this conclusion. All the derived PNECs would be protective for the mechanistic effects observed in the FLCTT.

7.3.6 Propiconazole

Propiconazole is positive for aromatase inhibition (AC_{50} = 29.5 μM) and shows AR antagonist activity (bioactivity model result of 0.167). The dataset is shown in Table 27 (with further details in Appendix 236.).

Table 27 NOEC data for PNEC derivation for propiconazole

Species	Study type	NOEC (µg/L)
Algae (*R. subcapitata*)	Growth inhibition (OECD TG 201)	460
Daphnia magna	Reproduction (OECD TG 211)	310
Amphibian Level 3	AMA (OECD TG 231)	56 (SVL and weight)
Fish Level 3 (*P. promelas*)	FSTRA (OECD TG 229) or Fish Screening Assay (OECD TG 230)	Study 1: 120 (female VTG inhibition, decreased reproduction and decreased female survival) Study 2: 53 (fecundity) Study 2: 5.8 (female VTG inhibition)
Fish Level 4 (*P. promelas*)	FELS (OECD TG 210)	EC_{10}: 380 (survival)
Fish Level 5 (*P. promelas*)	FLCTT (OPPTS 850.1500)	EC_{10}: 119 (reduced egg production) 7.8 (increased F1 male nuptial tubercle score)
Fish Level 5 (*C. variegatus*)	FLCTT (OPPTS 850.1500)	68 (reproduction)

PNECs were derived to account for ED properties by applying the AFs proposed in Table 16 to the NOECs in Table 27. The PNECs are compared with a freshwater deterministic PNEC derived according to ECHA (2008), including with an additional AF following the decision tree in James *et al*. (2023). The dataset for propiconazole also includes data for marine species. The PNECs are given in Table 28.

Table 28 PNECs for propiconazole

Basis for PNEC	Assessment factor	PNECs
Lowest NOEC algae, *Daphnia* and fish (ECHA, 2008) (freshwater species)	10	FLCTT NOEC (*C. variegatus*) = 68/10 = **6.8 µg/L**
James *et al.* (2008)	10 AND 5	FLCTT (*P. promelas*) includes sex ratio and SSC AMA (*X. laevis*) includes T-mediated endpoints 68/50 = **1.36 µg/L**
One NOEC from a Level 3 fish study (Level 3 PNEC)	1 000	FSTRA NOECs = 120 or 53/1 000 = **0.12 or 0.053 µg/L**
One NOEC/EC_{10} from a Level 4 fish study (Level 4 PNEC)	100	FELS NOEC = 380/100 = **3.8 µg/L**
One NOEC/EC_{10} from a Level 5 fish study (Level 5 PNEC)	10	FLCTT NOEC = 68 (*C. variegatus*)/10 = **6.8 µg/L**
NOEC for mechanistic effects		**5.8 µg/L** (VTG inhibition in a FSTRA).

The PNECs derived for propiconazole to account for ED properties based on OECD CF Level 3 and 4 studies are *protective* compared with the PNEC derived according to ECHA (2008). Interestingly, the FSTRA both generated lower NOECs compared with the Level 4 FELS with the same species (*P. promelas*), resulting in arguably over precautionary PNECs. This is because reproduction in the FSTRA was more sensitive than the endpoints for development and growth in the FELS test for this substance. The OECD CF Level 5 PNEC was the same as the ECHA (2008) PNEC because they are both based on the NOEC from the OECD CF Level 5 FLCTT test with *C. variegatus*.

Confidence in the ED PNEC is increased because a result from a second FLCTT study is also available. A PNEC of 6.8 µg/L is slightly above the lowest NOEC for mechanistic effects of 5.8 µg/L based on inhibition of VTG in female fish, in a FSTRA with test concentrations separated by a factor of 10.

The AMA was not used for PNEC derivation as this is a Level 3 screening study with wide spacing of test concentrations and therefore is not designed for this purpose. There were no effects on T-mediated endpoints in the AMA. This is the only validation substance where the EAS and T pathways have both been investigated. Nevertheless, following the decision tree of James *et al*. (2023) an additional AF of 5 is applied to the ECHA (2008) PNEC. This additional AF does not seem to be justified given that the dataset includes results from two FLCTT covering the relevant EAS mediated pathway.

Reproduction (in both the FSTRA and FLCTT) were important endpoints for assessing the ED effects of propiconazole.

7.3.7 Prochloraz

Prochloraz is a well-studied substance which has been included in OECD test method validation exercises as a model substance for effects via aromatase inhibition. In ToxCast the AC_{50} for aromatase inhibition of prochloraz is 1.51 μM which is more potent compared with other substances included in this review (i.e. cyproconazole AC_{50} = 33.9 μM and propiconazole AC_{50} = 29.5 μM). It also has anti-androgenic activity with a bioactivity model result of 0.295 and reduces levels of steroid hormones including testosterone and oestradiol in an *in vitro* steroidogenesis assay. The dataset for prochloraz is shown in Table 29 (with additional details in Appendix 241).

Table 29 NOEC data for PNEC derivation for prochloraz

Species	Study type	NOEC (μg/L)
Algae	Growth inhibition (OECD TG 201)	3.2
Daphnia magna	Reproduction (OECD TG 211)	22.2
Fish Level 3 (*O. latipes*)	FSTRA (OECD TG 229)	20.6 (female VTG inhibition and fertility)
Fish Level 4 (*P. promelas*)	FELS (OECD TG 210)	48.5
Fish Level 4 (*P. promelas*)	FSDT (OECD TG 234)	96 – 106 (sex ratio) < 29 – 31 (VTG inhibition)
Fish Level 4 (*D. rerio*)	FSDT (OECD TG 234)	48 - < 60 (sex ratio) 15 - 60 (VTG inhibition)

Species	Study type	NOEC (µg/L)
Fish Level 5 (*O. latipes*)	MEOGRT (OECD TG 240)	17.5 (fecundity) 5.3 (decreased growth although biological relevance unclear) 5.3 (VTG in females and SSC in males)
Fish Level 5 (*P. promelas*)	FLCTT (OPPTS 850.1500)	24.9 (reproduction)

The NOEC to be used in PNEC derivation from the MEOGRT is unclear, since there are available NOECs of 17.5 µg/L based on fecundity, and 5.3 µg/L based on reduced growth of sub-adult fish. The biological/population relevance for reduced growth was considered by the authors of the paper to be equivocal because the effect was slight and not observed in adult fish (Flynn *et al*., 2017). PNECs to account for ED properties were derived for prochloraz by applying the AFs proposed in Table 16 to the NOECs in Table 29. The PNECs are given in Table 30. The PNECs were derived using both NOECs generated in the MEOGRT.

Table 30 PNECs for prochloraz

Basis for PNEC	Assessment factor	PNECs
Lowest NOEC algae, *Daphnia* and fish (ECHA, 2008) (freshwater species)	10	Algae NOEC = 3.2/10 = **0.32 µg/L**
James *et al*. (2008)	10 AND 5	MEOGRT test available covering EAS mediated endpoints 3.2/50 = **0.064 µg/L**
One NOEC from a Level 3 fish study	1 000	FSTRA NOECs = 20.6/1 000 = **0.021 µg/L**
One NOEC/EC$_{10}$ from a Level 4 FELS study	100	FELS NOEC = 48.5/100 = **0.49 µg/L**

Basis for PNEC	Assessment factor	PNECs
One NOEC/EC_{10} from a Level 4 FSDT study	100	FSDT NOEC = 48 – 96/100 = **0.48 – 0.96 µg/L**
One NOEC/EC_{10} from a Level 5 fish study	10	FLCTT NOECs = 24.9 (*P. promelas*) or MEOGRT NOECs 5.3/17.5 (*O. latipes*)/10 = **2.49 or 0.53/1.75 µg/L**
NOEC for mechanistic effects		**5.3 µg/L** (VTG inhibition in females and SSC in males in the MEOGRT).

Some of the PNECs derived based on OECD CF Level 3, 4 or 5 fish studies for prochloraz were *not protective* compared with that derived using ECHA (2008). This is because, despite the well documented endocrine effects of prochloraz, algae are the most sensitive species for this substance. The PNEC derived using a Level 3 FSTRA study is *protective* compared with all the other PNECs by approximately one order of magnitude. Similar to propiconazole the FSTRA was more sensitive to prochloraz compared with the FELS test (and also for prochloraz the FSDT). As the FSDT is an extended FELS test there is an assumption that the result from this study might be most sensitive. However, for *P. promelas* the NOEC from a FELS test is more sensitive compared with two FSDT with the same species. The FELS test with *P. promelas* has broadly equivalent sensitivity to the FSDT with *D. rerio*. The PNECs derived by applying AFs to the lowest NOECs from the FELS or FSDT studies were *more protective* compared with the PNEC derived using the results from the Level 5 studies. Reproductive endpoints were the most sensitive in the Level 5 fish studies suggesting that reproduction is more sensitive for the aromatase inhibition MoA. A PNEC derived using either Level 5 fish study would be protective for the lowest NOEC based on mechanistic data (VTG inhibition in females and SSC in males in the MEOGRT). This suggests that a sufficient dataset for prochloraz should include the result from a Level 5 test and that application of an AF of 10 was suitable regardless of the study used for deriving the PNEC.

Even though Level 5 fish data is available for prochloraz, the PNEC derived according to ECHA (2008) is the most sensitive and it is recommended should be applied in order to ensure protection for algae, however the actual PNEC used may depend on the regulatory context. Similar to propiconazole the application of an additional AF of 5 following the decision tree of James *et al*. (2023) would not seem to be justified given that the dataset includes results from two FLCTT covering the relevant EAS mediated pathway.

7.3.8 Vinclozolin

Vinclozolin's main mechanism of endocrine activity is as an anti-androgen (bioactivity model result = 0.416). The ToxCast *in vitro* steroidogenesis data suggest that it will inhibit some steroid hormones, including testosterone. The effect on aromatase inhibition is not considered to be reliable due to data flags and possible cytotoxicity. The dataset for vinclozolin is shown in Table 31 (with further details in Appendix 245). Vinclozolin was used as a validation substance for the Medaka Multigeneration Test (MMT) and showed low sensitivity for population-relevant effects in these three studies (NOECs between 197 and 360 µg/L). This suggests that the MMT was not sufficiently sensitive for detecting the population-relevant effects of vinclozolin, and hence the test design was refined and replaced with the MEOGRT (which is now a validated OECD test method (OECD TG 240)). The main difference between the MMT and the MEOGRT is that the number of control replicates used for measuring reproduction was increased from 6 to 12 pairs to improve the statistical power for detecting reproductive effects. The MMT studies are therefore not reliable for ED PNEC derivation but can be used as supporting information (e.g. for mechanistic effects). This provides a further example of test design affecting the study outcome.

Table 31 NOEC data for PNEC derivation for vinclozolin

Species	Study type	NOEC (µg/L)
Algae	Growth inhibition (OECD TG 201)	NOEC: 1 020 growth rate
Daphnia magna	Reproduction (OECD TG 211)	790
Fish Level 3 (*O. latipes*)	FSTRA (OECD TG 229)	Study 1: 137 (female VTG inhibition and fertility) Study 2: 189 (reduced female VTG and fertility)
Fish Level 3 (*P. promelas*)		178 (based on ↓ female GSI and delayed gonadal development and ↑ male E2)
Fish Level 3 (G. *aculeatus*)	AFSS	NOEC/IC_{10}: 4.25 – 25 (based on ↓ spiggin)
Fish Level 5 (*O. latipes*)	MEOGRT (OECD TG 240)	NOEC: 17.3 (based on F1 and F2 larvae survival, ↑ F0 female SSC and ↓ F2 male SSC) < 17.3 (based on ↓ F1 male SSC)

Species	Study type	NOEC (µg/L)
Fish Level 5 (*O. latipes*)	MMT lifecycle validation	Three validation tests without significant effects on population-relevant endpoints at the highest tested concentrations of 197, 333 or 360. Nevertheless, in two of the three studies effects on male SSC were reported at the lowest test concentrations of 5 and 9
Fish Level 5 (*P. promelas*)	FLCTT (OPPTS 850.1500)	50 (based on F0 fecundity and delayed gonadal development and F1 survival, ↓ F1 length and F1 sex ratio)

PNECs to address ED properties were derived by applying the AFs proposed in Table 16 to the NOECs in Table 31. The PNECs are given in Table 32. There are two reliable Level 5 tests, a MEOGRT (NOEC = 17.3 µg/L based on reduced embryo survival) and a FLCTT with *P. promelas* (NOEC = 50 µg/L based on reduced reproduction). These PNECs are compared with a deterministic PNEC derived according to ECHA (2008), including with an additional AF following the decision tree in James *et al*. (2023).

Table 32 PNECs for vinclozolin

Basis for PNEC	Assessment factor	PNECs
Lowest NOEC algae, *Daphnia* and fish (ECHA 2008) (freshwater species)	10	Fish NOEC (*O. latipes*, MEOGRT) = 17.3/10 = **1.7 µg/L**
James *et al*. (2023)	10 AND 5	EAS endpoints in a MEOGRT and FLCTT 17.3/50 = **0.34 µg/L**
One NOEC from a Level 3 fish study (ED PNEC)	1 000	FSTRA NOECs = 137 - 189/1 000 = **0.14 – 0.19 µg/L**
One NOEC/EC_{10} from a Level 5 fish study (ED PNEC)	10	FLCTT NOECs = 50 (*P. promelas*) or MEOGRT NOECs 17.3 (*O. latipes*)/10 = **5.0 or 1.73 µg/L**

Basis for PNEC	Assessment factor	PNECs
NOEC for mechanistic effects		**IC_{10} = 4.25 µg/L** (decreased spiggin in an AFSS) or **NOEC < 5** based on F2 male (phenotypic) SSC in an MMT.

PNECs from OECD CF Level 3 studies were at least as protective asthe PNEC derived following ECHA (2008), which was based on the most sensitive fish study in the dataset (the MEOGRT NOEC of 17.3 µg/L). Since the Level 5 MEOGRT is the most sensitive result then the Level 5 PNEC based on the NOEC from this study is the same as for ECHA (2008) PNEC.

The NOEC from the Level 3 FSTRA were both more sensitive compared with the results from the MMT. The PNECs derived based on the FSTRA are *protective* compared with the PNECs derived using NOECs from the MEOGRT or the FLCTT with *P. promelas*. However, the Level 5 PNECs may *not be protective* for mechanistic effects. This is because the NOECs for effects on male SSC are unbounded (< 17.3 µg/L in the MEOGRT; ≤ 9 µg/L in the MMT), and therefore an ETNC has not been defined based on mechanistic endpoints. Mechanistic endpoints were not measured in the FLCTT. This does not mean that the proposed AF are not viable, but does highlight the need for high quality fish data with robust supporting mechanistic data.

It could be argued that the additional AF of 5 applied using the decision tree in James *et al*. (2023) could be justified in the present example, where there are some uncertainties based on evidence for mechanistic effects at concentrations below the proposed PNEC There should potentially be some flexibility in the proposed approach to use expert judgement to add additional AF where there is evidence that a PNEC may not be sufficiently protective for mechanistic effects.

The outcomes for this substance also suggest that the SETAC Pellston Workshop® criteria should also be considered when applying a deterministic approach to deriving a PNEC (as well as for a probabilistic approach). Nevertheless, a sufficient dataset for vinclozolin could consist of a Level 5 fish study if the test design was sufficiently robust to detect a NOEC for mechanistic changes.

8. Synthesis of the validation exercise

In this section, we consider the number and type of fish studies that are required to derive a reliable PNEC for ED properties focussed on EAS modalities, based on the findings in the previous sections.

8.1 Review of PNECs for validation substances

Three of the eight validation substances have ER agonist activity (methyl paraben, propyl paraben and 3-BC), and the rest have other EAS activity (fenbuconazole, cyproconazole, propiconazole, prochloraz, and vinclozolin). The substances have variable amounts of available data with which to conduct the assessments.

A summary of the PNECs that have been derived for ER agonist substances based on different approaches is given in Figure 7. It includes data for NP and OP and is presented on a logarithmic scale to account for the different concentration ranges. Data for EE2 were not included because of the lack of standardised studies. However, it is notable that a PNEC of 0.005 ng/L for EE2 is at least three orders of magnitude lower than any PNEC derived for these chemicals (based on any of the approaches followed).

It is worth noting that fish studies incorporating ED relevant endpoints can be challenging tests to conduct and interpret. Even when studies are considered reliable without restrictions there may be differences between NOEC values from different tests, even for the same species. Reasons for differences include inter-laboratory variation, inherent biological variability, differences in concentration spacing in the various studies, and the endpoints measured in the study. Furthermore, no independent assessment of study reliability was conducted for the validation substances for the purposes of this report. These factors therefore could affect the PNECs derived and therefore any comparisons made between them. Nevertheless, the available PNECs are still considered to be useful as a guide for assessing the suitability of the proposed AF(s).

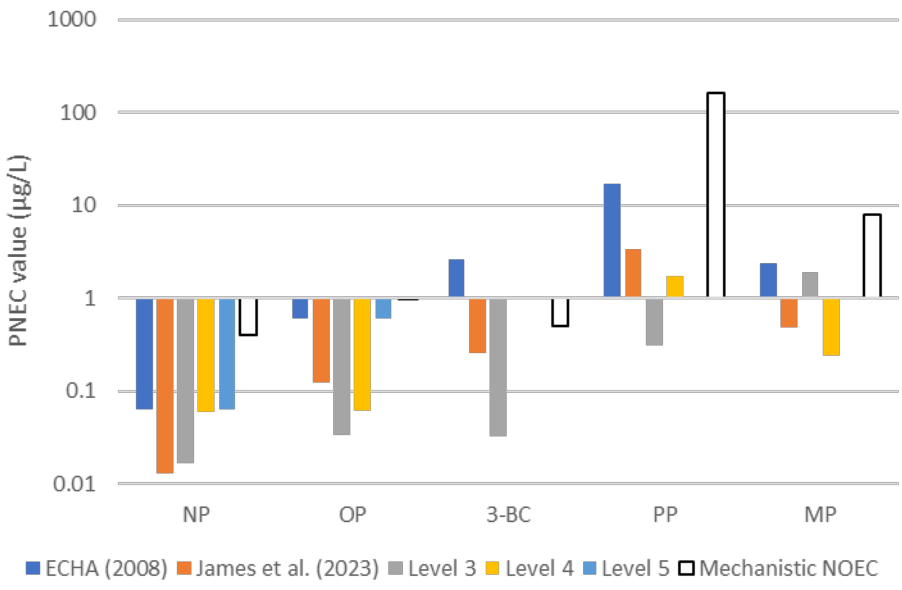


Figure 7 A comparison of PNECs derived using different methods for substances that are ER agonists

Note: PP = propyl paraben, MP = methyl paraben. The mechanistic NOEC for OP is 1 µg/L.

Figure 8 provides an overview of the PNECs derived for substances affecting other EAS modalities. Again, it is presented on a logarithmic scale to account for the different concentration ranges.

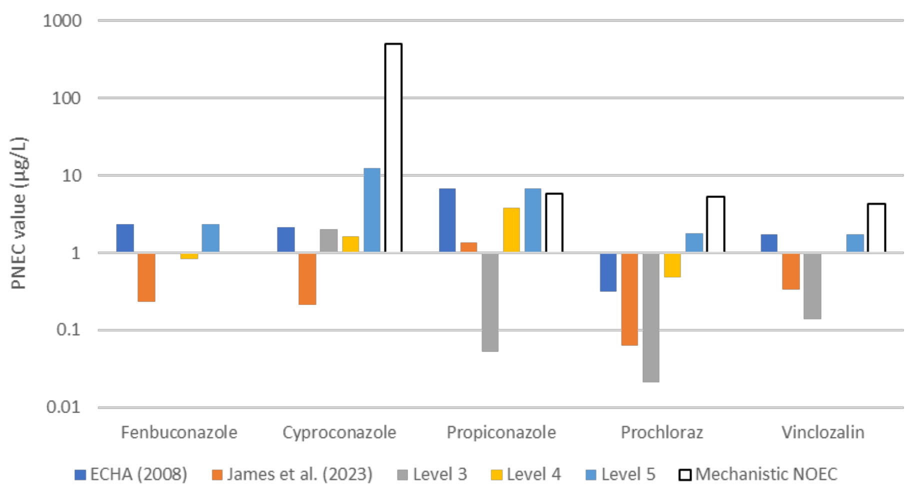


Figure 8 A comparison of PNECs derived using different methods for substances with EAS mechanisms of action other than ER agonist

8.1.1 How useful were PNECs derived using Level 3 NOECs?

The FSTRA is a mechanistic screening study in Level 3 of the OECD CF. It usually has only three test concentrations which are widely spaced (typically a factor 10). There is generally low replication for detecting reproductive effects in such studies. FSTRA studies would not usually be considered relevant for PNEC derivation because they possess low resolution and sensitivity for population level effects. Nevertheless, population-relevant endpoints (reproduction) are measured in the FSTRA. The study design also allows for endocrine mechanistic effects to be determined, and for some substances (e.g. 3-BC) it represents the only available fish data. A highly precautionary AF of 1 000 is proposed to be applied to the NOEC from a FSTRA to derive a viable ED PNEC, in the absence of other relevant data. This approach is considered analogous to using acute toxicity data for deriving a PNEC, in the absence of chronic data, as described in ECHA (2008). Level 3 data are not proposed to be used for refining a PNEC, i.e. when Level 4 or 5 data are available.

A FSTRA has been conducted for seven of the nine validation substances plus there are data for NP and OP (resulting in a Level 3 PNEC of 0.0165 µg/L for NP based on reproductive effects in a FSTRA with *O. latipes* (Ishibash *et al*., 2006), and a PNEC of 0.0336 µg/L for OP based on egg viability in a FSTRA with *C. variegatus* (Karels *et al*., 2003)). The PNECs derived using Level 3 data were always more precautionary compared with PNECs derived according to ECHA (2008), using Level 5 fish data or the mechanistic NOEC (except for vinclozolin where the NOEC for mechanistic effects is unbounded), usually with a large margin of safety (Figure 7 and Figure 8). For methyl paraben, there was no Level 5 fish study and the FSTRA PNEC was *less precautionary* compared with the PNEC derived using the NOEC from the Level 4 FSDT. Effects in the FSDT occurred at a test concentration lower than might be predicted based on ER agonist activity *in vitro* (in the Bioactivity model or in a medaka specific ERα assay) or *in vivo* in the FSTRA, resulting in a PNEC that was lower than expected (i.e. lower than propyl paraben which is predicted to have more potent ER agonist activity). As the NOEC from the FSDT with methyl paraben is based on survival and growth endpoints (compared with an effect on sex ratio for propyl paraben) it may be that an alternative (non-ED) MoA is involved. The results might also be affected by quality issues that have not been reviewed for the purposes of this report (for instance, there were issues maintaining test solution concentrations in the FSDT for methyl paraben as discussed in Section 87).

Even though no population level or mechanistic effects were reported in the FSTRA for cyproconazole at the highest concentration of 2 000 µg/L (possibly due to low sensitivity of the test), the PNEC based on the NOEC from this study is still protective compared with the PNEC based on the result from a FLCTT. The NOEC used is, however, an unbounded (greater-than) result.

It is worth highlighting that, despite the screening nature of the assay, the Level 3 FSTRA was more sensitive compared with Level 4 studies for propiconazole (FELS test) and prochloraz (FELS test or FSDT) (Table 27 and Table 29, respectively), both of which are aromatase inhibitors and anti-androgenic *in vitro*. Reproduction would be expected to be a sensitive endpoint based on the aromatase inhibition MoA, and none of the Level 4 studies available assessed reproduction for these two substances. Aromatase inhibition *in vitro* is also associated with VTG inhibition *in vivo* (Brown *et al*., 2023). Nevertheless, cyproconazole, which is also an aromatase inhibiter *in vitro*, was insensitive in the FSTRA and therefore caution should be applied when interpreting the data from this study for PNEC derivation.

As a PNEC based on a FSTRA is only intended to be a precautionary approach in the absence of other data, there is no need to change the proposed AF of 1 000. It is highly unusual for an SVHC to be identified solely on the basis of OECD CF Level 3 data, but it has happened in the past and might be relevant if a read across argument is being considered. It is therefore important to have a mechanism to calculate a PNEC for such substances should they be added to REACH Annex 14 or subject to a restriction proposal in future.

8.1.2 How comparable were PNECs derived using Level 4 versus Level 5 NOECs?

The datasets for NP and OP do not provide much clarity on the sufficiency of using a Level 4 study to derive an ED PNEC. For OP, the results from FSDT studies were more sensitive compared to Level 5 studies with the same species, whereas for NP the Level 5 MEOGRT was more sensitive compared to Level 4 FELS studies with two different fish species. There were three FSDTs available for NP, all with *O. latipes;* the lowest NOEC is 11.6 µg/L (Seki *et al*., 2005) which is considerably higher than the NOEC of 0.64 µg/L from the MEOGRT (Watanabe *et al*., 2017). In fact, FELS studies with *P. promelas* (NOEC = 7.4 µg/L, Ward and Boeri, 1991a) and *On. mykiss* (NOEC = 6 µg/L, Spehar *et al*., 2010) both had lower NOECs for NP compared with the FSDTs with *O. latipes*. The available data from EE2 demonstrate that exposure during the sensitive window for sexual differentiation is important for assessing the effects of an ER agonist substance, but as both the FELS test and FSDT expose embryos they are expected to cover this.

An FSDT would be expected to be more sensitive compared with the FELS test, because it has a longer duration (at least for the same species) and includes endpoints for sexual development. According to OECD GD 150, the preferred study to use for investigating a positive result in a FSTRA is the MEOGRT or, when available, the Zebrafish extended one generation test (OECD, 2018). However, the FSDT is recommended where sexual development is expected to give a response at lower concentrations than reproduction (although it is unclear how this would be assessed). In OECD validation the FSDT was responsive for an ER agonist mechanism (based on studies with OP, 4-tert-pentylphenol and 17β-oestradiol) and AR agonist mechanism (based on studies with dihydrotestosterone (DHT)) (OECD, 2012), although relative sensitivity compared with effects in other test designs is not discussed. In addition, the FSDT detected effects on sex ratio of the aromatase inhibitor prochloraz. None of the OECD validation studies with the AR antagonist flutamide were reliable and therefore it is unknown whether the FSDT can detect this MoA, as discussed in Panter *et al*. (2023).

Table 33 compares the PNECs derived from FELS or FSDT studies with PNECs derived from datasets which include Level 5 studies. In all cases the factor of 100 applied to the Level 4 FELS result would be considered acceptable. Only the PNEC for OP would seem overly protective but this is because the result from a Level 4 FELS test was more sensitive compared with the Level 5 tests. All PNECs are also protective for mechanistic effects, except for propiconazole. For propiconazole the lowest NOEC for mechanistic effects was from a FSTRA with a factor of 10 spacing of test concentrations and therefore the NOEC value from this study is conservative.

Table 33 PNECs from FELS and FSDT studies compared with PNECs from FLCTT

Substance	PNEC Level 4 NOEC[1] (µg/L)	PNEC Level 5 NOEC[2] (µg/L)	Lowest NOEC for mechanistic effects[3] (µg/L)
NP	0.06 – 0.076 (FELS) 0.12 (FSDT)	0.064	0.4
OP	0.061 (FELS) 0.074 (FSDT)	**0.61**	1
Fenbuconazole	0.82 (FELS)	2.3	-
Cyproconazole	1.6 (FELS)	12.5	500
Propiconazole	3.8 (FELS)	6.8-11.9	5.8
Prochloraz	0.49 (FELS) 0.48 – 0.96 (FSDT)	0.53 – 2.49	5.3

[1] ED PNEC derived by applying a factor of 100 to the NOEC from a Level 4 test.
[2] ED PNEC derived by applying a factor of 10 to the NOEC from the most sensitive result from a Level 4 or 5 test.
[3] The lowest NOEC from any study regardless of OECD CF Level.
Numbers in bold are where the Level 4 NOEC was used to derive the PNEC. This was only the case for OP.

Four of the validation substances had both Level 4 and Level 5 data, and for three of these substances the Level 4 studies were FELS studies rather than FSDTs. For prochloraz, a NOEC of 48.5 µg/L based on growth in a FELS study with *P. promelas* was more sensitive than the effects determined in FSDTs with the same species (NOEC = 96 to 106 µg/L based on sex ratio), and similar in sensitivity to the NOECs of 48 – < 60 µg/L in FSDTs with *D. rerio* (also based on sex ratio). There are likely to be few other substances which have available data for an FSDT and a Level 5 test, because both are assumed relevant for PNEC derivation and demonstrating an ED MoA, and therefore only one or the other is required for regulatory purposes. For example, methyl paraben and propyl paraben have results from an FSDT but not a Level 5 test, whereas fenbuconazole, cyproconazole, propiconazole, and vinclozolin have Level 5 data but not an FSDT.

Overall, where FELS and FSDT data are available, the sensitivity in these studies appears to be similar (differences may be explained by laboratory or species differences or spacing of test concentrations). These findings suggest that an AF of 100 would be both proportionate and protective when deriving a PNEC for ED properties from the result of a Level 4 FELS test or FSDT. However, although a PNEC to address ED properties can be derived based on the result from a FELS study, it does not include mechanistic endpoints which are important for interpreting the ED outcomes. Therefore, the FSDT is the preferred Level 4 test for a substance with an ED concern.

The application of an AF of 100 to the NOEC from a FELS study or a FSDT could be considered equivalent to the proposals from James *et al*. (2023), who recommend applying an additional AF of 10 where no ED data are available.

8.1.3 How do PNECs derived using Level 3, 4 or 5 data compare to deterministic PNECs derived according to ECHA (2008)?

The deterministic PNECs (based on the substance-specific maximum available dataset for each validation substance) have been compared with a PNEC derived according to ECHA (2008) (Figure 7 and Figure 8). The following scenarios were identified:

- Where there is a Level 5 fish study and fish are the most sensitive taxon, the PNEC derived according to ECHA (2008) and the Level 5 PNEC were identical (NP, OP, fenbuconazole, propiconazole and vinclozolin);
- Where there is a Level 5 fish study and algae were the most sensitive taxon, the PNEC derived according to ECHA (2008) was lower than the Level 5 PNEC (cyproconazole and prochloraz); and
- Where only Level 4 fish data are available, the PNEC derived to account for ED properties was lower than the PNEC derived according to ECHA (2008) due to the additional AF of 10 proposed to be applied to the Level 4 test.

There were scenarios where the ECHA (2008) PNEC was not protective for the NOEC for mechanistic effects, most notably for 3-BC where the chronic dataset was based on QSAR estimates. The other exception was propiconazole (where the NOEC for mechanistic effects is just above the PNEC derived from the Level 5 fish studies based on ECHA 2008). Otherwise, all ECHA (2008) PNECs were protective of the mechanistic NOEC.

AFs applied according to ECHA (2008) are intended to account for a range of different uncertainties for a huge range of chemicals, including inter- and intra-laboratory variation, extrapolation from laboratory to field, and the very narrow range of species usually present in the available ecotoxicity dataset. It is clear that an AF of 10 applied to the standard chronic dataset of algae, *Daphnia* and a FELS study for fish would not be protective for ED properties, because a multigeneration study covering both developmental and reproductive endpoints is required to fully investigate ED properties. Species differences are covered in the next section.

8.1.4 Are species differences accounted for?

There are data from Level 5 tests with more than one fish family for three of the validation substances (propiconazole, prochloraz and vinclozolin) as well as EE2 and OP. These data are shown in Table 34.

Table 34 Comparison of NOEC data from Level 5 tests with different species

Species	OP (µg/L)	EE2[a] (ng/L)	Propiconazole (µg/L)	Prochloraz (µg/L)	Vinclozolin (µg/L)
***O. latipes* (MOEGRT)**	25.1	-	-	5.3 or 17.5	17.3
***P. promelas* (FLCTT)**	-	0.76	119	24.9	50
***D. rerio* (FLCTT)**	12	0.05, 0.1, 0.31	-	-	-
***On. mykiss* (partial FLCTT)**	-	0.4	-	-	-
***C. variegatus* (FLCTT)**	-	-	68	-	-
Factor difference	**2.1**	**15.2**	**1.8**	**4.7 or 1.4**	**2.9**

[a] Studies where embryo life stages were not exposed are not included due to their low relevance for the ER agonist MoA.

Both the MEOGRT study with NP (Section 24; Watanabe *et al.*, 2017, Appendix 179) and the MEOGRT for prochloraz, where a NOEC of 5.3 µg/L based on reduced growth is considered equivocal because the effect was slight illustrate the difficulty of interpretation of these studies. For the latter, while effects were observed in the sub-adult fish, these were not reported for the equivalent adult fish (Flynn *et al.*, 2017). The experience of the authors of this report in study monitoring MEOGRTs also suggests that the performance of such studies requires careful optimisation in some areas on a laboratory-specific basis, and some laboratories can have particular issues with obtaining sufficiently high hatching of eggs to meet the test validity criteria. Nevertheless, the MEOGRT reliably detects ED effects for prochloraz and vinclozolin (Flynn *et al*., 2017), as well as for NP (Watanabe *et al*., 2017), and has an experimental design that is sensitive for the effects of EDCs.

Although based on just five substances with only two or three species (also from different families), the factor difference between the Level 5 studies is < 5, except for EE2. The EE2 dataset should however be interpreted with caution due to the variable experimental designs of the fish studies and the large spacing between test concentrations (as discussed in Section 68).

As discussed in Section 77, data are not available to assess the sensitivity of fish species with alternative reproductive strategies, long-lived fish species and marine species. In addition, any proposed PNECs also need to be protective of other aquatic taxa such as amphibians. There are currently no standardised multigeneration OECD test guidelines covering such species for a targeted comparison of species sensitivity. Where data are available for amphibians and fish with alternative reproductive strategies, they were not the most sensitive in the dataset (although this is based on a limited number of studies and may be related to test design since no multigeneration studies were available). For effects on longer lived fish species, the available data regarding sensitivity is inconclusive. For OP a level 4 FELS test with *On. mykiss* (ABC labs, 1984, cited in EA 2005 and ECHA 2011) is more sensitive compared with a Level 5 MEOGRT study (Flynn *et al*., 2017), whereas for NP a Level 5 study with *On. mykiss* (Schwaiger *et al*., 2002, used as supporting information only) was similar in sensitivity to the MEOGRT (Watanabe *et al*., 2017). However, for EE2, a Level 5 study with *D. rerio* was more sensitive (Larsen *et al*., 2008) compared with a Level 5 test with *On. mykiss* (Brown *et al*., 2007). For marine verses freshwater species, *C. variegatus* (marine) was more sensitive to propiconazole compared with *P. promelas* (freshwater) in Level 5 FLCTTs, but only by a factor of < 2.

Based on the data reviewed in this report, an AF of 10 should be adequate to account for differences between species and families of fish (assuming that the fish studies follow robust standard Level 5 fish test designs), and there appears to be minimal benefit to increasing the dataset to include multiple additional fish species.

Given that the proposed AFs are already designed to be precautionary for mechanistic effects, it seems justified that a sufficient dataset can be based on a result from a single Level 5 fish study which covers the whole life cycle, including the sensitive window for sexual differentiation and reproductive effects, with an AF of sufficient magnitude to account for species differences. This obviously assumes that fish are the most sensitive species, which is not always the case as demonstrated for cyproconazole and prochloraz. Therefore, the recommended dataset for an EDC should also include the standard data requirements for algae and invertebrates so that where fish are not the most sensitive, a PNEC can be derived using the approach followed for industrial chemicals as detailed in ECHA (2008).

We specifically make no recommendation for a preferred Level 5 fish study, or species. In fact, there are advantages in having scope to target the experimental design to detect outcomes based on an expected MoA and/or species of concern. Nevertheless, testing should ideally follow regulatory accepted test designs (i.e. OECD TG 240 or the Office of Prevention, Pesticides and Toxic Substances (OPPTS) fish life cycle toxicity test guideline 850.1500 (US EPA, 1996)). These provide the principles for a test exposing fish from one stage of the life cycle to at least the same stage of the next generation, covering sensitive embryo life stages and the window for sexual differentiation as well as reproductive endpoints. Mechanistic endpoints should also be included to provide context.

8.1.5 Were ED PNECs protective for mechanistic effects?

As discussed in Section 82, seven of the eight validation substances had *in vivo* mechanistic data (only fenbuconazole did not – see Section 91). For five of the substances, the lowest mechanistic NOEC came from the following studies:

- Propyl paraben: female biased sex ratio in the FSDT (Level 4) – same value as the adverse effect in the FSDT.
- Methyl paraben: VTG induction in the female FSDT (Level 4) – lower than the adverse effect in the FSDT.
- 3-Benzylidene camphor: VTG induction in the female FSTRA (Level 3) – lower than the adverse effect from the Level 3 study used for PNEC derivation.
- Cyproconazole: inhibition of female VTG in a FLCTT (Level 5) – higher than the adverse effects in the Level 5 study.
- Prochloraz: VTG inhibition in females and SSC in males in the MEOGRT (Level 5) – lower than the adverse effects in the Level 5 study.

The ED PNECs derived were lower than the NOEC based on mechanistic effects for all of these substances, and this can be seen in the summaries provided by Figure 7 and Figure 8.

There were two substances where this was not the case:

- The PNEC for propiconazole of 6.8 μg/L was slightly above the lowest NOEC for VTG inhibition from a (Level 3) FSTRA of 5.8 μg/L, but this was considered an artifact of the wide spacing of test concentrations in this study where the LOEC for effects on VTG was 53 μg/L. The PNEC was protective of mechanistic effects from Level 5 studies.
- For vinclozolin, none of the Level 5 studies with *O. latipes* gave a NOEC for endocrine activity based on reduced male SSC. Population level effects appear to occur at much higher test concentrations compared with the mechanistic effects, and it is assumed that if a lower concentration range was tested, a NOEC based on SSC would be derived. Nevertheless, the PNEC for vinclozolin (based on results from two Level 5 studies with different species) is an order of magnitude lower compared with the LOEC for effects on SSC.

The possibility of including an additional AF in the PNEC derivation where it is not adequately protective for mechanistic effects is discussed in Section 115.

8.1.6 How important is the ED MoA?

A robust test design covering sexual differentiation and reproduction has been identified as being important for identifying ER agonist effects and deriving a sensitive population-relevant NOEC for NP, OP and EE2. The lowest NOECs in the dataset were based on fecundity/fertility for NP, growth for OP and sex ratio or fecundity for EE2. For EE2 and NP these align with the known AOPs (Figure 14, Appendix K). Effects of OP on sex ratio and reproduction were reported in FSDT and FLCTT studies, respectively. For the validation substances the most sensitive endpoint depended on test design. There were only FSDTs available for methyl paraben and propyl paraben, where the NOECs were based on survival/growth or sex ratio, respectively. In the FSTRA with both these substances the NOEC was based on a reduction in the number of fertilised eggs.

Fenbuconazole, cyproconazole, propiconazole and prochloraz all have some evidence for aromatase inhibition *in vitro* and (except for fenbuconazole where no data were available) inhibition of VTG in female fish *in vivo*. The lowest NOEC for all these substances is reduced reproduction in an OECD CF Level 5 test. Of these substances there are only comparative data for effects in an FSDT for prochloraz, and here the effect on sex ratio in the FSDT was relatively insensitive compared with reduced reproduction in the FLCTT. For the anti-androgen vinclozolin the lowest NOEC was based on larval survival in a MEOGRT, with larval survival and fecundity being affected at the lowest test concentration in the second FLCTT.

These data (and knowledge from AOPs) suggest that there will be differences in endpoint sensitivity depending on the endocrine MoA as well as potency. However, it can be difficult to predict with certainty what the most sensitive endpoint will be for any particular substance because it may vary based on test design (e.g. part of the lifecycle assessed) or species (e.g. sex ratio was less sensitive to ER agonist activity in an FSDT with *G. aculeatus* compared with *O. latipes* and *D. rerio,* which is recognised in the TG).

8.2 Final reflections

8.2.1 What dataset is sufficient for PNEC derivation for EDCs?

There does not appear to be any reason why threshold values for EAS modalities cannot be derived in principle (Matthiessen *et al*., 2017). UBA (2014) concluded that NMDR is of low relevance for the environment. The findings for the substances considered in this report suggest that NMDR can in any case be addressed within the existing risk assessment framework (e.g. increased growth in fish as a possible secondary effect of the ER agonist MoA). Concerns regarding threshold levels for EDCs are driven by the high potency of some receptor mediated substances, like EE2. The present assessment suggests the variable quality of the EE2 dataset may have contributed to uncertainties around threshold levels, because none of the studies were conducted for regulatory purposes. Nevertheless, there is no strong evidence for any effect of EE2 on mechanistic or population-relevant endpoints in fish below 0.01 ng/L, and this can therefore be considered a benchmark for other EDCs with EAS modalities.

This is a notably low concentration, and in practical risk assessment terms is almost equivalent to saying EE2 has no safe level of exposure. However, potency is an important factor in receptor mediated activity as demonstrated by *in vitro* data, and this is reflected in the different PNECs derived for the three case study substances. Assuming that all EDCs are equally toxic for risk management purposes is not scientifically defensible.

The proposal for a sufficient dataset for an EDC acting via an EAS MoA is an OECD CF Level 5 fish study. Testing should cover the sensitive window for sexual differentiation as well as reproduction, and include mechanistic endpoints relevant for the MoA of concern to provide context and confidence about the outcome. Ideally the study would follow regulatory accepted test designs (i.e. OECD TG 240 or OPPTS 850.1500) and be performed according to the principles of Good Laboratory Practice. A recognised test design would address the issue raised by the non-standard study for EE2 discussed in Section 6.1. Based on this review, a regulatory standard Level 5 fish study is often the most sensitive result in the dataset for an EDC. Even where the Level 5 NOEC is not the most sensitive result, the availability of such a study should provide confidence that many of the concerns associated with ED properties have been addressed, in line with the SETAC Pellston Workshop® criteria. For example, relevant taxa and endpoints are addressed, and sensitive life-stages and multigenerational effects are considered.

An AF of 10 applied to the result from an OECD CF Level 5 study may be assumed to be protective for effects on other gill-breathing vertebrates in freshwater aquatic environments, by analogy with the approach taken for chemicals in general. Mechanistic effects often occur at lower test concentrations compared with population-relevant endpoints. Therefore, if the derived PNEC is lower than the NOEC for mechanistic effects there can be increased confidence that the value will be protective for ED properties. It is nevertheless acknowledged that further evidence of responses in species with alternative reproductive strategies (such as live bearers) or longer life spans than the species used in regulatory testing would provide further reassurance on this point.

The unknown relative sensitivity of marine fish species is an additional uncertainty that may need to be considered for marine environments. As discussed in Section 110, the available evidence does not suggest that marine species will be markedly more sensitive compared with freshwater species, but there is only relevant comparative data based on similar study designs for one substance (propiconazole, where the difference between the two NOECs is < 2). The approach used for standard risk assessment (ECHA, 2008) could be followed, whereby an additional AF of 10 is applied to the PNEC for freshwater to account for effects in marine species in the absence of sufficient marine organism toxicity data. However, the additional AF for marine systems is intended to cover the increased diversity of marine taxa in general, particularly invertebrates which have low relevance for endocrine effects based on EAS modalities.

It is suggested that the additional AFs proposed by James *et al*. (2023) are not required, provided a fully reliable regulatory standard Level 5 test is available. PNECs derived following the decision tree in James *et al*. (2023) were the most precautionary, with the exception of the Level 3 PNEC, for substances where Level 5 fish data are available (Figure 7 and Figure 8). However, the additional AFs are applied based on whether endocrine mediated endpoints have been addressed or not, essentially giving data from an FSDT or MEOGRT equivalent weight. This may not be sufficiently protective in cases where only the FSDT is available (see methyl paraben and propyl paraben in Figure 7). Especially for the ER agonist MoA (where fish are usually the most sensitive species), applying a default AF of 10 to a dataset without OECD CF Level 4 or 5 fish data, wherein ED endpoints are not explored, would not provide the necessary reassurance that the PNEC is protective for ED effects. The approach proposed in this report is therefore considered to be more refined.

It is acknowledged that existing test guidelines do not address potentially subtle effects of EDCs, such as changes in behaviour or an increased susceptibility to diseases as discussed in UBA (2014). However, chemicals with non-ED MoAs might also cause such effects, so this issue needs to be considered at a more general level.

8.2.2 Limited datasets

An FSDT covers sensitive life stages and incorporates mechanistic endpoints, but does not address concerns for multigeneration effects or effects on reproduction. Comparing the sensitivity of the FSDT with Level 5 studies is difficult because of the limited number of tests conducted, but an additional AF of 10 (total AF of 100) is recommended to be applied to the NOEC from a Level 4 fish study to account for the additional uncertainty if a Level 5 fish test is not available. A FELS test could be used for this purpose, because the findings of this report suggest that the FSDT is not necessarily more sensitive when population relevant endpoints are considered. A more detailed review of the sensitivity of the FELS study versus the FSDT with a wider range of substances may provide more clarity on the use of these studies in deriving a PNEC for an EDC. That said, it is highly unusual for an SVHC identified as an environmental EDC (based on aquatic ecotoxicity data) to have neither an FSDT nor Level 5 study.

An AF of 1 000 is proposed to be applied to the Level 3 FSTRA. This screening assay is usually conducted with wide spacing of test concentrations and does not cover sensitive embryo life stages. The AF is therefore deliberately precautionary. As discussed in Section 105, there is a need for this option to be retained in circumstances where only Level 3 FSTRA data are available.

Derivation of a PNEC for an EDC using relatively small datasets is inherently uncertain. The outcomes should therefore be subject to a degree of expert judgement to ensure that they appear sensible and sufficiently protective. This is likely to be particularly important when there are differences regarding the interpretation of the results of key studies. This is demonstrated with the MEOGRT study outcomes for prochloraz, where refinement of the PNEC depends on which NOEC is selected for use in the assessment. Therefore, while default additional AFs are not considered necessary as standard as part of the proposed approach, additional uncertainty could be accounted for by using a small additional AF of 5 on a case-by-case basis. Additional AFs could be applied at any tier but would need to be fully justified. Justifications would include evidence for a NMDR or lack of adequate protection for mechanistic effects.

9. Conclusions and recommendations

9.1 Conclusions

PNECs were derived for NP, OP and EE2 using datasets which consisted of OECD CF Level 4 and 5 vertebrate data as shown in Table 35. Invertebrates were initially included in the datasets for NP and OP to maximise the species diversity but the resulting probabilistic HC_5 value was not protective for the ED properties. Since invertebrates are not known to have the relevant receptors for EAS mediated effects, they were subsequently excluded from further consideration.

Probabilistic assessments were conducted for OP and EE2, with both datasets being based on NOEC values from six aquatic vertebrate species, which is the minimum required for a probabilistic assessment using ssdtools. Ssdtools was chosen for the probabilistic assessment because it allows fitting of multiple models to the dataset. This was considered useful for demonstrating that a change in the model fit has the potential to impact the HC_5 value for small datasets and hence the outcome of the assessment. For both OP and EE2 the average model fit was used for the probabilistic assessment. There were insufficient data available for NP to allow a probabilistic PNEC to be derived.

Table 35 PNECs for NP, OP and EE2

Substance	Type of PNEC	PNEC value [a]	Biomarker NOEC	Unit	AF	Vertebrate dataset
NP	Deterministic	**0.064**	0.4	µg/L	10	1 MEOGRT 2 FELS
OP	Deterministic	0.61	1	µg/L	10	7 FSDT 1 FELS 3 LAGDA 1 MEOGRT 1 FLCTT
	Probabilistic	**0.77**		µg/L	5	
EE2	Deterministic	**0.005**	0.025	ng/L	10 [b]	2 Level 4 with fish [c] 1 Level 4 with amphibians 5 FLCTT [c]
	Probabilistic	0.006		ng/L	5	

[a] Preferred PNECs are in bold.
[b] A safety margin is already provided by the NOEC from the key study, since the exposure concentrations are widely spaced (other supporting data for the same species suggest that the NOEC could be higher).

[s] Included insensitive NOECs from studies without embryo exposure.

The vertebrate datasets were assessed for suitability for PNEC derivation against the SETAC Pellston Workshop® criteria. These require the dataset to cover sensitive life stages and endpoints, multiple generations, and latent effects. It is also important to consider evidence for NMDR and for whether an ETNC can be determined (i.e. whether effects might be expected at test concentrations lower than those that have been tested). All the datasets were considered acceptable as all included multigeneration studies with fish. There were no concerns for NMDR which would not be addressed as part of the risk assessment (some evidence for increased growth of fish potentially related to the ER agonism MoA). For EE2 there were some initial concerns regarding the ETNC because despite the extensive number of available studies, many were unbounded and there were wide confidence limits around the HC_5 value from the SSD. Furthermore, consideration of the reliable data for EE2 found that studies for two of the species did not include exposure of embryos and were relatively insensitive. The driving NOEC value, from a study with *D. rerio*, was very sensitive but this was because there was a factor of 10 spacing between treatments, compared with the factor of 2 applied to the NOEC from an unbounded study with a similar percentage level of effect. The sensitive NOEC at the low end of the SSD led to uncertainty in the HC_5 value. When the nine reliable NOECs were considered (including multiple Level 5 studies with some species) there was no evidence for either biomarker or population-relevant effects of EE2 below 0.01 ng/L (which is similar to the EU EQS value). The preferred PNEC for EE2 is the deterministic PNEC of 0.005 ng/L because, when two species were excluded from the dataset owing to irrelevant study design (i.e. equivalent to Level 3 tests), there were insufficient species for an SSD.

The PNEC values for EE2, NP and OP were consistent with the relative potency of the substances for ER agonist activity based on the ToxCast bioactivity model and VTG induction in the available FSTRAs. Threshold values were therefore considered possible based on individual substance data (PNECs) and for the ER agonist MoA.

The PNECs in Table 35 are based on applying a standard default AF of 10 for deterministic assessment and 5 for a probabilistic assessment. Given the perceived hazard caused by EDCs, some researchers have suggested that additional AFs are needed, but this assessment has concluded this is generally over-precautionary because all PNECs were considered protective for the NOECs for mechanistic “biomarker” endpoints (when assessed in the context of study design). There may, however, be a case for using an additional AF of 5, for example where the dataset does not meet the SETAC Pellston Workshop® criteria (e.g. evidence for NMDR) or where the PNEC is not protective for relevant mechanistic effects.

The datasets for NP, OP and EE2 were used to propose AFs for use with chemicals with a limited dataset of studies to account for the ED properties: an AF of 10 when the dataset includes an OECD CF Level 5 study, 100 for a CF level 4 study and 1 000 for an OECD CF Level 3 FSTRA. These proposals were reviewed using eight ‘validation’ substances with EAS activity and different amounts of reliable fish data, with PNECs given in Table 36.

Table 36 PNECs for validation substances

Substance	Maximum ED dataset	Assessment factor	PNEC (µg/L)
Methyl paraben	FSDT	100	0.24
Propyl paraben	FSDT	100	1.65
3-Benzylidene camphor	FSTRA	1 000	0.033
Fenbuconazole	Level 5 FLCTT	10	2.3
Cyproconazole	Level 5 FLCTT	10 (algae)	2.1
Propiconazole	Two Level 5 FLCTT	10	6.8
Prochloraz	Level 5 FLCTT and MEOGRT	10 (algae)	0.32
Vinclozolin	Level 5 FLCTT and MEOGRT	10	1.73

Note: For cyproconazole and prochloraz, fish were not the most sensitive species in the dataset and so the AF is applied to the result from the lowest NOECs, which were for algae.

Except for vinclozolin (where the NOECs based on mechanistic data were unbounded) all PNECs were considered protective for mechanistic effects.

The PNECs could not always be explained by potency of the ED mechanism. For example, methyl paraben should be less potent compared with propyl paraben based on the ToxCast bioactivity model and *in vivo* effects in the FSTRA. The most sensitive population-relevant NOEC for methyl paraben in the FSDT was for survival/growth, in contrast to sex ratio for propyl paraben. The differences could potentially be explained by an alternative MoA, but this is unconfirmed. This suggests caution is necessary when using relative potency to read across ED property data between substances.

For substances where there was more than one regulatory style Level 5 study available, a factor of 10 was protective for species differences. It was not possible to compare the sensitivity of the FSDT versus a Level 5 study but in general FSDT and FELS studies showed similar sensitivity. Since these Level 4 studies do not cover effects on reproduction or multigeneration effects, an additional AF of 10 seems justified.

Overall, the proposed AFs appear suitable for addressing the ED properties in fish and the approach gives more weight to the availability of relevant ED data compared with alternative approaches which propose an additional default AF to account for the ED properties.

9.2 Recommendations

An OECD CF Level 5 fish study covering the sensitive window for sexual differentiation and reproductive effects is considered ‘sufficient’ with an AF of 10 for deriving a reliable PNEC to address ED properties via EAS modalities. It is expected that such studies would include relevant mechanistic data. Such tests would be expected to follow the principles of OECD TG 240 or OPPTS 850.1500 (and ideally Good Laboratory Practice). Non-standard Level 5 studies that were not performed according to these principles require additional scrutiny to ensure acceptable lifecycle coverage.

A higher AF of 100 is proposed when only data from OECD CF Level 4 are available (i.e. the FSDT) to account for sensitivity differences relating to test design. A FELS test does not provide mechanistic information and therefore is not sufficient for assessing ED properties, but the available data suggest that an AF of 100 applied to the NOEC for population-relevant parameters from this study may be adequately protective. Whilst this is unlikely to be relevant for environmental EDCs identified as SVHCs (which generally have data from FSDT or Level 5 studies), it might be useful if a group assessment is being performed.

Very few SVHCs only have Level 3 FSTRA data available. A precautionary AF of 1 000 can be applied to the NOEC from such studies to derive a PNEC for screening risk assessment purposes, although the result should be carefully considered alongside other relevant data (*in vitro* data or data for similar substances) for plausibility.

The recommended AFs are summarised in Table 37.

Table 37 AFs to be applied to fish toxicity data for deriving a deterministic PNEC to address ED properties via EAS modalities

Available data	Assessment factor [a]
Lowest NOEC/EC_{10} from a Level 3 FSTRA	1 000
Lowest NOEC/EC_{10} from a Level 4 fish study (FSDT preferred [b])	100
Lowest NOEC/EC_{10} from a Level 5 fish study covering the sensitive window for sexual differentiation and reproduction	10

[a] A small additional AF of 5 could be applied at any tier on a case-by-case basis but would need to be justified. Justifications would include evidence for a NMDR or lack of adequate protection for mechanistic effects.
[b] The available data suggest that FELS tests and FSDTs may have similar sensitivity for population-relevant endpoints.

This proposal is justified based on applying the SETAC Pellston Workshop® criteria and on demonstrating that an ETNC can be established based on receptor-mediated endocrine activity (e.g. ER agonism). Where additional uncertainties are identified within the reliable and relevant dataset, an additional AF of up to 5 could be applied (e.g. to account for NMDR or where a PNEC is not protective for mechanistic effects). Additional AFs should be fully justified.

A future applicant for authorisation who wishes to make an argument about adequate control would be expected to describe how the SETAC Pellston Workshop® criteria apply to the data for their substance. If they cannot show that the criteria are met, they would have to use the socioeconomic route described in Section 11. Applicants would need clear regulatory guidance to follow, which could be based on this report. Such guidance would also allow the regulator to judge when the database for a chemical is sufficient for quantitative risk assessment purposes, which is relevant to restrictions of EDCs under REACH, and to guide Substance Evaluation decisions.

The approach does not account for mixtures of substances acting via the same ED mechanism (for a discussion of this issue, see EA, 2022).

Refinement of a PNEC using probabilistic methods is plausible following the approaches taken in this report, including assessing the quality of the dataset against the SETAC Pellston Workshop® criteria. However, a dataset of six reliable and relevant Level 4 and 5 fish studies will rarely be available, and it is not ethically justified to generate new data for this purpose. There are general ambitions to phase out or reduce animal testing in the long-term. Until there is a better understanding of how threshold levels for an EDC can be derived through alternative methods, it is difficult to justify extrapolation from shorter, less animal intensive studies or even from non-animal data. Nevertheless, this project suggests it is possible to derive precautionary threshold values for an ED using data from FSTRA or FELS studies if the limitations of such an approach are acknowledged.

When using a PNEC derived from ED properties in fish for risk assessment, data for other taxonomic groups should also be checked in case fish are not the most sensitive taxa. For example, algae were the most sensitive taxa for the pesticides cyproconazole and prochloraz, so an ED PNEC would not be sufficiently protective. In this regard, it is worth noting that some EDCs have been developed to target specific taxonomic groups (e.g. micro-organisms in the case of azole fungicides).

9.3 Suggestions for further work

There have been many advances in fish testing for EDCs over the last twenty years and confidence in the outcomes of regulatory assessments of EDCs has therefore increased.

However, from this analysis there remain uncertainties and the following further work could be conducted to address them:

1. There remain data gaps for EAS modalities (based on the present analysis) for comparative Level 5 studies measuring relevant effects on marine species, long lived species, species with alternative reproductive strategies and amphibians. There may be existing data in the scientific literature to fill these knowledge gaps especially for the ER agonist MoA (e.g. for oestradiol, oestrone or 4-tert-pentylphenol) to improve confidence that the proposed AF will cover these groups. Comparative data should ideally cover the sensitive windows for sexual differentiation and reproductive effects.
2. The application of the approach to a larger number of validation substances would provide additional reassurance in its acceptability.
3. Although there are uncertainties in the data quality for deriving a PNEC for EE2, it has been included in “real world” exposure scenarios (e.g. Kidd *et al*., 2014) albeit at concentrations above the proposed PNEC. A review of population models associated with this dataset may provide additional confidence for the proposed threshold level.
4. The FSDT is often used to investigate ED effects, but the present review has found that it can be less sensitive compared with Level 5 tests (and even equivalent level 4 FELS studies), that sensitivity may vary depending on the EAS mechanism and that the guidance for when this study should be required instead of a Level 5 test is unspecific. We recommend conducting a review of the FSDT to compare the sensitivity of this study with other fish tests for substances with different EAS MoA. This review could improve confidence in PNECs derived using this study, and provide more guidance on which endocrine mechanisms are reliably detected using the FSDT.
5. Further work could be done to explore what dataset (and associated AFs) are considered sufficient for risk assessment of thyroid active substances. A data gap for assessing effects via the thyroid is the lack of mechanistic endpoints which can currently be measured in fish.
6. The FSTRA was not sensitive in detecting VTG inhibition in a life cycle test for the pesticide cyproconazole. This is an aromatase inhibitor with relatively low potency. This might be related to variability in VTG measurements. A detailed assessment of the study (including assessment of proficiency criteria for VTG measurement according to Burden *et al*., 2023) could be conducted to assess its reliability for PNEC derivation assuming the data are available. To improve general confidence in VTG measurements it is recommended to take into consideration the outcome of an ongoing OECD review.
7. Issues have been identified with interpreting the outcomes from the MEOGRT, for example inconsistent statistically significant effects on growth parameters at different stages of the life cycle leading to uncertainty regarding biological relevance (Flynn *et al*., 2017; Watanabe *et al*., 2017). Efforts to improve the MEOGRT or develop validated alternatives to this study are therefore recommended (there are potential advantages to designing bespoke fish studies to investigate a specific MoA or concern).

8. Further consideration could be given to whether PNECs for air-breathing organisms can be derived using a similar approach. This is relevant for SVHCs identified as environmental EDCs based on mammalian or avian toxicity.

10. References

ABC LABORATORIES. 1984. Dynamic 14-day acute toxicity of octylphenol to rainbow trout (*Salmo gairdneri*). ABC Lane. Columbia, Missouri 65202, United States. Study number 31911.

ABC Laboratories. 2012. 4-tert-Octylphenol: Larval Amphibian Growth and Development Assay using *Xenopus laevis*. ABC Laboratories, Inc. (In-Life Phase), 7200 E. ABC Lane. Columbia, Missouri 65202, United States. Study number 66086. 28 December 2012.

ACKERMANN GE, SCHWAIGER J, NEGELE RD, FENT K. 2002. Effects of long-term nonylphenol exposure on gonadal development and biomarkers of estrogenicity in juvenile rainbow trout *Oncorhynchus mykiss*. Aquatic Toxicology, 60(3-4), 203-21.

ANKLEY, G.T., BENNETT, R.S., ERICKSON, R.J., HOFF, D.J., HORNUNG, M.W., JOHNSON, R.D., MOUNT, D.R., NICHOLS, J.W., RUSSOM, C.L., SCHMIEDER, P.K., SERRRANO, J.A., TIETGE, J.E., VILLENEUVE, D.L. 2010. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. Environmental Toxicology and Chemistry, 29(3), 730-41.

ANKLEY, G.T., LALONE, C.A., GRAY, E.L., VILLENEUVE, D.L., HORNUNG, M.W. 2016. Evaluation of the scientific underpinnings for identifying estrogenic chemicals in nonmammalian taxa using mammalian test systems. Environmental Toxicology and Chemistry, 35, 2806–2816.

ANKLEY, G.T., BLACKWELL, B.R., CAVALLIN, J.E., DOERING, J.A., FEIFAREK, D.J., JENSEN, K.M., KAHL, M.D., LALONE, C.A., POOLE, S.T., RANDOLPH, E.C., SAARI, T.W., VILLENEUVE, D.L. 2020. Adverse Outcome Pathway network-based assessment of the interactive effects of an androgen receptor agonist and an aromatase inhibitor on fish endocrine function, Environmental Toxicology and Chemistry, 39(4), 913-922.

ANKLEY, G.T., SANTANA-RODRIGUEZ, K., JENSEN, K.M., MILLER, D.H., VILLENEUVE, D.L. 2023. AOP Report: Adverse Outcome Pathways for aromatase inhibition or androgen receptor agonism leading to male-biased sex ratio and population decline in fish. Environmental Toxicology and Chemistry, 42(4), 747-756.

ASHFIELD, L.A., POTTINGER, T.G., SUMPTER, J.P. 1998. Exposure of female juvenile rainbow trout to alkylphenolic compounds results in modification to growth and ovosomatic index. Environmental Toxicology and Chemistry, 17(3), 679-686.

BALCH, G.C., MACKENZIE, C.A., METCALFE, C.D. 2004. Alterations to gonadal development and reproductive success in Japanese medaka (*Oryzias latipes*) exposed to 17α-ethinylestradiol. Environmental Toxicology and Chemistry, 23, 782-791.

BALCH, G. AND METCALFE, C. 2006. Developmental effects in Japanese medaka (*Oryzias latipes*) exposed to nonylphenol ethoxylates and their degradation products. Chemosphere, 62, 1214-1223.

BALDWIN, W., GRAHAM, S., SHEA, D., LE BLANC, G. 1997. Metabolic androgenization of female *Daphnia magna* by the xenoestrogen 4-nonylphenol. Environmental Toxicology and Chemistry, 16(9), 1905-1911.

BECHMANN, R. 1999. Effect of the endocrine disrupter nonylphenol on the marine copepod *Tisbe battagliai*. Science of the Total Environment, 23, 33-46.

BEYDOUN, S., CHOI, H.S., DELA-CRUZ, G., KRUEMPEL, J., HUANG, S., BAZOPOULOU, D., MILLER, H.A., SCHALLER, M.L., EVANS, C.R., LEISER, S.F. 2021. An alternative food source for metabolism and longevity studies in *Caenorhabditis elegans*. Communications in Biology, 4, 258.

BOSKER, T., SANTORO, G., MELVIN, S.D. 2017. Salinity and sensitivity to endocrine disrupting chemicals: A comparison of reproductive endpoints in small-bodied fish exposed under different salinities. Chemosphere, 183, 186-196.

BRENNAN, S., BROUGHAM, C., ROCHE, J., FOGARTY, A. 2006. Multi-generational effects of four selected environmental oestrogens on *Daphnia magna*. Chemosphere, 64, 49-55.

BROOKE, L.T. 1993a. Acute and chronic toxicity of nonylphenol to ten species of aquatic organisms. Report no. 68-C1-0034. Environmental Health Laboratory, Lake Superior Research Institute, University of Wisconsin-Superior.

BROOKE, L.T. 1993b. Accumulation and lethality for two freshwater fishes (fathead minnow and bluegill) to nonylphenol. Report no.: 68-C1-0034. Environmental Health Laboratory, Lake Superior Research Institute, University of Wisconsin-Superior.

BROWN, K.H., SCHULTZ, I.R., NAGLER, J.J. 2007. Reduced embryonic survival in rainbow trout resulting from paternal exposure to the environmental estrogen 17α-ethinylestradiol during late sexual maturation. Reproduction, 134, 659-666.

BROWN, R.J., PANTER, G.H., BURDEN, N., SALINAS, E.R., WELTJE, L., WHEELER, J.R., LAGADIC, L. 2023. Are changes in vitellogenin concentrations in fish reliable indicators of chemical-induced endocrine activity? Ecotoxicology and Environmental Safety 266, 115563.

BROWNE, P., JUDSON, R. S., CASEY, W. M., KLEINSTREUER, N. C., THOMAS, R. S. 2015. Screening Chemicals for Estrogen Receptor Bioactivity Using a Computational Model, Environmental Science and Technology 49(14), 8804–14.

BURNHAM, K. P., AND ANDERSON, D.R. eds. 2002. Model Selection and Multimodel Inference. New York, NY: Springer New York. https://doi.org/10.1007/b97636.

CRANE, M., DUNGEY, S., LILLICRAP, A., THOMPSON, H., WELTJE, L., WHEELER, J.R., LAGADIC, L. 2022. Commentary: Assessing the endocrine disrupting effects of chemicals on invertebrates in the European Union. Environmental Sciences Europe, 34, 36.

DEPIEREUX, S., LIAGRE, M., DANIS, L., DE MEULDER, B., DEPIEREUX, E., SEGNER, H., KESTEMONT, P. 2014. Intersex occurrence in rainbow trout (*Oncorhynchus mykiss*) male fry chronically exposed to ethynylestradiol, PLoS ONE, 9.

DIAMANTI-KANDARAKIS, E., BOURGUIGNON, J.P., GIUDICE, L.C., HAUSER, R., PRINS, G.S., SOTO, A.M., ZOELLER, R.T., GORE, A.C., 2009. Endocrine-Disrupting Chemicals: An Endocrine Society Scientific Statement. Endocrine Reviews, 30, 293-342.

DUCROT, V., TEIXEIRA-ALVES, M., LOPES, C., DELIGNETTE-MULLER, M-L., CHARLES, S., LAGADIC, L. 2010. Development of partial life cycle experiments to assess the effects of endocrine disruptors on the freshwater gastropod *Lymnaea stagnalis*: a case-study with vinclozolin. Ecotoxicology, 19, 1312–1321.

CAMPOS, B., JORDAO, R., RIVETTI, C., LEMOS, M.F.L., SOARES, A.M.V.M., TAULER, R., BARATA, C. 2016. Two generational effects of contaminants in *Daphnia magna*: Effects of offspring quality. Environmental Toxicology and Chemistry, 3, 1470-1477.

COMBER, M.H.I., WILLIAMS, T.D. AND STEWART, K.M. 1993. The effects of nonylphenol on *Daphnia magna*. Water Research, 27 (2), 273-276.

CRUMP, D., LEAN, D., TRUDEAU, V. L. 2002. Octylphenol and UV-B radiation alter larval development and hypothalamic gene expression in the leopard frog (*Rana pipiens*). Environmental Health Perspectives, 110(3), 277-284.

DALGARNO, S. 2018. ssdtools: A shiny web app to analyse species sensitivity distributions. Prepared by Poisson Consulting for the Ministry of the Environment, British Columbia. https://bcgov-env.shinyapps.io/ssdtools/

DRÈZE, V., MONOD, G., CRAVEDI, J-P., BIAGIANTI-RISBOURG, S., LE GAC, F. 2000. Effects of 4-nonylphenol on sex differentiation and puberty in mosquitofish (*Gambusia holbrooki*), *Ecotoxicology* 9, 93-100.

ENVIRONMENT AGENCY (EA). 2002. The Identification of Oestrogenic Effects in Wild Fish – Phase II, R&D Technical Report W2-014/TR. National Centre for Ecotoxicology and Hazardous Substances. Environment Agency, Bristol.

ENVIRONMENT AGENCY (EA). 2004. Causes and Consequences of Feminisation of Male Fish in English Rivers, Science Report SC030285/SR. Environment Agency, Bristol.

ENVIRONMENT AGENCY (EA). 2005. Environmental Risk Evaluation Report: 4-tert-octylphenol. Environment Agency, Bristol. https://assets.publishing.service.gov.uk/media/5a7c78c6e5274a559005a192/scho0405biyz-e-e.pdf (accessed 07 March 2023).

ENVIRONMENT AGENCY (EA), 2010. Exposure Assessment of Landfill Sites Volume 1: Main report, Science Report P1-396. Environment Agency, Bristol.

ENVIRONMENT AGENCY (EA). 2020a. Advice on fish and amphibian vitellogenin assays used for the regulation of endocrine disrupting substances. Project number: C7172G. April 2020. Environment Agency, Bristol.

ENVIRONMENT AGENCY (EA). 2020b. Overview of Endocrine Disrupting Chemicals (EDCs) in the UK: Parabens. Chief Scientist's Group internal discussion report. Unpublished report dated July 2020.

ENVIRONMENT AGENCY (EA). 2021. Overview of Endocrine Disrupting Chemicals (EDCs) in the UK: Camphor/ bornane derivatives. Chief Scientist's Group internal discussion report. Unpublished report dated May 2021.

ENVIRONMENT AGENCY (EA). 2022. Evaluation of the potential approaches to risk assessment of unintentional chemical mixtures for future UK REACH assessments. Environment Agency, Bristol.

ENVIRONMENT AGENCY (EA). 2023. Overview of Endocrine Disrupting Chemicals (EDCs) in the UK: Azoles Chief Scientist's Group internal discussion report. Unpublished report, dated August 2023.

EUROPEAN COMMISSION (EC). 2002. Study on the Scientific Evaluation of 12 Substances in the Context of Endocrine Disruption Priority List of Actions. WRc-NSF Ref: 6052. November 2002.

EUROPEAN COMMISSION (EC). 2016. Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee in accordance with Article 138(7) of REACH to review if the scope of Article 60(3) should be extended to substances identified under Article 57(f) as having endocrine disrupting properties with an equivalent level of concern to other substances listed as substances of very high concern. COM/2016/0814 final.

EUROPEAN COMMISSION (EC). 2018. Common Implementation Strategy for the Water Framework Directive (2000/60/EC) Guidance Document No. 27 Technical Guidance for Deriving Environmental Quality Standards. Technical Report for Approval by Water Directors -WD 2018-1-1. Office for Official Publications in the European Communities, Luxembourg.

EUROPEAN COMMISSION (EC). 2022. Nonylphenol_Final Dossier after SCHEER final opinion. November 2022.

EUROPEAN COMMISSION (EC). 2023. Commission Delegated Regulation (EU) 2023/707 of 19 December 2022 amending Regulation (EC) No 1272/2008 as regards hazard classes and criteria for the classification, labelling and packaging of substances and mixtures. Official Journal of the European Union L93/7. 31 March 2023.

ECOTOXICOLOGY AND TOXICOLOGY OF CHEMICALS (ECETOC). 2014. Estimating toxicity thresholds for aquatic ecological communities from sensitivity distributions. ECETOC workshop report 28. 11-13 February 2014, Amsterdam.

EUROPEAN CHEMICALS AGENCY (ECHA). 2008. Guidance on information requirements and chemical safety assessment. Chapter R.10: Characterisation of dose [concentration]-response for environment. Guidance for the implementation of REACH May 2008.

EUROPEAN CHEMICALS AGENCY (ECHA). 2011. Member State Committee, Support document for identification of 4-(1,1,3,3-tetramethylbutyl)phenol, 4-tert-octylphenol as a substance of very high concern because its endocrine disrupting properties cause probable serious effects to the environment which give rise to an equivalent level of concern [online]. Available from: https://echa.europa.eu/documents/10162/4c6cccfd-d366-4a00-87e5-65aa77181fb6 [Accessed 4 November 2022].

EUROPEAN CHEMICALS AGENCY (ECHA). 2012a. Member State Committee, Support document for identification of 4-nonylphenol, branched and linear as substances of very high concern because due to their endocrine disrupting properties they cause probable serious effects to the environment which give rise to an equivalent level of concern to those of CMRs and PBTs/vPvBs [online]. Available from: https://echa.europa.eu/candidate-list-table/-/dislist/details/0b0236e1807db370 [Accessed 4 November 2022].

EUROPEAN CHEMICALS AGENCY (ECHA). 2012b. Annex XV dossier. Proposal for identification of a substance as a CMR 1a or 1b, PBT, vPvB or a substance of an equivalent level of concern. Substance Name(s): 4-(1,1,3,3-tetramethylbutyl)phenol, ethoxylated - covering well-defined substances and UVCB substances, polymers and homologues. [Accessed 7 March 2023].

EUROPEAN CHEMICALS AGENCY (ECHA). 2013a. Member State Committee, Support document for identification of 4-nonylphenol, branched and linear, ethoxylated as substances of very high concern because, due to their degradation to substances of very high concern (4-nonylphenol, branched and linear) with endocrine disrupting properties, they cause probable serious effects to the environment which give rise to an equivalent level of concern to those of CMRs and PBTs/vPvBs [online]. Available from: https://echa.europa.eu/documents/10162/9af34d5f-cd2f-4e63-859c-529bb39da7ae [Accessed 4 November 2022].

EUROPEAN CHEMICALS AGENCY (ECHA). 2013b. Committee for Risk Assessment (RAC) opinion proposing harmonised classification and labelling at EU level of etridiazole [online]. CLH-O-0000002504-80-02/F Adopted 4 June 2013. Available from: https://echa.europa.eu/documents/10162/1800a914-0898-fd24-b689-55a4f0053321 [Accessed 9 November 2022].

EUROPEAN CHEMICALS AGENCY (ECHA). 2014a. Committee for Risk Assessment (RAC) and Committee for Socio-economic Analysis (SEAC) Background document to the Opinion on the Annex XV dossier proposing restrictions on NONYLPHENOL and NONYLPHENOL ETHOXYLATES [online]. ECHA/RAC/ RES-O-0000005317-74-01/F and ECHA/SEAC/ RES-O-0000005317-74-02/F. Available from:

https://echa.europa.eu/documents/10162/e6556367-e450-54a9-01e8-fbc2851e171f [Accessed 4 November 2022].

EUROPEAN CHEMICALS AGENCY (ECHA). 2014b. Opinion of the Committee for Risk Assessment (RAC) and Committee for Socio-economic Analysis (SEAC) on an Annex XV dossier proposing restrictions on NONYLPHENOL and NONYLPHENOL ETHOXYLATES [online]. ECHA/RAC/ RES-O-0000005317-74-01/F and ECHA/SEAC/ RES-O-0000005317-74-02/F. Available from: https://echa.europa.eu/documents/10162/e6556367-e450-54a9-01e8-fbc2851e171f.

EUROPEAN CHEMICALS AGENCY (ECHA). 2014c. Support document for identification of bis(2-ethylhexyl) phthalate (DEHP) as a substance of very high concern [online]. https://echa.europa.eu/documents/10162/fa429d23-21e7-4764-b223-6c8c98f8a01c [Accessed 14 February 2020].

EUROPEAN CHEMICALS AGENCY (ECHA). 2016. Support Document for Identification of 3-BC as a Substance of Very High Concern. European Chemicals Agency, Helsinki, Finland. https://echa.europa.eu/documents/10162/84e7e600-0f73-6a4a-b246-19482576aa20 [Accessed 9 November 2022].

EUROPEAN CHEMICALS AGENCY (ECHA). 2018. Substance evaluation conclusion as required by REACH Article 48 and evaluation report for Nonylphenol, branched, ethoxylated, EC No 500-209-1. CAS No 68412-54-4. Evaluating Member State(s): United Kingdom February 2018.

EUROPEAN FOOD STANDARDS AGENCY (EFSA)/ EUROPEAN CHEMICALS AGENCY (ECHA). 2018. Guidance for the identification of endocrine disruptors in the context of Regulations (EU) No 528/2012 and (EC) No 1107/2009. EFSA Journal, 16, 5311, 135 pp. https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2018.5311

EUROPEAN FOOD STANDARDS AGENCY (EFSA). 2010a. Conclusion on the peer review of the pesticide risk assessment of the active substance fenbuconazole, EFSA Journal, 8(4), 1558. Accessed 25 February 2022. https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2010.1558

EUROPEAN FOOD STANDARDS AGENCY (EFSA). 2010b. Conclusion on the peer review of the pesticide risk assessment of the active substance cyproconazole. EFSA Journal, 8(11), 1897. Accessed 25 February 2022. https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2010.1897

EUROPEAN FOOD STANDARDS AGENCY (EFSA). 2010c. Conclusion on the peer review of the pesticide risk assessment of the active substance etridiazole. EFSA Journal, 8(10), 1823. Accessed 25 February 2022. https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2010.1823

EUROPEAN FOOD STANDARDS AGENCY (EFSA) 2011 Conclusion on the peer review of the pesticide risk assessment of the active substance prochloraz. EFSA Journal, 9(7),

2323. Accessed 25 February 2022.
https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2011.2323

EUROPEAN FOOD STANDARDS AGENCY (EFSA). 2017. Conclusion on the peer review of the pesticide risk assessment of the active substance propiconazole. EFSA Journal, 15(7), 4887. Accessed 9 November 2022.
https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2017.4887

EUROPEAN MEDICINES AGENCY (EMA). 2018. Guideline on the environmental risk assessment of medicinal products for human use. Committee for Medicinal Products for Human Use (CHMP). EMEA/CHMP/SWP/4447/00 Rev. 1. 15 November 2018.

ENGLAND, D.C. 1995. Chronic toxicity of nonylphenol to *Ceriodaphnia dubia*. Report no. 41756. Report prepared for the Chemical Manufacturers Association by ABC Laboratories Inc.

FLIEDNER, A. 1993. *Daphnia magna*, verlängerter toxizitätstest testsubstanz: Nonylphenol. Fraunhofer-Institut für Umweltchemie und Ökotoxikologie, 1-51.

FLYNN, K., LOTHENBACH, D., WHITEMAN, F., HAMMERMEISTER, D., TOUART, L.W., SWINTEK, J., TATARAZAKO, N., ONISHI, Y., IGUCHI, T., JOHNSON, R. 2017. Summary of the development the US Environmental Protection Agency's Medaka Extended One Generation Reproduction Test (MEOGRT) using data from nine multigenerational medaka tests. Environmental Toxicology and Chemistry, 36 (12), 3387–3403.

FOX, D.R., VAN DAM, R.A., FISHER, R., BATLEY, G.E., TILLMANNS, A.R., THORLEY, J. SCHWARZ C.J., SPRY, D.J., MCTAVISH, K. 2021. Recent developments in species sensitivity distribution modelling. Environmental Toxicology and Chemistry, 40(2), 293–308.

GIBBLE, R. AND BAER K.N. 2003. Effects of 4-nonylphenol on sexual maturation in *Daphnia magna*. Bulletin of Environmental Contamination and Toxicology, 70, 315-321.

GODFRAY, H. C. J., STEPHENS, A. E. A., JEPSON, P. D., JOBLING, S., JOHNSON, A. C., MATTHIESSEN, P., SUMPTER, J. P., TYLER, C. R., MCLEAN, A. R. 2019. A restatement of the natural science evidence base on the effects of endocrine disrupting chemicals on wildlife. Proceedings of The Royal Society, 286, 2018−2416.

GONZALEZ-PEREZ, B.K., SARMA, S.S.S., CASTELLANOS-PAEZ, M.E., NANDINI, S. 2021. Effects of the endocrine disruptor 4-nonylphenol on the demography of rotifers *Plationus patulus* and *Brachionus havanaensis*: a multigenerational study. Journal of Environmental Science and Health, Part A, 56 (13), 1357-1366.

GORE, A.C., CREWS, D., DOAN, L.L., LA MERRILL, M., PATISAUL, H., ZOTA, A. 2014. Introduction to endocrine disrupting chemicals (EDCs) – A guide for public interest organisations and policy-makers. Joint Endocrine Society-IPEN Report, December 2014. https://www.endocrine.org/-/media/endosociety/files/advocacy-and-outreach/important-documents/introduction-to-endocrine-disrupting-chemicals.pdf. Accessed 17 February 2021.

GRANMO, A., EKELUND, R., MAGNUSSON, K., BERGGREN, M. 1989. Lethal and sublethal toxicity of 4-nonylphenol to the common mussel (*Mytilus edulis* L.). Environmental Pollution, 59(2), 115-27.

GRAY, M.A., NIMI, A.J., METCALFE, C.D. 1999a. Factors affecting the development of testis-ova in medaka, *Oryzias latipes*, exposed to octylphenol. Environmental Toxicology and Chemistry, 18 (8), 1835-1842.

GRAY, M.A., TEATHER, K.L., METCALFES, C.D. 1999b. Reproductive success and behavior of Japanese medaka (*Oryzias latipes*) exposed to 4-tert-octylphenol. Environmental Toxicology and Chemistry, 18 (11), 2587-2594.

GRONEN, S., DENSLOW, N.D., MANNING, S., BARNES, S., BARNES, D., BROUWER, M. 1999. Serum vitellogenin levels and reproductive impairment of male Japanese medaka (*Oryzias latipes*) exposed to 4-tert-octylphenol. Environmental Health Perspectives, 107 (5), 385-390.

GIUSTI, A., LAGADIC, L., BARSI, A., THOME, J.P., JOAQUIM-JUSTO, C., DUCROT, V. 2014. Investigating apical adverse effects of four endocrine active substances in the freshwater gastropod *Lymnaea stagnalis*. Science of the Total Environment 493, 147-155.

HARRIES, J.E., RUNNALLS, T., HILL, E., HARRIES, C.A., MADDIX, S., SUMPTER, J.P., TYLER, C.R. 2000. Development of a reproductive performance test for endocrine disrupting chemicals using pair-breeding fathead minnows (*Pimephales promelas*). Environmental Science and Technology, 34 (14), 3003-3011.

HARRIS, C.A., SANTOS, E.M., JANBAKHSH, A., POTTINGER, T.G., TYLER, C.R., SUMPTER, J.P. 2001. Nonylphenol affects gonadotropin levels in the pituitary gland and plasma of female rainbow trout. Environmental Science and Technology, 35 (14), 2909-2916.

HASELMAN, J.T., KOSIAN, P.A., KORTE, J.J., OLMSTEAD, A.W., IGUCHI, T.B., JOHNSON, R.D., DEGITZ, S.J. 2015. Development of the Larval Amphibian Growth and Development Assay: Effects of chronic 4-tert-octylphenol or 17β-trenbolone exposure in *Xenopus laevis* from embryo to juvenile. Journal of Applied Toxicology, 36, 1639–1650.

HOGAN, N. S., LEAN, D. R. S., TRUDEAU, V. L. 2006. exposures to estradiol, ethinylestradiol and octylphenol affect survival and growth of *Rana pipiens* and *Rana sylvatica* tadpoles. Journal of Toxicology and Environmental Health, Part A, 69(16), 1555-1569.

HOLBECH, H., KINNBERG, K.L., BRANDE-LAVRIDSEN, N., BJERREGAARD, P., PETERSEN, G.I., NORRGREN, L., ÖRN, S., BRAUNBECK, T., BAUMANN, L., BOMKE, C., DORGERLOH, M., BRUNS, E., RUEHL-FEHLERT, C., GREEN, J.W, SPRINGER, T.A., GOURMELON, A. 2012. Comparison of zebrafish (*Danio rerio*) and fathead minnow

(*Pimephales promelas*) as test species in the Fish Sexual Development Test (FSDT). Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology, 155(2), 407-415.

HORIE, Y., KANAZAWA, N., TAKAHASHI, C., TATARAZAKO, N., IGUCHI, T. 2021. Exposure to 4-nonylphenol induces a shift in the gene expression of gsdf and testis-ova formation and sex reversal in Japanese medaka (*Oryzias latipes*). Journal of Applied Toxicology, 41, 399-409.

HÖSS, S., JÜTTNER, I., TRAUNSPURGER, W., PFISTER, G., SCHRAMM, K.W. AND STEINBERG, C. 2002. Enhanced growth and reproduction of *Caenorhabditis elegans* (Nematoda) in the presence of 4-nonylphenol. Environmental Pollution, 120, 169-172.

HULS, 1992. Datasheet for 4-(1,1,3,3,tetramethylbutyl)phenol), Long-term toxicity to aquatic invertebrates, 002 Key: Experimental result [online]. ECHA. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/15074/6/2/5/?documentUUID=77a1b7c3-f114-4f2f-b2ba-a1e0640596f6 [Accessed 3 November 2022].

INERIS (French National Institute for Industrial Environment and Risks), 2011. Normes de Qualite Environnementale: Cyproconazole – CAS No. 9436106-5 [online]. Available from: https://substances.ineris.fr/fr/substance/2198/3 [Accessed 9 November 2022].

ISHIBASHI, H., HIRANO, M., MATSUMURA, N., WATANABE, N., TAKAO, Y., ARIZONO, K. 2006. Reproductive effects and bioconcentration of 4-nonylphenol in medaka fish (*Oryzias latipes*). Chemosphere, 65 (6), 1019-1026.

IUCLID, 2000. Datasheet for 4-(1,1,3,3,tetramethylbutyl)phenol), International Uniform Chemical Information Database on High Production Volume Chemicals (HPVCs) reported by European Industry in the frame of the EU Existing Chemicals Risk Assessment Programme. EUR 19559 EN.

IUCLID , 2014. Datasheet for Phenol, 4-nonyl-, branched, Long-term toxicity to fish, 011 Supporting: Experimental results [online]. ECHA. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/15896/6/2/3/?documentUUID=03102f7e-a2ec-4044-a859-e28f4e788e9d [Accessed 3 November 2022].

IUCLID, 2019. Datasheet for Phenol, 4-nonyl-, branched, Long-term toxicity to aquatic invertebrates, 008 Supporting: Experimental results [online]. ECHA. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/15896/6/2/5/?documentUUID=dd46495e-bc5c-4d95-808e-64b27c283262 [Accessed 3 November 2022].

IUCLID , 2020. Datasheet for Phenol, 4-nonyl-, branched, Long-term toxicity to aquatic invertebrates, 001 Key: Experimental results [online]. ECHA. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/15896/6/2/5/?documentUUID=30c48638-600b-4808-b7b8-23d7161cfb84 [Accessed 3 November 2022].

IUCLID, 2021a. Datasheet for Phenol, 4-nonyl-, branched [online]. ECHA. Last modified 22 December 2021. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/15896/ [Accessed 4 November 2022].

IUCLID, 2021b. Datasheet for propyl 4-hydroxybenzoate [online]. ECHA. Last modified 28 October 2021. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/13890/ [Accessed 13 December 2022].

IUCLID, 2021c. Datasheet for 4-(4-chlorophenyl)-2-phenyl-2-[(1H-1,2,4-triazol-1-yl)methyl]butanenitrile [online]. ECHA. Last modified 7 September 2021. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/32833/6/2/6 [Accessed 10 November 2022].

IUCLID, 2022a. Datasheet for 4-(1,1,3,3,tetramethylbutyl)phenol) [online]. ECHA. Last modified 22 September 2022. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/15074/ [Accessed 4 November 2022].

IUCLID, 2022b. Datasheet for methyl 4-hydroxybenzoate [online]. ECHA. Last modified19 December 2022. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/14310/ [Accessed 24th January 2023].

IUCLID, 2022c. Datasheet for Bis(2-ethylhexyl) phthalate [online]. ECHA. Last modified 19 January 2022. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/15358/6/2/6 [Accessed 10 November 2022].

IUCLID, 2024. Datasheet for cyproconazole [online]. ECHA. Last modified 19 May 2023. Available from: https://echa.europa.eu/registration-dossier/-/registered-dossier/32337/6/2/3 [Accessed 06 March 2023].

IVEY, C.D., WANG, N., ALVAREZ, D., HAMMER, E.J., BAUER, C.R. 2018. Chronic toxicity of 4-nonylphenol to two unionid mussels in water-only exposures. Bulletin of Environmental Contamination and Toxicology, 101, 423-427.

JAMES, A., KROLL, A. MINIER, C. T. 2023. Towards a better consideration of endocrine disruption within the technical guidance for deriving environmental quality standards, Regulatory Toxicology and Pharmacology, 143, 105457.

JAPAN. MINISTRY OF THE ENVIRONMENT. 2002. SPEED98_Strategic_Programs_on_Environmental_Endocrine_Disrupters Annex 7.cited in ECHA. 2011. Available from: https://echa.europa.eu/documents/10162/4c6cccfd-d366-4a00-87e5-65aa77181fb6 [Accessed 4 November 2022].

JOBLING, S., SHEAHAN, D., OSBORNE, J.A, MATTHIESSEN, P., SUMPTER, J.P. 1996. Inhibition of testicular growth in rainbow trout (*Oncorhynchus mykiss*) exposed to estrogenic alkylphenolic chemicals. Environmental Toxicology and Chemistry, 15 (2), 194-202.

JOBLING, S., CASEY, D., RODGERS-GRAY, T., OEHLMANN, J., SCHULTE-OEHLMANN, U., PAWLOWSKI, S., BAUNBECK, T., TURNER, A. P., TYLER, C. R. 2003. Comparative responses of molluscs and fish to environmental estrogens and an estrogenic effluent. Aquatic Toxicology, 65(2), 205-220.

JOLLY, C., KATSIADAKI, I., MORRIS, S., LE BELLE, N., DUFOUR, S., MAYER, I., POTTINGER, T.G., SCOTT, A.P. 2009. Detection of the anti-androgenic effect of endocrine disrupting environmental contaminants using *in vivo* and *in vitro* assays in the three-spined stickleback. Aquatic Toxicology, 92 (4), 228-239.

JOINT RESEARCH CENTRE (JRC). 2016. Monitoring-based Exercise: Second Review of the Priority Substances List under the Water Framework Directive. European Commission, Ispra, Italy. Accessed February 2022. https://circabc.europa.eu/sd/a/7fe29322-946a-4ead-b3b9-e3b156d0c318/Monitoring-based%20Exercise%20Report_FINAL%20DRAFT_25nov2016.pdf.

JOINT RESEARCH CENTRE (JRC). 2020. Selection of substances for the 3rd Watch List under the Water Framework Directive. European Commission, Luxembourg. Accessed February 2022. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC121346/third_watch_list_report_pdf.pdf.

JOINT RESEARCH CENTRE (JRC). 2022. EQS datasheet – Environmental Quality Standard: 17 Alpha Ethinylestradiol (EE2) [online]. October 2022.

JUN, F.L.M., STALLING, D.L., JOHNSON, J.L. 1972. Phthalate esters as environmental contaminants. Nature, 238:411-413.

JUNG, J.W., KANG, J.S., CHOI, J., PARK, J.W. 2020. Chronic toxicity of endocrine disrupting chemicals used in plastic products in Korean resident species: Implications for aquatic ecological risk assessment. Ecotoxicology and Environmental Safety, 192, 110309.

KAHL, M., MAKYNEN, E., KOSIAN, P., ANKLEY, G. 1997. Toxicity of 4-nonylphenol in a life cycle test with the midge *Chironomus tentans*. Ecotoxicology and Environmental Safety, 38, 155-160.

KAMEL, A., MATTEN, S.R., LYNN, S.G., WOLF, J.C. AND FORT, D.J. 2022. Amphibian Metamorphosis Assay: Investigation of the potential effects of five chemicals on the hypothalamic-pituitary thyroid axis of *Xenopus laevis*. Regulatory Toxicology and Pharmacology, 134, 105241.

KANG, I.J., YOKOTA, H., OSHIMA, Y., HANO, T., TSURUDA, Y., MAEDA, M., IMADA, N., TADOKORO, H., HONJO, T. 2003. Effects of 4-nonylphenol on reproduction of Japanese medaka, *Oryzias latipes*. Environmental Toxicology and Chemistry, 22 (10), 2438-2445.

KARELS, A.A., MANNING, S., HOEXUM BROUWER, T., BROUWER, M. 2003. Reproductive effects of estrogenic and antiestrogenic chemicals on sheepshead minnows (*Cyprinodon variegatus*). Environmental Toxicology and Chemistry, 22 (4), 855-865.

KATSIADAKI, I., MORRIS, S., SQUIRES, C., HURST, M.R., JAMES, J.D., SCOTT, A.P. 2006. Use of the three-spined stickleback (*Gasterosteus aculeatus*) as a sensitive *in vivo* test for detection of environmental antiandrogens. Environmental Health Perspectives, 114 (Supp 1), 115-121.

KAWASHIMA, Y., ONISHI, Y., TATARAZAKO, N., YAMAMOTO, H., KOSHIO, M., OKA, T., HORIE, Y., WATANABE, H., NAKAMOTO, T., YAMAMOTO, J., ISHIKAWA, H., SATO, T., YAMAZAKI, K., IGUCHI, T. 2021. Summary of 17 chemicals evaluated by OECD TG 229 using Japanese medaka, *Oryzias latipes* in EXTEND 2016. Journal of Applied Toxicology, 2, 750-777.

KIDD, K.A., PATERSON, M.J., RENNIE, M.D., PODEMSKI, C.L., FINDLAY, D.L., BLANCHFIELD, P.J., LIBER, K. 2014. Direct and indirect responses of a freshwater food web to a potent synthetic oestrogen. Philosophical Transactions of the Royal Society B, 369, 20130578

KINNBERG, K., KORSGAARD, B., BJERREGAARD, P. 2003. Effects of octylphenol and 17(beta)-estradiol on the gonads of guppies (*Poecilia reticulata*) exposed as adults via the water or as embryos via the mother. Comparative Biochemistry and Physiology Part C, Toxicology and Pharmacology, 134 (1), 45-55.

KLEINSTREUER, N.C., CEGER, P., WATT, E.D., MARTIN, M., HOUCK, K., BROWNE, P., THOMAS, R.S., CASEY, W.M., DIX, D.J., ALLEN, D., SAKAMURU, S., XIA, M., HUANG, R., JUDSON, R. 2017. Development and validation of a computational model for androgen receptor activity. Chemical Research in Toxicology, 30, 946–964.

KLIMISCH, H.J., ANDREAE, M., TILLMANN, U., 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regulatory Toxicology and Pharmacology, 25(1), 1-5.

KLOAS, W., LUTZ, I., EINSPANIER, R. 1999. Amphibians as a model to study endocrine disruptors: II. Estrogenic activity of environmental chemicals *in vitro* and *in vivo*. The Science of The Total Environment, 225(1-2), 59-68.

KNÖRR, S. AND BRAUNBECK, T. 2002. Decline in reproductive success, sex reversal, and developmental alterations in Japanese medaka (*Oryzias latipes*) after continuous exposure to octylphenol. Ecotoxicology and Environmental Safety, 51 (3), 187-196.

KOPF, W. 1997. Wirkung endokriner stoffe in biotests mit wasserogranismen. In Stoffe mit endokriner wirkung in wasser. Bayerisches landesamt für wasserwirtschaft, Institut für Wasserforschung München.

KRISTENSEN, T., BAATRUP, E., BAYLEY, M. 2005. 17α-ethinylestradiol reduces the competitive reproductive fitness of the male guppy (*Poecilia reticulata*). Biology of Reproduction, 72, 150-156.

KUHN, A., MUNNS JR., W.R., CHAMPLIN, D., MCKINNEY, R., TAGLIABUE, M., SERBST, J., GLEASON, T. 2001. Evaluation of the efficacy of extrapolation population modeling to predict the dynamics of *Americamysis bahia* populations in the laboratory. Environmental Toxicology and Chemistry, 20 (1), 213-221.

KUNZ, P.Y., GRIES, T., FENT, K. 2006. The ultraviolet filter 3-benzylidene camphor adversely affects reproduction in fathead minnow (*Pimephales promelas*). Toxicological Sciences, 93, 311-321.

LÄNGE, R., HUTCHINSON, T.H., CROUDACE, C.P., SIEGMUND, F., SCHWEINFURTH, H., HAMPE, P., PANTER, G.H., SUMPTER, J.P. 2001. Effects of the synthetic estrogen 17α-ethinylestradiol on the life cycle of the fathead minnow (*Pimephales promelas*). Environmental Toxicology and Chemistry, 20 (6), 1216-1217.

LANGE, A., KATSU, Y., ICHIKAWA, R., PAULL, G.C., CHIDGEY, L.L., COE, T.S., IGUCHI, T., TYLER, C.R. 2008. Altered sexual development in roach (*Rutilus rutilus*) exposed to environmental concentrations of the pharmaceutical 17α-ethinylestradiol and associated expression dynamics of aromatases and estrogen receptors. Toxicological Sciences, 106 (1), 113-123.

LARSEN, M.G., HANSEN, K.B., HENRIKSEN, P.G., BAATRUP, E. 2008. Male zebrafish (*Danio rerio*) courtship behaviour resists the feminising effects of 17α-ethinyloestradiol-morphological sexual characteristics do not. Aquatic Toxicology, 87, 234-244.

LI, X., SHEN, Y., LANG, B., ZHAO, J., WANG, H., ZHANG, Y. 2021. Influence of octylphenol on gene expression of gonadotropins and their receptors, testicular structure and mating behavior of male *Rana chensinensis*. Environmental Toxicology and Pharmacology, 87, 103694.

LIN, L.L. AND JANZ, D.M. 2006. Effects of binary mixtures of xenoestrogens on gonadal development and reproduction in zebrafish. Aquatic Toxicology, 80 (4), 382-395.

LUSSIER, S. M., CHAMPLIN, D., LIVOLSI, J., POUCHER, S. PRUELL, R. J. 2000. Acute toxicity of para-nonylphenol to saltwater animals. Environmental Toxicology and Chemistry, 19, 617–621.

MAACK, G., AND SEGNER, H. 2004. Life-stage-dependent sensitivity of zebrafish (*Danio rerio*) to estrogen exposure. Comparative Biochemistry and Physiology - C Toxicology and Pharmacology, 139, 47-55.

MAKYNEN, E.A., KAHL, M.D., JENSEN, K.M., TIETGE, J.E., WELLS, K.L., VAN DER KRAAK, G., ANKLEY, G.T. 2000. Effects of the mammalian antiandrogen vinclozolin on development and reproduction of the fathead minnow (*Pimephales promelas*). Aquatic Toxicology, 48 (4), 461-475.

MARCIAL, H.S., HAGIWARA, A., SNELL, T.W. 2003. Estrogenic compounds affect development of harpacticoid copepod *Tigriopus japonicus*. Environmental Toxicology and Chemistry, 22 (12), 3025-3030.

MARINI, J.P., WOLF, J.C., MINGO, V., SAYERS, L.E., JAMIESON, S.Y., WHEELER, J.R. 2023. An Amphibian Metamorphosis Assay dietary restriction study: Lessons for data interpretation. Environmental Toxicology and Chemistry. 42(5), 1061-1074.

MARTINOVIĆ, D., BLAKE, L.S., DURHAN, E.J., GREENE, K.J., KAHL, M.D., JENSEN, K.M., MAKYNEN, E.A., VILLENEUVE, D.L., ANKLEY, G.T. 2008. Reproductive toxicity of vinclozolin in the fathead minnow: Confirming an anti-androgenic mode of action. Environmental Toxicology and Chemistry, 27 (2), 478-488

MATTHIESSEN, P., ANKLEY, G.T., BIEVER, R.C., BJERREGAARD, P., BORGERT, C., BRUGGER, K., BLANKINSHIP, A., CHAMBERS, J., COADY, K.K., CONSTANTINE, L., DANG, Z., DENSLOW, N.D., DREIER, D.A., DUNGEY, S., GRAY, L.E., GROSS, M., GUINEY, P.D., HECKER, M., HOLBECH, H., IGUCHI, T., KADLEC, S., KAROUNA-RENIER, N.K., KATSIADAKI, I., KAWASHIMA, Y., KLOAS, W., KRUEGER, H., KUMAR, A., LAGADIC, L., LEOPOLD, A., LEVINE, S.L., MAACK, G., MARTY, S., MEADOR, J., MIHAICH, E., ODUM, J., ORTEGO, L., PARROTT, J., PICKFORD, D., ROBERTS, M., SCHAEFERS, C., SCHWARZ, T., SOLOMON, K., VERSLYCKE, T., WELTJE, L., WHEELER, J.R., WILLIAMS, M., WOLF, J.C., YAMAZAKI, K. 2017. Recommended approaches to the scientific evaluation of ecotoxicological hazards and risks of endocrine-active substances. Integrated Environmental Assessment and Management, 13 (2), 267-279.

MEHINTO, A.C., KROLL, K.J., JAYASINGHE, B.S., LAVELLE, C.M., VAN DERVORT, D., ADEYEMO, O.K., BAY, S.M., MARUYA, K.A., DENSLOW, N.D. 2018. Linking *in vitro* estrogenicity to adverse effects in the inland silverside (*Menidia beryllina*). Environmental Toxicology and Chemistry, 37, 884-892.

METCALFE, T.L., METCALFE, C.D., KIPARISSIS, Y., NIIMI, A.J., FORAN, C.M., BENSON, W.H. 2000. Gonadal development and endocrine responses in Japanese medaka (*Oryzias latipes*) exposed to o,p’-DDT in water or through maternal transfer. Environmental Toxicology and Chemistry, 19(7), 1893–1900.

MILES-RICHARDSON, S.R., PIERENS, S.L., NICHOLS, K.M., KRAMER, V.J., SNYDER, E.M., SNYDER, S.A., RENDER, J.A., FITZGERALD, S.D., GIESY, J.P. 1999. Effects of waterborne exposure to 4-nonylphenol and nonylphenol ethoxylate on secondary sex characteristics and gonads of fathead minnows (*Pimephales promelas*). Environmental Research, 80 (2), S122-S137.

MOERMOND, C., KASE, R., KORKARIC, M., ÅGERSTRAND, M. 2016. CRED – Criteria for Reporting and evaluating ecotoxicity Data. Environmental Toxicology and Chemistry, 35, 1297-1309.

MORSHEAD, M.L., JENSEN, K.M., ANKLEY, G.T., VLIET, S., LALONE, C.A., ALLER, A.V., WATANABE, K.H., VILLENEUVE, D.L. 2023. Putative adverse outcome pathway development based on physiological responses of female fathead minnows to model estrogen versus androgen receptor agonists. Aquatic Toxicology, 61, 106607.

NASH, J.P., KIME, D.E., VAN DER VEN, L.T.M., WESTER, P.W., BRION, F., MAACK, G., STAHLSCHMIDT-ALLNER, P., TYLER, C.R. 2004. Long-term exposure to environmental concentrations of the pharmaceutical ethynylestradiol causes reproductive failure in fish. Environmental Health Perspectives, 112 (17), 1725-1733.

NAKAMURA, A., TAKANOBU, H., TAMURA, I., YAMAMURO, M., IGUCHI, T., TATARAZAKO, N. 2014. Verification of responses of Japanese medaka (*Oryzias latipes*) to anti-androgens, vinclozolin and flutamide, in short-term assays. Journal of Applied Toxicology, 34 (5), 545-553.

NIELSEN, L. AND BAATRUP, E. 2006. Quantitative studies on the effects of environmental estrogens on the testis of the guppy, *Poecilia reticulata*. Aquatic Toxicology, 80 (2), 140-148.

NATIONAL INSTITUTE OF ENVIRONMENTAL HEALTH SCIENCES (NIEHS) 2024. Endocrine Disruptors and Your Health. Webpage: https://www.niehs.nih.gov/health/topics/agents/endocrine. Accessed 15th August 2025.

NIMROD, A.C. AND BENSON, W.H. 1998. Reproduction and development of Japanese medaka following an early life stage exposure to xenoestrogens. Aquatic Toxicology, 44 (1-2), 141-156.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). 2000. Test No. 215: Fish, Juvenile Growth Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264070202-en.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). 2006. Report of the validation of the 21-day fish screening assay for the detection of endocrine substances (Phase 1B). Report no. JT03213438.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD).2010. Validation report of the 21-day androgenized female stickleback screening assay [online]. OECD series on testing and assessment. No. 128. Paris. ENV/JM/MONO(2010)19. Accessed 16 December 2022. Available from: https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2010)19&doclanguage=en.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). 2011. Test No. 234: Fish Sexual Development Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264122369-en.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). 2012. Validation report (Phase 2) for the fish sexual development test for the detection of endocrine active substances (Updated 2012). In: Series on Testing and assessment No. 142.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). 2013. Test No. 210: Fish, Early-life Stage Toxicity Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264203785-en.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). 2015., Test No. 240: Medaka Extended One Generation Reproduction Test (MEOGRT), OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264242258-en.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD).2016., Test No. 242: *Potamopyrgus antipodarum* Reproduction Test. OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264264311-en.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). 2018. Revised Guidance Document 150 on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption, OECD Series on Testing and Assessment, No. 150, OECD Publishing, Paris, https://doi.org/10.1787/9789264304741-en.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). 2022. Validation of the Rapid Androgen Disruption Activity Reporter (RADAR) assay for the detection of androgen axis active substances (draft report) [online]. Accessed 16 December 2022. Available from: https://www.oecd.org/chemicalsafety/testing/draft-validation-report-rapid-androgen-disruption-activity-reporter-assay.pdf.

OEHLMANN, J., SCHULTE-OEHLMANN, U., TILLMANN, M., MARKERT, B. 2000. Effects of endocrine disruptors on prosobranch snails (mollusca: gastropoda) in the laboratory. Part I: Bisphenol A and octylphenol as xeno-estrogens. Ecotoxicology, 9(6), 383-397.

OGINAH, S.A., POSTHUMA, L., HAUSCHILD, M., SLOOTWEG, J., KOSNIK, M., FANTKE, P. 2023. To split or not to split: Characterizing chemical pollution impacts in aquatic ecosystems with species sensitivity distributions for specific taxonomic groups. Environmental Science and Technology, 57(39), 14526-14538.

ONISHI, Y., TATARAZAKO, N., KOSHIO, M., OKAMURA, T., WATANABE, H., SAWAI, A., YAMAMOTO, J., ISHIKAWA, H., SATO, T., KAWASHIMA, Y., YAMAZAKI, K., IGUCHI, T. 2021. Summary of reference chemicals evaluated by the fish short-term reproduction assay, OECD TG229, using Japanese medaka, *Oryzias latipes*. Journal of Applied Toxicology, 41(8), 1200-1221.

PANTER, G.H., BROWN, R.J., JONES, A., KÖRNER, O., LAGADIC, L., WELTJE, L. 2022. Detection of anti-androgenic activity in chronic fish studies: Critical Reviews in Toxicology, 53 (5), 326-338.

PARROTT, J.L., AND BLUNT, B.R. 2005. Life cycle exposure of fathead minnows (*Pimephales promelas*) to an ethinylestradiol concentrations below 1 ng/L reduces egg fertilisation success and demasculinizes males. Environmental Toxicology, 20 (2), 131-141.

PETERS, R.E.M., COURTENAY, S.C., HEWITT. L.M., MACLATCHY, D.L. 2010. Effects of 17α-ethynylestradiol on early-life development, sex differentiation and vitellogenin induction in mummichog (*Fundulus heteroclitus*). Marine Environmental Research, 69, 178-186.

PETTERSSON, I., AND BERG, C. 2007. Environmentally relevant concentrations of ethinylestradiol cause females-biased sex ratios in *Xenopus tropicalis* and *Rana temporaria*. Environmental Toxicology and Chemistry, 26 (5), 1005-1009.

PORTER, K.L., OLMSTEAD, A.W., KUMSHER, D.M., DENNIS, W.E., SPRANDO, R.L., HOLCOMBE, G.W., KORTE, J.J., LINDBERG-LIVINGSTON, A., DEGITZ, S.J. 2011. Effects of 4-tert-octylphenol on *Xenopus tropicalis* in a long term exposure. Aquatic Toxicology, 103 (3–4), 159-169

POSTHUMA, L., VAN GILS, J., ZIJP, M.C, VAN DE MEENT, D., DE ZWART, D., 2019. Species sensitivity distributions for use in environmental protection, assessment, and management of aquatic ecosystems for 12 386 chemicals. Environmental Toxicology and Chemistry, 38(4), 905-917.

POTTINGER, T.G., KATSIADAKI, I., JOLLY, C., SANDERS, M., MAYER, I., SCOTT, A.P., MORRIS, S. KORTENKAMP, A., SCHOLZE, M. 2013. Anti-androgens act jointly in suppressing spiggin concentrations in androgen-primed female three-spined sticklebacks - prediction of combined effects by concentration addition. Aquatic Toxicology, 15, 140-141, 145-56

RASMUSSEN, T.H., ANDREASSEN, T.K., PEDERSEN, S.N., VAN DER VEN, L.T.M., BJERREGAARD, P., KORSGAARD, B. 2002. Effects of waterborne exposure of octylphenol and oestrogen on pregnant viviparous eelpout (*Zoarces viviparus*) and her embryos in ovario. Journal of Experimental Biology, 205 (24), 3857-3876.

RASMUSSEN, T.H., TEH, S.J., BJERREGAARD, P., KORSGAARD, B. 2005. Anti-estrogen prevents xenoestrogen-induced testicular pathology of eelpout (*Zoarces viviparus*). Aquatic Toxicology, 72 (3), 177-194.

ROBINSON, C.D., BROWN, E., CRAFT, J.A., DAVIES, I.M., MOFFAT, C.F., PIRIE, D., STAGG, R.M., STRUTHERS, S. 2003. Effects of sewage effluent and ethynyl oestradiol upon molecular markers of oestrogenic exposure, maturation and reproductive success in the sand goby (*Pomatoschistus minutus*, Pallas). Aquatic Toxicology, 62 (2), 119-134.

ROBINSON, C.D., BROWN, E., CRAFT, J.A., DAVIES, I.M., MOFFAT, C.F., 2004. Effects of prolonged exposure to 4-tert-octylphenol on toxicity and indices of oestrogenic exposure in the sand goby (*Pomatoschistus minutus*, Pallas). Marine Environmental Research, 58 (1), 19-38.

ROHR, J.R., ELSKUS, A.A., SHEPHERD, B.S., CROWLEY, P.H., MCCARTHY, T.M., NIEDZWIECKI, J.H., SAGER, T., SIH, A., PALMER, B.D., 2003. Lethal and sublethal effects of atrazine, carbaryl, endosulfan, and octylphenol on the streamside salamander (*Ambystoma barbouri*). Environmental Toxicology and Chemistry, 22 (10), 2385-2392.

ROUTLEDGE, E.J., SHEAHAN, D., DESBROW, C., BRIGHTY, G.C., WALDOCK, M., SUMPTER, J.P., 1998. Identification of estrogenic chemicals in STW effluent. 2. In vivo responses in trout and roach. Environmental Science and Technology, 32 (11), 1559-1565.

ROYAL SOCIETY OF CHEMISTRY, 2025. Chemical mixtures in aquatic environments. https://royalsociety.org/-/media/policy/projects/environmental-chemical-mixtures/chemical-mixtures-report.pdf

RUTHERFORD, R., LISTER, A., BOSKER, T., BLEWETT, T., MEINA, E.G., CHEHADE, I., KANAGASABESAN, T., MACLATCHY, D., 2020. Mummichog (*Fundulus heteroclitus*) are less sensitive to 17α-ethinylestradiol (EE2) than other common model teleosts: A comparative review of reproductive effects. General and Comparative Endocrinology, 289, 113378.

SCHÄFERS C., TEIGELER, M., WENZEL, A., MAACK, G., FENSKE, M., SEGNER, H., 2007. Concentration- and time-dependent effects of the synthetic estrogen, 17α-ethinylestradiol, on reproductive capabilities of the zebrafish, *Danio rerio*. Journal of Toxicology and Environmental Health - Part A: Current Issues, 70, 768-779.

SCHEER [Scientific Committee on Health, Environmental and Emerging Risks]. 2022. Final opinion on Nonylphenol, 11 November 2022. https://health.ec.europa.eu/publications/scheer-scientific-opinion-draft-environmental-quality-standards-priority-substances-under-water-7_en.

SCHOENFUSS, H.L., BARTELL, S.E., BISTODEAU, T.B., CEDIEL, R.A., GROVE, K.J., ZINTEK, L., LEE, K.E., BARBER, L.B. 2008. Impairment of the reproductive potential of male fathead minnows by environmentally relevant exposures to 4-nonylphenol. Aquatic Toxicology, 86, 91-98.

SCHOLZ, N. 1992. Bestimmung der auswirkungen von nonylphenol auf die reproduktion von *Daphnia magna*. Study number: 2468, Report no.: DL-143. Testing laboratory: Huels AG. Owner company: Sasol Germany GmbH.

SCHOLZ, S AND GUTZEIT, H.O. 2000. 17-α-ethinylestradiol affects reproduction, sexual differentiation and aromatase gene expression of the medaka (*Oryzias latipes*). Aquatic Toxicology, 50 (4), 363-373.

SCHUG, T., JOHNSON, A.F., BIRNBAUM, L.S., COLBORN, T., GUILETTE, L.J., CREWS, D.P., COLLINS, T., SOTO, A.N., VOM SAAL, F.S., MCLACHLAN, J.A., SONNENSCHEIN, C., HEINDEL, J.J., 2016. Mini-review: Endocrine Disruptors: Past Lessons and Future Directions. Molecular Endocrinology, 30:833-847.

SCHULTZ, I.R., SKILLMAN, A., NICOLAS, J.-M., CYR, D.G., NAGLER, J.J. 2003. Short term exposure to 17-alpha-ethinylestradiol decreases the fertility of sexually mature male rainbow trout (*Oncorhynchus mykiss*). Environmental Toxicology and Chemistry, 22, 1272-1280.

SCHWAIGER, J., MALLOW, U., FERLING, H., KNOERR, S., BRAUNBECK, T., KALBFUS, W., NEGELE, R.D. 2002. How estrogenic is nonylphenol? A transgenerational study using rainbow trout (*Oncorhynchus mykiss*) as a test organism. Aquatic Toxicology, 59 (3-4), 177-189.

SEGNER, H., NAVAS, J.M., SCHÄFERS, C., WENZEL, A. 2003. Potencies of estrogenic compounds in *in vitro* screening assays and in life cycle tests with zebrafish *in vivo*. Ecotoxicology and Environmental Safety, 54, 315-322.

SEKI, M., YOKOTA, H., MAEDA, M., TADOKORO, H., KOBAYASHI, K., 2003. Effects of 4-nonylphenol and 4-tert-octylphenol on sex differentiation and vitellogenin induction in medaka (*Oryzias latipes*). Environmental Toxicology and Chemistry, 22 (7), 1507-1516.

SHIRDEL, I. AND KALBASSI, M.R. 2016. Effects of nonylphenol on key hormonal balances and histopathology of the endangered Caspian brown trout (*Salmo trutta caspius*), Comparative Biochemistry and Physiology, Part C, 183-184, 28-35.

SIFAKIS, S., ANDROUTSOPOULOS, V.P., TSATSAKIS, A.M., SPANDIDOS, D.A., 2017. Human Exposure to Endocrine Disrupting Chemicals: Effects on the Male and Female Reproductive Systems. Environmental Toxicology and Pharmacology, 51, 56-70.

SOARES, J, COIMBRA, A.M., REIS-HENRIQUES, M.A., MONTEIRO, N.M., VIEIRA, M.N., OLIVERIA, J.M.A., GUEDES-DIAS, P., FONTAINHAS-FERNANDES, A., PARRA, S.S., CARVALHO, A.P., CASTRO, L.F.C., SANTOS, M.M. 2009. Disruption of zebrafish (*Danio rerio*) embryonic development after full life cycle parental exposure to low levels of ethinylestradiol. Aquatic Toxicology, 95, 330-338.

SPADOTO, M., SUEITT, A.P.E., GALINARO, C.A., DA SILVA PINTO, T., POMPEI, C.M.E., BOTTA, C.M.R., VIEIRA, E.M. 2018. Ecotoxicological effects of bisphenol A and nonylphenol on the freshwater cladocerans *Ceriodaphnia silvestrii* and *Daphnia similis*. Drug and Chemical Toxicology, 41 (4), 449-458.

SPEHAR, R.L., BROOKE, L.T., MARKEE, T.P., KAHL, M.D., 2010. Comparative toxicity and bioconcentration of nonylphenol in freshwater organisms. Environmental Toxicology and Chemistry, 29 (9), 2104-2111.

SUN, D., CHEN, Q., HE, N., DIAO, P.P., JIA, L.X., DUAN, S.S., 2017. Effect of environmentally-relevant concentrations of nonylphenol on sexual differentiation in zebrafish: a multi-generational study. Nature Scientific Reports, 7, 42907.

SUN, D., CHEN, Q., ZHU, B., ZHAO, H., DUAN, S., 2021. Multigenerational reproduction and developmental toxicity, and HPG axis gene expression study on environmentally-relevant concentrations of nonylphenol in zebrafish. Science of the Total Environment, 764, 144259.

SUN, H., AND GU, X., 2005. Comprehensive toxicity study of nonylphenol and short-chain nonylphenol polyethoxylates on *Daphnia magna*. Bulletin of Environmental Contamination and Toxicology, 75 (4), 677-683.

THORLEY, J., AND SCHWARZ C. 2018. ssdtools: An R package to fit Species Sensitivity Distributions. Journal of Open Source Software, 3(31), 1082

TOFT, G. AND BAATRUP, E. 2001. Sexual characteristics are altered by 4-tert-octylphenol and 17(beta)-estradiol in the adult male guppy (*Poecilia reticulata*). Ecotoxicology and Environmental Safety, 48 (1), 76-84.

TOFT, G. AND BAATRUP, E. 2003. Altered sexual characteristics in guppies (Poecilia reticulata) exposed to 17(beta)-estradiol and 4-tert-octylphenol during sexual development. Ecotoxicology and Environmental Safety, 56 (2), 228-237.

UBA (German Federal Environment Agency). 2014. Substances of very high concern under REACH – an evaluation of uncertainties in the environmental risk assessment of endocrine active substances. Final report. TEXTE 37/2014.

UBA (German Federal Environment Agency). 2017. Validation of reproduction tests with molluscs – Establishment of an OECD test guideline for the identification of endocrine and reproductive toxic effects in snails. Final report. TEXTE 34/2017.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY (US EPA). 1996. Ecological Effects Test Guidelines. OPPTS 850.1500 Fish life cycle toxicity. EPA712–C–96–122 April 1996.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY (US EPA). 2013. Validation of the Larval Amphibian Growth and Development Assay: Integrated Summary Report. U.S. Environmental Protection Agency, Endocrine Disruptor Screening Program, Washington, D.C. 28 May 2013.

VAN DEN BELT, K., VERHEYEN, R., WITTERS, H. 2001. Reproductive effects of ethynylestradiol and 4t-octylphenol on the zebrafish (*Danio rerio*). Archives of Environmental Contamination and Toxicology, 41 (4), 458-467.

VAN DER VOET, M., TEUNIS, M., LOUTER-VAN DE HAAR, J., STIGTER, N, BHALLA, D., ROOSEBOOM, M., WEVER, K.E., KRUL, C., PIETERS, R., WILDWATER, M., VAN NOORT, V. 2021. Towards a reporting guideline for developmental and reproductive toxicology testing in *C. elegans* and other nematodes. Toxicology Research, 28, 10(6), 1202-1210.

VAN VLAARDINGEN, P.L.A., TRAAS, T.P., WINTERSEN, A.M., ALDENBERG, T. 2004. ETX 2.3. A program to calculate hazardous concentrations and fraction affected, based on normally distributed toxicity data. Bilthoven, the Netherlands: National Institute for Public Health and the Environment (RIVM). Report no. 601501028/2004, 68 pp.

WARD, T.J. AND BOERI, R.L. 1991a. Early life stage toxicity of nonylphenol to the fathead minnow (*Pimephales promelas*). Study No 8979-CMA. Report prepared for Chemical Manufacturers Association by Resource Analysts.

WARD T.J. AND BOERI, R.L. 1991b. Chronic toxicity of nonylphenol to the mysid (*Mysidopsis bahia*). Study No 8977-CMA. Report prepared for Chemical Manufacturers Association by Resource Analysts.

WATANABE, H., HORIE, Y., TAKANOBU, H., KOSHIO, M., FLYNN, K., IGUCHI, T., TATARAZAKO, N. 2017. Medaka extended one-generation reproduction test evaluating 4-nonylphenol. Environmental Toxicology and Chemistry, 36, 3254–3266.

WENZEL, A., SCHÄFERS, C., VOLLMER, G., MICHNA, H., DIEL, P. 2001. Research efforts towards the development and validation of a test method for the identification of endocrine disrupting chemicals. Final Report Contract B6-7920/98/000015.

WORLD HEALTH ORGANISATION (WHO)/INTERNATIONAL PROGRAMME ON CHEMICAL SAFETY (IPCS). 2002. Global Assessment of the State-Of-The-Science of Endocrine Disruptors. An assessment prepared by an expert group on behalf of the World Health Organization, the International Labour Organisation, and the United Nations Environment Programme. 13 July 2002.

XIE, L., LI, X-Y., LIANG, K., WU, C., WANG, H-Y., ZHANG, Y-H. 2019. Octylphenol influence growth and development of *Rana chensinensis* tadpoles via disrupting thyroid function. Ecotoxicology and Environmental Safety, 169, 747–755.

YANG, F.X., XU, Y., WEN, S. 2005. Endocrine-disrupting effects of nonylphenol, bisphenol A, and p,p'-DDE on *Rana nigromaculata* tadpoles. Bulletin of Environmental Contamination and Toxicology, 75(6), 1168-75.

YANG, L., LIN, L., WENG, S., FENG, Z., LUAN, T. 2008. Sexually disrupting effects of nonylphenol and diethylstilbestrol on male silver carp (*Carassius auratus*) in aquatic microcosms. Ecotoxicology and Environmental Safety, 71(2), 400-411.

YOKOTA, H., SEKI, M., MAEDA, M., OSHIMA, Y., TADOKORO, J., HONJO, T., KOBAYASHI, K. 2001. Life cycle toxicity of 4-nonylphenol to medaka (*Oryzias latipes*). Environmental Toxicology and Chemistry, 20 (11), 2552-2560.

YOUNG, B.J., CRISTOS, D.S., CRESPO, D.C., SOMOZA, G.M., CARRIQUIRIBORDE, P. 2020. Effects of 17α-ethinylestradiol on sex ratio, gonadal histology and perianal hyperpigmentation of *Cnesterodon decemmaculatus* (Pisces, Poeciliidae) during a full-lifecycle exposure. Ecotoxicology and Environmental Safety, 205, 111176.

ZHA, J., WANG, Z., WANG, N., INGERSOLL, C. 2007. Histological alternation and vitellogenin induction in adult rare minnow (*Gobiocypris rarus*) after exposure to ethynylestradiol and nonylphenol, Chemosphere, 66 (3), 488-495.

ZHA, J., SUN, L., ZHOU, Y., SPEAR, P.A., MA, M., WANG, Z. 2008a. Assessment of 17a-ethynlestradiol effects and underlying mechanisms in a continuous, multigeneration exposure of the Chinese rare minnows (*Gobiocypris rarus*). Toxicology and Applied Pharmacology, 226 (3), 298-308.

ZHA, J., SUN, L., SPEAR, P., WANG, Z. 2008b. Comparison of ethinylestradiol and nonylphenol effects on reproduction of Chinese rare minnows (*Gobiocypris rarus*). Ecotoxicology and Environmental Safety, 71, 390-399.

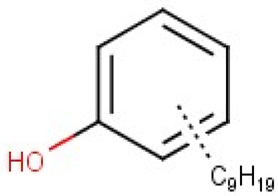
ZHOU S, DI PAOLOA C, WU X, SHAO S, SEILERA T-B, HOLLERT H. 2019. Optimization of screening-level risk assessment and priority selection of emerging pollutants – The case of pharmaceuticals in European surface waters. Environment International, 128,1-10.

ZILLIOUX, E.J., JOHNSON, I.C., KIPARISSIS, Y., METCAKLE, C.D., WHEAT, J.V., WARD, S.G., LIU, H., 2001. The sheepshead minnow as an in vivo model for endocrine disruption in marine teleosts: A partial life cycle test with 17α-ethynylestradiol. Environmental Toxicology and Chemistry, 20, 1968-1978.

Appendix A: Case Study Substance Information

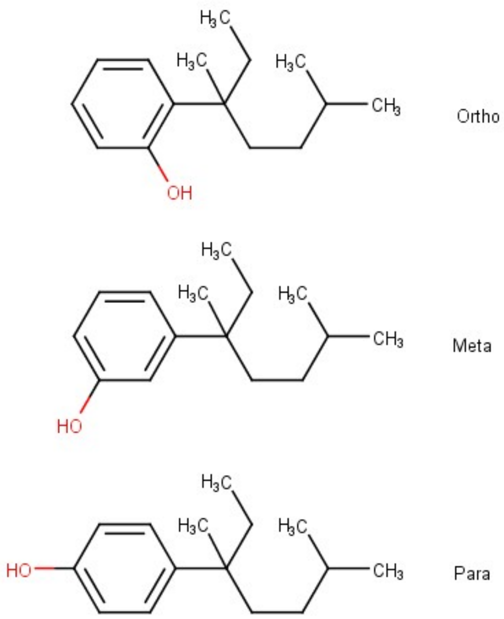
A.1 Nonylphenol

This assessment covers NP and its isomers and NPEOs, as outlined in the ECHA restriction dossier. The term "NP" may apply to a large number of substances and/or constituents thereof with the general molecular formula $C_6H_4(OH)C_9H_{19}$, in which an alkyl chain with the carbon number of nine is "attached" to the phenol (2014a). The inherent properties are however likely to be similar for all of them. The commercially produced NPs are predominantly 4-NP (also known as para-NP) with a varied and undefined degree of branching in the alkyl group. This assessment primarily focuses on 4-NP as this is the substance with the most information available. The generic structural formula for NP is shown below.



Nonylphenol isomers can vary in two primary ways:

- The position of the substitution of the nonyl group on the phenol group (ortho, meta and para positioning; shown below); and
- The degree of branching on the nonyl group.



In total, over 550 isomers are possible. It has been shown that structural features of different alkylphenols affect their oestrogenic activity, and the oestrogenic effect of an individual NP isomer is heavily dependent upon the structure of the side chain (ECHA, 2014). A list of NP isomers identified in ECHA (2014a) are shown in Table 38.

Table 38 Nonylphenol substances listed in ECHA (2014a)

Name	CAS no.	EC no.
Nonylphenol	25154-52-3	246-672-0
Isononylphenol	11066-49-2	234-284-4
Phenol, nonyl-, branched	90481-04-2	291-844-0
Phenol, 4-nonyl-, branched	84852-15-3	284-325-5
p-Nonylphenol	104-40-5	203-199-4
p-Isononylphenol	26543-97-5	247-770-6
p-(Nonan-2-yl))phenol	17404-66-9	241-427-4
p-(2-Methyloctan-2-yl)phenol	30784-30-6	250-339-5
4-(3-Methyloctan-3-yl)phenol	52427-13-1	257-907-1
o-Nonylphenol	136-83-4	205-263-7
o-Isononylphenol	27938-31-4	248-741-0
Phenol, 2-nonyl-, branched	91672-41-2	294-048-1
m-Nonylphenol	139-84-4	205-376-1
Neononylphenol-	1196678-78-0	-
4-(3,5-Dimethylheptan-3-yl)phenol	186825-36-5	-
4-(3,6-Dimethylheptan-3-yl)phenol	142731-63-3	-
2-(Nonan-2-yl)phenol	17404-45-4	-

Name	CAS no.	EC no.
Phenol, 2-tert-nonyl-	89585-68-2	-
Phenol, sec-nonyl-	97372-03-7	-
Phenol, 4-tert-nonyl-	58865-77-3	-
Phenol, o-sec-nonyl-	27214-48-8	-
Phenol, p-sec-nonyl-	27072-91-9	-

The physicochemical properties for 4-NP are shown in Table 39. Data are from ECHA (2014a).

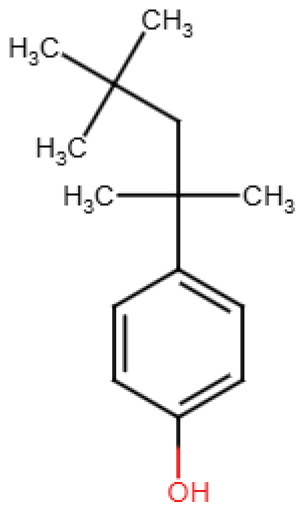
Table 39 Physicochemical properties of 4-nonylphenol (CAS: 104-40-5) based on data from ECHA (2014a)

Property	Value
Molecular weight	220.34 g/mol
Relative density	0.95 at 20 °C
Vapour pressure	0.3 Pa at 25 °C
Partition coefficient	Log K_{OW}: 4.48 – 5.4[1]
Water solubility	6 mg/L at 20 °C

[1] Range of Log K_{OW} values may relate to the variability of the commercial 4-nonyl (branched) compound, including its impurities

A.2 Octylphenol

For OP, this assessment primarily focuses on 4-tert-octylphenol (CAS 140-66-9; EC: 205-426-2) as this is the substance with the most information available and which has known oestrogenic activity. The molecular formula for 4-tert-octylphenol is $C_{14}H_{22}O$, and the structural is shown below.



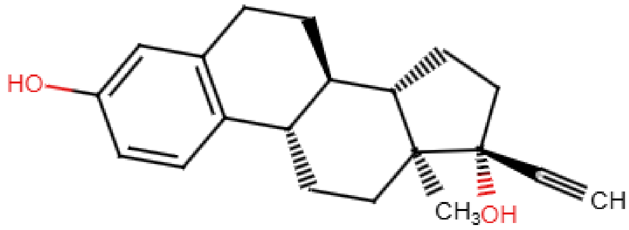
The physicochemical properties of 4-tert-octylphenol are given in the Table 40 based on data from ECHA (2012).

Table 40 Physicochemical properties of 4-tert-octylphenol (CAS 140-66-9) based on data from ECHA (2012)

Property	Value
Molecular weight	206.32 g mol^{-1}
Relative density	950 kg m^{-3} (temperature not indicated)
Vapour pressure	1.0 Pa at 20 °C
Partition coefficient	Log K_{OW}: 4.12 at 20.5 °C
Water solubility	19 mg/L at 22 °C

A.3 17α-Ethinyloestradiol

This assessment focuses on EE2 (CAS: 57-63-6; EC: 200-342-2). 17α-ethinyloestradiol is classed as an oestrogen/synthetic steroid and has the molecular formula $C_{19}H_{24}O_2$. The structural formula for EE2 is shown below.



The physicochemical properties of EE2 are given in the Table 41 based on data from JRC (2022).

Table 41 Physicochemical properties of EE2 (CAS 57-63-6) based on data from JRC (2022)

Property	Value[1]
Molecular weight	296.41 g/mol
Vapour pressure	6 $x10^{-9}$ Pa at 25 °C
Partition coefficient	Log K_{OC}: 2.92 – 5.44 Log K_{OW}: 4.2 at 25 °C
Water solubility	4.3 – 19 mg/L at 20 °C

[1] Data taken from Good Laboratory Practice (GLP) studies where available, otherwise data given as range

Appendix B: OECD Conceptual Framework for Endocrine Disruptors

The revised Conceptual Framework for Testing and Assessment of Endocrine Disrupters (OECD 2018) is given in Table 42.

Table 42 OECD Revised Conceptual Framework for Testing and Assessment of Endocrine Disrupters (OECD 2018)

	Mammalian and non-mammalian toxicology
Level 1: Existing data and existing or new non-test information	• Physical & chemical properties, e.g. MW reactivity, volatility, biodegradability • All available (eco)toxicological data from standardised or non-standardised tests • Read across, chemical categories, QSARs and other in silico predictions, and ADME model predictions
Level 2: *In vitro* assays providing data about selected endocrine mechanism(s) / pathways(s)	• Oestrogen (OECD TG 493) or androgen receptor binding affinity (US EPA TG OPPTS 890.1150) • Oestrogen receptor transactivation (OECD TG 455), yeast oestrogen screen (ISO 19040-1,2&3) • Androgen receptor transactivation (OECD TG 458) • Steroidogenesis *in vitro* (OECD TG 456) • Aromatase assay (US EPA TG OPPTS 890.1200) • Thyroid disruption assays (e.g. thyroperoxidase inhibition, transthyretin binding) • Retinoid receptor transactivation assays • Other hormone receptors assays as appropriate • High-throughput screens (See OECD GD No. 211 describing non-guideline *in vitro* test methods)

	Mammalian toxicology[3]	Non-mammalian toxicology[3]
Level 3: *In vivo* assays providing data about selected endocrine mechanism(s) / pathway(s)[1]	• Uterotrophic assay (OECD TG 440) • Hershberger assay (OECD TG 441)	• Amphibian metamorphosis assay (AMA) (OECD TG 231) • Fish short term reproduction assay (FSTRA) (OECD TG 229)[2] • 21 day fish assay (OECD TG 230) • Androgenised female stickleback screen (AFSS) (OECD GD 148) • EASZY assay. Detection of substances acting through oestrogen receptors using transgenic cyp19a1b GFP zebrafish embryos. (OECD TG 250) • *Xenopus* embryonic thyroid signalling assay XETA) (OECD TG 248) • Juvenile medaka anti-androgen screening assay (JMASA) (draft OECD GD) • Short-term juvenile hormone activity screening assay using *Daphnia magna* (draft OECD TG) • Rapid Androgen Disruption Adverse Outcome Reporter (RADAR) Assay (OECD TG 251)
Level 4: *In vivo* assays providing data on adverse effects on endocrine relevant endpoints[2]	• Repeated dose 28 day study (OECD TG 407) • Repeated dose 90 day study (OECD TG 408) • Pubertal development and thyroid function assay in peripubertal male rats (PP male assay) (US EPA TG OPPTS 890.1500) • Pubertal development and thyroid function assay in peripubertal female rats (PP female assay) (US EPA TG OPPTS 890.1450) • Prenatal developmental toxicity study (OECD TG 414)	• Fish sexual development test (FSDT) (OECD TG 234) • Larval amphibian growth & development assay (LAGDA) (OECD TG 241) • Avian reproduction assay (OECD TG 206) • Fish early life stage (ELS) toxicity test (OECD TG 210) • New guidance document on harpacticoid copepod development and reproduction test with *Amphiascus* (OECD GD 201)[2] • *Potamopyrgus antipodarum* reproduction test (OECD TG 242)[4] • *Lymnaea stagnalis* reproduction test (OECD TG 243)[4] • Chironomid toxicity test (OECD TG 218-219)[4]

	• Combined chronic toxicity and carcinogenicity studies (OECD TG 451-3) • Reproduction/developmental toxicity screening test (OECD TG 421) • Combined repeated dose toxicity study with the reproduction/developmental toxicity screening test (OECD TG 422) • Developmental neurotoxicity study (OECD TG 426) • Sub-chronic dermal toxicity: 90 day study (OECD TG 411) • Sub-chronic inhalation toxicity: 90 day study (OECD TG 413) • Repeated dose 90-day oral toxicity study in non-rodents (OECD TG 409)	• *Daphnia magna* reproduction test (with male induction) (OECD TG 211)[4] • Earthworm reproduction test (OECD TG 222)[4] • Enchytraeid reproduction test (OECD TG 220)[4] • Sediment water *Lumbriculus* toxicity test using spiked sediment (OECD TG 225)[4] • Predatory mite reproduction test in soil (OECD TG 226)[4] • Collembolan reproduction test in soil (OECD TG 232)[4]
Level 5: *In vivo* assays providing more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism[2]	• Extended one-generation reproductive toxicity study (EOGRTS) (OECD TG 443)[5] • 2-Generation reproduction toxicity study (OECD TG 416 most recent update)	• Fish lifecycle toxicity test (FLCTT) (US EPA TG OPPTS 850.1500) • Medaka extended one-generation reproduction test (MEOGRT) (OECD TG 240) • Avian 2 generation toxicity test in the Japanese quail (ATGT) (US EPA TG OCSPP 890.2100/740-C-15-003) • Sediment water chironomid life cycle toxicity test (OECD TG 233)[4] • *Daphnia* multigeneration test for assessment of EDCs (draft OECD TG)[4] • Zebrafish extended one generation reproduction test (ZEOGRT) (draft OECD TG)

[1] Some assays may also provide some evidence of adverse effects.
[2] Effects can be sensitive to more than one mechanism and may be due to non-ED mechanisms.
[3] Depending on the guideline/protocol used, the fact that a substance may interact with a hormone system in these assays does not necessarily mean that when the substance is used it will cause adverse effects in humans or ecological systems.

[4] At present, the available invertebrate assays solely involve apical endpoints which are able to respond to some EDCs and some non-EDs. Those in Level 4 are partial lifecycle tests, while those in Level 5 are full- or multiple lifecycle tests.
[5] The new EOGRT study (OECD TG 443) is preferable for detecting endocrine disruption because it provides an evaluation of a number of endocrine endpoints in the juvenile and adult F1, which are not included in the 2-generation study (OECD TG 416) adopted in 2001.

Notes to the OECD Revised Conceptual Framework:

Note 1: Entering at all levels and exiting at all levels is possible and depends upon the nature of existing information and needs for testing and assessment.

Note 2: The assessment of each chemical should be made on a case by case basis, taking into account all available information.

Note 3: The framework should not be considered as all-inclusive at the present time. It includes assays that are either available, or for which validation is under way. With respect to the latter, these are provisionally included. At Level 2 some assays are not (yet) proposed for validation but are included because they may provide information on important molecular interactions.

Appendix C: Nonylphenol dataset

Table 43 Summary of Level 3 fish and amphibian studies, and fish growth studies for nonylphenol

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Oryzias latipes	Fish short term reproduction assay Test item: 4-NP-branched	1 Test guideline study meeting validity criteria. Results are transparent and well reported (this report)	5.63, 18.8, 51.8, 170	VTG induction in males 5.63; and females 18.8 HSI in males 51.8 Reproduction 51.8	VTG induction in males < 5.63; and females 5.63 HSI in males 18.8 Reproduction 18.8	Kawashima *et al*., 2021 [30]
Oryzias latipes	Fish short term reproduction assay (21-day exposure). Semi-static exposure. Test item: 4-NP	2 (ECHA, 2012)	5.4, 16.5, 61.2	Hepatic VTG 5.4 Hepatosomatic index 16.5 Fertility 61.2 Hatchability 61.2	Hepatic VTG < 5.4 Hepatosomatic index 5.4 Fertility 16.5 Hatchability 16.5	Ishibashi *et al*., 2006 [59]
Oryzias latipes	Modified fish short term reproduction assay (21-day exposure). Flow-through exposure.	2 (ECHA, 2012)	24.8, 50.9, 101, 184	Hepatic VTG 50.9 Fecundity 101 Fertility 184	Hepatic VTG 24.8 Fecundity 50.9 Fertility 101	Kang *et al*., 2003 [60]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	Test item: 4-NP, mixture of isomers			Gonadosomatic index 184	Gonadosomatic index 101	
Pimephales promelas	Fish short term reproduction assay (21-day exposure). Flow-through exposure. Test item: 4-NP, technical grade, branched chain isomers	2 (ECHA, 2012) Data from experiment 2 not included because the assessment from ECHA (2012) was that the information should be used with care due to low reproduction performance in controls	71	Plasma VTG 71 Spawned eggs 71	Growth > 71 Plasma VTG < 71 Spawned eggs < 71	Harries *et al.*, 2000 [61]
Pimephales promelas	Fish short term reproduction assay (42-day exposure). Flow-through exposure. Test item: 4-NP	2 (ECHA, 2012)	0.05, 0.16, 0.4, 1.6, 3.4	Testis histology 1.6	Testis histology 0.4 Secondary sex characteristics > 3.4	Miles-Richardson *et al.*, 1999 [69]
Pimephales promelas	Similar to fish juvenile growth test. 28-day exposure. Flow-through exposure. Test item: 4-NP	1 (ECHA, 2014)	Not in ECHA summary	Mortality 193 Growth 77.5	Mortality 77.5 Growth 38	Brooke *et al.*, 1993b [67]
Pimephales promelas	Reproduction screening assay. 28-day exposure. Flow-through exposure. Test item: 4-NP	1 (ECHA, 2014)	Not in ECHA summary	-	Mortality > 3.2 Gonadosomatic index > 3.2 VTG > 3.2 Histology > 3.2	Schoenfuss *et al.*, 2008 [68]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Oncorhynchus mykiss	Reproduction screening assay in females (18-weeks exposure). Flow-through exposure. Test item: 4-NP, mixture of isomers	2 (ECHA, 2012)	0.7, 8.3, 85.6	Plasma VTG 8.3 Plasma oestradiol 85.6 Gonadosomatic index 85.6	Plasma VTG 0.7 Plasma oestradiol 8.3 Gonadosomatic index 8.3	Harris *et al*., 2001 [62]
Oncorhynchus mykiss	Reproduction screening assay in males (3-weeks exposure). Flow-through exposure. Test item: 4-NP, 4-substituted isomers	2 (ECHA, 2012)	0.24, 1.06, 1.85, 5.02, 20.3, 54.3	VTG 20.3 Gonadosomatic index 54.3	VTG 5.02 Gonadosomatic index 20.3	Jobling *et al*., 1996 [63]
Gobiocypris rarus	Similar to fish short term reproduction assay (21-day exposure). Flow-through exposure. Test item: 4-NP (90% 4-NP, 10% 2-NP) and 2% dinonylphenol).	1 (ECHA, 2014)	4.52, 9.13, 18.53	Testis ova 18.5 Liver lesions 9.13 Male plasma VTG 4.52	Mortality >1 8.5 Fertility > 18.5 Reproduction > 18.5 Testis ova 9.13 Liver lesions 4.52 Male plasma VTG < 4.52	Zha *et al*., 2008 [64]
Carassius auratus	Adult male (2 -3 years) exposure. 21-day semi-static Test item: NP technical	2 (ECHA, 2012)	1.39, 8.57, 64.5	Testis histopathology 64.5 Plasma oestrone and	Testis histopathology 8.57 Plasma oestrone and	Yang *et al*., 2008 [87]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
				17β-oestradiol induced 8.57 Testosterone reduced 8.57	17β-oestradiol induced 1.39 Testosterone reduced 1.39	
Lepomis macrochirus	Similar to fish juvenile growth test. 28-day exposure. Flow-through exposure. Test item: 4-NP	1 (ECHA, 2014)	Not in ECHA summary	Mortality 126	Mortality 59.5	Brooke *et al*., 1993b [65]
Salmo trutta	Similar to fish juvenile growth test. 21-day exposure. Semi-static exposure. Test item: NP, branched, technical mixture (tNP)	2 (IUCLID) 3 Not relevant due to growth being the only population-relevant endpoint. For biomarker endpoints the methods are not sufficiently well described and therefore there are many assumptions. Analytical verification of test substances is minimal assumed only at test start. Controls for statistical comparisons unclear. Replication is low (3 males and 3 female) Results are inconsistent i.e. different effects in the sexes, increased mucus does not	0.95, 8.72, 91.67 (1, 10, 100 nominal)	Gill and intestine lesions 1 Increased plasma calcium 100 Decreased TSH (not a concentration response) 1 Increased Somatotropin 1 (inverted concentration response)	Weight > 100 Length > 100 Gill and intestine lesions < 1 Increased plasma calcium 10 Decreased TSH (not a concentration response) < 1 Increased Somatotropin < 1 (inverted concentration response)	Shirdel *et al*., 2016 [69]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		reflect gill histopathology, thyroid hormone changes inconsistent (this report)				
Moxostoma hubbsi	Fish growth test. 21-day exposure. Flow-through exposure. Test item: 4-NP	2 (IUCLID)	13.2, 23.6, 49.8, 71.3		Mortality > 50	IUCLID, 2014 [70]
Menidia beryllina	Fish growth test. 28-day exposure to juveniles. Semi-static exposure. Test item: n-NP	3 Growth test not relevant. The study uses a semi-static test design (where a flow through would be preferred), without detailed reporting of phys chem conditions (e.g. oxygen levels) or biomass loading. There is a lack of reporting of endpoints (e.g. actual growth measurements are not reported only the result relative to the control), and replication is not clear (e.g. different numbers of fish per tank compared with numbers used for growth assessment). There were no concentration dependent adverse effects on growth or sex ratio (this report)	0.85, 10.58, 47.86 (1, 10, 50 nominal)	VTG induction 49.8	Mortality > 71.3 Growth > 71.3 Sex ratio > 71.3 VTG induction 23.6	Mehinto *et al.*, 2018 [35]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Xenopus laevis	Amphibian metamorphosis assay Test item: 4-NP-branched	2 Guideline study which meets all relevant validity criteria. Results are transparent and well reported (this report)	1.82, 5.17, 12.5, 48.2	Inc in SVL at all test concentrations at Day 21 (not adverse) 1.82 (11.14 – 23.48%)	< 1.82 (Inc in SVL not adverse)	Kamel *et al*., 2022 [29]

[a] Indicates the source for the reliability assessment

[b] Population-relevant in black, mechanistic endpoints in red

Table 44 Summary of Level 4 studies for nonylphenol in fish

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Oncorhynchus mykiss	Early life stage test. 91 days exposure under flow-through conditions. Growth endpoints measured at 51 days post-hatch. Test item: 4-NP	2 (ECHA, 2014)	6.0, 10.3, 23.1, 53.0, 114	Growth (reduced standard length and weight) 10.3 Larval abnormalities 53.0 Survival 23.1	Growth (reduced standard length and weight) 6 Larval abnormalities 23.1 Survival 10.3	Brooke *et al.*, 1993a; Spehar *et al.*, 2010 [1]
Oncorhynchus mykiss	Similar to fish sexual development test. 1 year exposure to embryonic, larval and juvenile stages under flow-through conditions. Test item: 4-NP	2 (ECHA, 2012)	1.05, 10.17	Hepatic VTG 1.05 Zona radiata protein 10.17	Hepatic VTG < 1.05 Zona radiata protein 1.05 Mortality > 10.17 Growth (body weight) > 10.17 Development (hatching rate, sex ratio) > 10.17 Histology: gonadal maturity and testis-ova > 10.17	Ackermann *et al.*, 2002 [2]
Oncorhynchus mykiss	Similar to fish sexual development test. 35 day exposure. Test item: 4-NP	3 (ECHA, 2014)	1, 10, 30 (nominal)	Growth (weight, length) 10 Ovosomatic index 30	Growth (weight, length) 1 Ovosomatic index 10	Ashfield *et al.*, 1998 [17]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Oryzias latipes	OPPTS 850.1500 (Fish Life Cycle Toxicity). 28-day exposure under flow-through conditions. Test item: 4-NP	1 (ECHA, 2014)	0.44, 0.54, 0.77, 1.93	-	Mortality > 1.9 Growth > 1.9	Nimrod and Benson, 1998 [3]
Oryzias latipes	Similar to fish sexual development test. 60 days exposure under flow-through conditions. Started with fertilised eggs/embryos. Test item: 4-NP	2 (ECHA, 2012)	3.30, 6.08, 11.6, 23.5, 44.7	Mortality > 44.7 Growth (length) 44.7 Growth (body weight) 23.5 External sex characteristics 44.7 Testis-ova induction 11.6 VTG induction 11.6	Mortality > 44.7 Growth (length) 23.5 Growth (body weight) 11.6 External sex characteristics 23.5 Testis-ova induction 6.08 VTG induction 6.08	Seki *et al*., 2003 [5]
Oryzias latipes	Similar to fish sexual development test. 100 days exposure under semi-static conditions. Started with fertilised eggs/embryos. Test item: 4-NP	2 (ECHA, 2012)	0.29, 0.87, 2.9, 8.7, 29	Mixed secondary sex characteristics 8.7 Testis-ova induction 29	Mixed secondary sex characteristics 2.9 Testis-ova induction 8.7 Sex ratio > 29	Balch and Metcalfe, 2006 [6]
Oryzias latipes	Fish sexual development test. 60 days exposure under flow through conditions. Started	1 Well reported study to a standardised OECD test guideline. Some minor data omissions related to oxygen content during the study	1.1, 3.0, 11.7, 31, 104.9	Number of fin ray joint plates showing papillary processes in	Number of fin ray joint plates showing papillary processes in	Horie *et al*., 2021 [32]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	with fertilised eggs/embryos. Test item: 4-NP	and non-affected endpoints (this report)		genetic male fish 3.0 Testis-ova induction 31 Sex reversal 104.9	genetic male fish 1.1 Testis-ova induction 11.7 Sex reversal 31	
Pimephales promelas	Similar to fish early life stage test. 33 days exposure to > 24 hour old embryos under flow-through conditions. Test item: 4-NP	1 (ECHA, 2014)	2.8, 4.5, 7.4, 14, 23	Mortality 14	Mortality 7.4 Growth > 23	Ward and Boeri, 1991a [7]
Danio rerio	Similar to fish early life stage test. Exposure two to 60 days post hatch. Test item: 4-NP	2 (ECHA, 2012) 3 Does not meet relevance criteria due to lack of analytical verification so used as supporting study only (this report)	10, 100 (nominal)	Fecundity 100 Gonad histology 100 VTG 100 Sex ratio 10	Fecundity 10 Gonad histology 10 VTG 10 Sex ratio < 10 Mortality > 100 Growth > 100 Hatchability > 10	Lin and Janz, 2006 [16]

[a] Indicates the source for the reliability assessment
[b] Population-relevant in black, mechanistic endpoints in red

Table 45 Summary of Level 5 lifecycle studies for nonylphenol in fish

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Oryzias latipes	Similar to OPPTS 850.1500 (Fish Life Cycle Toxicity). Exposure from ≤ 24 hour post-fertilisation for 104 days (F0 generation) and 60 days (F1 generation). Flow-through conditions. Test item: 4-NP	2 (ECHA, 2012)	4.2, 8.2, 17.7, 51.5, 183	Post-swim up mortality (F0) 17.7 Testis-ova induction 17.7 Fertility 17.7 Sex ratio 17.7 Secondary-sex characteristics 17.7	Post-swim up mortality (F0) 8.2 Testis-ova induction 8.2 Fertility 8.2 (not statistically significant) Sex ratio 8.2 Secondary-sex characteristics 8.2 Growth > 183 Mortality > 183	Yokota *et al*., 2001 [4]
Oryzias latipes	Medaka Extended One Generation Reproduction Test (MEOGRT), OECD TG 240 Test item: 4-NP	2 An early version of the MEOGRT study. Biological validity criteria parameters were met but there were some minor deviations based on physicochemical parameters (e.g. temperature above the recommended range) and the test solutions were not strictly maintained within ± 20% nominal. Some of the endpoints did not follow a concentration response (most notably	1.27, 2.95, 9.81, 27.8, 89.4	Survival (F0 males and females) > 89.4 Fertility (F0) > 89.4 Length and weight (F0) > 89.4 Intersex / sex reversal (F0) > 89.4 Hatching rate (F1) 89.4 Time-to-hatch (F1) > 89.4 Length and weight (F1 male and female sub-adults) 89.4 Intersex / sex reversal (F1 sub-adults) 27.8	Survival (F0 males and females) > 89.4 Fertility (F0) > 89.4 Length and weight (F0) > 89.4 Intersex / sex reversal (F0) > 89.4 Hatching rate (F1) 27.8 Time-to-hatch (F1) > 89.4 Length and weight (F1 adult and female sub-adults) 27.8 Intersex / sex reversal (F1 sub-adults) 9.81	Watanabe *et al*., 2017 [72] [c]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		growth of sub adults). Interpretation of the study is considered in more detail in Appendix D (this report)		Survival (F1 adult males) 89.4 Survival (F1 adult females) 9.81 Fecundity – eggs/female/day (F1) 1.27 Fertilised eggs/female/day (F1) 1.27 Fertility - % (F1) 27.8 Length (F1 adult males and females) > 89.4 Weight (F1 adult males and females) 89.4 Intersex / sex reversal (F1 adults) 89.4 Hatching rate (F2) 89.4 Time-to-hatch (F2) 27.8 Survival (F2) 89.4 HSI increased (F0 males) 9.81 HSI increased (F0 females) 1.27	Survival (F1 adult males) 27.8 Survival (F1 adult females) 2.95 Fecundity – eggs/female/day (F1) < 1.27 (LOEC/2 = 0.64) Fertilised eggs/female/day (F1) 1.27 (LOEC/2 = 0.64) Fertility - % (F1) 9.81 Length (F1 adult males and females) >89.4 Weight (F1 adult males and females) 27.8 Intersex / sex reversal (F1 adults) 27.8 Hatching rate (F2) 27.8 Time-to-hatch (F2) 9.81 Survival (F2) 27.8 HSI increased (F0 males) 2.95	

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
				VTG liver induction (F0 males) 2.95 HSI (increased F1 male and decrease female sub-adults) 89.4 GSI (F1 male sub-adults) > 89.4 GSI (decrease F1 female sub-adults) 89.4 SSC – anal fin papillae (F1 male sub-adults) 2.95 VTG liver induction (F1 male sub-adults) 9.81 HSI increase (F1 male adults) 27.8 HSI (F1 female adults) > 89.4 GSI increase (F1 male adults) 9.81 GSI increase (F1 female adults) 89.4 SSC – anal fin papillae (F1 male adults) 27.8 VTG liver induction (F1 male adults) 27.8	HSI increased (F0 females) < 1.27 VTG liver induction (F0 males) 1.27 HSI (increased F1 male and decrease female sub-adults) 27.8 GSI (F1 male sub-adults) > 89.4 GSI (decrease F1 female sub-adults) 27.8 SSC – anal fin papillae (F1 male sub-adults) 1.27 VTG liver induction (F1 male sub-adults) 2.95 HSI increase (F1 male adults) 9.81 HSI (F1 female adults) > 89.4 GSI increase (F1 male adults) 2.95 GSI increase (F1 female adults) 27.8 SSC – anal fin papillae (F1 male sub-adults) 9.81	

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
				VTG liver induction (F1 female adults) 89.4 Gonad histology (F1 adults) 27.8 Testis-ova induction (F1 adults) 27.8	VTG liver induction (F1 male adults) 9.81 VTG liver induction (F1 female adults) 27.8 Gonad histology (F1 adults) 9.81 Testis-ova induction (F1 adults) 9.81	
Danio rerio	Multigeneration test with zebrafish Test item: 4-NP branched	3 There are only three widely spaced test concentrations, exposed via a semi static method without robust analytical verification or details of physicochemical conditions. The study lacks key details on replication, statistical methods and information important for determining the validity of the test (e.g. survival and hatching success). The study only reports two endpoints relevant for PNEC derivation (growth and sex ratio). The result for growth	2.9, 20.3 and 182.6	Growth (F1, F2) 182.6 Sex ratio (increase in females) 20.3 (F1 but not F2 and not at 183)	Growth (F1, F2) 20.3 Sex ratio 2.9	Sun *et al*., 2017 [82]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		(reduction at 200 µg/L) is plausible whereas the effect on sex ratio did not follow a concentration response and was not observed in a subsequent generation (this report)				
Danio rerio	Multigeneration test with zebrafish Test item: 4-NP branched	3 There are only three widely spaced test concentrations, exposed via a semi static method without robust analytical verification. The source and acclimation of the animals used in the test is unclear. The study lacks key details on test design (tank size, number of animals per tank), replication and on how animals were sampled for analysis. Where replication is reported it seems low for this type of study (e.g. nine males or females per treatment for body length measurements). No validity criteria were	2.9, 20.3 and 182.6	Fertility (F1, F2, F3) 182.6 Hatching success (F2, F3) 20.3 Growth (F1 male) 182.6 Growth (F2 male) 20.3 Gonad damage (F1) 20.3	Fertility (F1, F2, F3) 20.3 Hatching success (F2, F3) 2.9 Growth (F1 male) 20.3 Growth (F2 male) 2.9 Gonad damage (F1) 2.9	Sun *et al.*, 2021 [83]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		reported. Performance criteria could be inferred for some of the endpoints based on the available data but because endpoints such as fertility, hatching and mortality were reported as a ratio criteria relating these to performance criteria can't be done (this report)				
Oncorhynchus mykiss	Reproduction screening assay. Intermittent exposure for 10 days / month for four months. Flow-through exposure. Test item: Technical NP consisting of 98% NP isomers (90% 4-nonylphenol, 10% 2-nonylphenol) and 2% dinonylphenol	1 (ECHA, 2014) 3 Does not meet relevance criteria due to only two test concentrations and intermittent dosing. It is therefore used as a supporting study only (this report)	1, 10 (nominal but confirmed by analyis)	Hatching rate (F0) 10 Mortality before eyed egg stage (F1) 1 Testis histopathology (F0) > 10 VTG induction male (F0) 1 VTG induction female (F1) 10 Oestradiol increase in males (F1) 10 Testosterone increase in females (F1) 10	Hatching rate (F0) 1 Mortality before eyed egg stage 1 (F1) < 1 Testis histopathology (F0) > 10 VTG Induction male (F0) < 1 VTG induction female (F1) 1 Oestradiol increase in males (F1) 1 Testosterone increase in females (F1) 1 Sex ratio based on histopathological examination > 10	Schwaiger *et al.*, 2002 [74]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	

[a] Indicates the source for the reliability assessment
[b] Population-relevant in black, mechanistic endpoints in red
[c] See Appendix 179 for summary of the outcomes from this study

Table 46 Summary of Level 4 and 5 studies for nonylphenol in other aquatic species

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Nematode						
Caenorhabditis elegans	Full life cycle test similar to ISO 10872. 72 hour exposure to juvenile nematodes under static conditions. Test item: 4-NP	1 (ECHA, 2014) 3 Recent attempts to standardise the *C. elegans* assay have found that the live food source (bacteria *Escherichia coli* at 10^9 cells/mL) may affect the study result either through metabolism of the test substance or indirectly if the substance has antibiotic activity (Beydoun *et al*., 2021; van der Voet *et al*., 2021). NP concentrations were not maintained within ± 20% in the Höss *et al*. (2002) study (mean measured concentrations being 65.6-80% of nominal over the 72-hour exposure period) and therefore it is plausible that the NP concentrations were reduced in the study due to bacteria utilising NP as a food source. Since it is not clear whether the observed effects on *C. elegans* are an adverse effect of the test item or are	40.2, 65.6, 106.5, 150.8, 189.2, 213.9, 235.2	Reproduction 40.2 Growth 65.6	Mortality > 235 Reproduction < 40.2 Growth 40.2	Höss *et al*., 2002 [8]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		an artifact of the test design the study is not reliable for PNEC derivation (this report)				
Mollusc						
Potamopyrgus antipodarum	OECD TG 242 Test item: 4-NP	1 (IUCLID)	3.82, 6.51, 16, 26.7, 72.3	Mortality 66.7 Shell length 24.7 Reproduction 14.7	Mortality 24.7 Shell length 14.7 Reproduction 6.01	IUCLID, 2020 [54]
Villosa iris	28-day growth study. Test item: 4-NP (94%)	3 A study to an ASTM test guideline with two freshwater mussels. The study is generally well reported. There were, however, issues with analytical verification which could not be explained by the authors (the highest test concentration of 200 ug/L was -85% of nominal and less than the measured concentrations at lower nominal test concentrations) therefore the actual NOEC values from these studies can't be confirmed (this report)	4, 23, 40, 83, 30[1] (nominal 200)	Reduced growth 83	Reduced growth 30 (nominal 200)	Ivey *et al*., 2018 [33]
Lampsilis siliquoidea	28-day growth study. Test item: 4-NP (94 %)	3 See above for study with *Villosa iris* (this report)	4, 23, 40, 83, 30[1] (nominal 200)	Reduced growth 83	Reduced growth 30 (nominal 200)	Ivey *et al*., 2018 [34]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Insect						
Chironomus tentans	OECD TG 219. 53-day exposure to first instar larva (< 24 hours old) under flow-through conditions. Test item: 4-NP	1 (ECHA, 2014)	12.5, 25, 50, 100, 200 (nominal)	Mortality 91	Mortality 42	Kahl *et al.*, 1997 [10]
Crustaceans						
Ceriodaphnia dubia	7-day exposure to neonates (< 24 hours old) under semi-static conditions. Test item: 4-NP	1 (ECHA, 2014)	Not in ECHA summary	Mortality 377 Reproduction 202	Mortality 202 Reproduction 88.7	England, 1995 [9]
Ceriodaphnia silvestrii	8-day exposure to juveniles under semi-static conditions according to the Brazilian National Standard (NBR 13373). Test item: NP	2 Guideline study meeting most of the required performance criteria. Data is not reported to confirm whether criteria were met for control mortality or analytical verification of test solutions and data is not reported for parent survival (this report)	15, 19, 24, 32, 42, 54	Reproduction (offspring per female) 19	Reproduction (offspring per female) 15	Spadoto *et al.*, 2018 [36]
Daphnia magna	Similar to OECD TG 211. 21-day exposure to neonates (< 24	1 (ECHA, 2014)	44.3, 63.1, 116, 215, 500	Mortality 215 Growth (length) 215	Mortality 116 Growth (length) 116	Brooke, 1993a [11] (assumed same as

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	hours old) under semi-static conditions. Test item: 4-nonylphenol			Reproduction 215	Reproduction 116	Spehar *et al.*, 2010)
Daphnia magna	Similar to OECD TG 202 (1984). 22-day exposure to neonates (< 24 hours old) under semi-static conditions. Test item: 4-NP	1 (ECHA, 2012)	1.55, 1.34, 3.45, 10.70, 47.81	Reproduction (offspring per survivor) 10.7 Reproduction (Total offspring d9) 47.81	Reproduction (offspring per survivor) 3.45 Reproduction (Total offspring d9) 10.7	Fliedner. 1993 [18]
Daphnia magna	14-day exposure to female neonates (< 24 hours old).	2 (ECHA, 2012)	100	-	Reproduction (deformed neonates) 100	Gibble and Baer. 2003 [19]
Daphnia magna	Similar to ASTM 1988. Life Cycle Toxicity Tests with Daphnia magna. Flow-through exposure to neonates (< 24 hours old). 21-day exposure. Test item: 4-NP	1 (ECHA, 2012)	44.3, 63.1, 116, 215, 500	Growth 215	Growth 116	Spehar *et al.*, 2010 [20]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Daphnia magna	Similar to OECD TG 211. 21-day exposure to neonates (< 24 hours old) under semi-static conditions.	1 (ECHA, 2012)	14, 24, 39, 71, 130, 250	Mortality 250 Growth 71 Reproduction 39	Mortality 130 Growth 39 Reproduction 24	Comber *et al*., 1993 [27]
Daphnia magna	Similar to OECD TG 211. 21-day exposure to neonates (< 24 hours old) under semi-static conditions. Test item: NP	1 (ECHA, 2014)	0.3, 1.0, 3.0, 10, 30, 100 (nominal concentrations but measured confirmed within ± 20%)	-	Mortality > 100 Reproduction > 100	Scholz, 1992 [28]
Daphnia magna	Similar to OECD TG 211 but extended to 2nd generation. 21-day exposure to neonates (< 24 hours old) (F0), under semi-static conditions. F1 generation based on three tests using the first, second and third broods of	2 Guideline study but details on physical chemical parameters, actual measured test concentrations and results for endpoints which were not affected (e.g. reproduction) are missing. It was not possible to confirm whether test validity criteria were met for reproduction, which was not affected, but survival was 100% in the controls (this report)	5, 10, 20, 40 (nominal but analytical verification confirmed ±10%)	Mortality 20 (F0) Growth 40 (F0) Reproduction > 40	Mortality 10 (F0) Growth 20 (F0) Reproduction > 40	Campos *et al*., 2016 [31]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	the F0 generation. Test item: 4-NP					
Daphnia magna	Similar to ASTM 1988. Life Cycle Toxicity Tests with *Daphnia magna*. Semi-static exposure to neonates (< 24 hours old). 21-day exposure. Test item: 4-NP	2 (ECHA, 2012)	6.2, 12, 25, 50, 100 (nominal)	Reproduction 100 Mortality > 100	Reproduction 50 Mortality > 100	Baldwin *et al*., 1997 [98]
Daphnia magna	Reproduction assay in *daphnia*. Static exposure to neonates (< 24 hours old). 21-day exposure. Test item: NP	3 (ECHA, 2014)	Not reported (ECHA 2014)	Reproduction 10	Reproduction 1	Kopf, 1997 [99]
Daphnia magna	Similar to OECD TG 211. Semi-static exposure to neonates (< 24 hours old). 21-day exposure. Test item: NP	2 (ECHA, 2012)	13, 25, 50, 100, 200 (nominal)	Reproduction 25 Mortality 50	Reproduction 13 Mortality 25	Sun and Gu, 2005 [100]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Daphnia magna	Similar to ISO 2000. Long Term Toxicity of Substances to *Daphnia Magna*. Semi-static exposure to neonates (< 24 hours old). 21-day exposure. Test item: 4-NP	2 (ECHA, 2012)	20, 40, 60, 80, 100 (nominal)	Mortality 40 Reproduction 80	Mortality 20 Reproduction 60	Brennan *et al*., 2006 [101]
Moina macrocopa	14-day study following similar principles to reproduction studies with other Cladocera. Starts with juveniles < 24 hours old. Test item: 4-NP technical mixture	3 No analytical verification (this report)	0.37, 1.1, 3.3, 10, 30 (nominal)	Reproduction 1.1	Reproduction 0.37	Jung *et al*., 2020 [41]
Rotifer						
Plationus patulus	Two generation life table study Test item: 4-NP	3 No analytical verification (this report)	15.6, 31.25, 62.5 (nominal)	Average lifespan 15.6 Reproduction 15.6	Average lifespan < 15.6 Reproduction < 15.6	Gonzalez-Perez *et al*., 2021 [73]
Brachionus havanaensis	Two generation life table study Test item: 4-NP	3 No analytical verification (this report)	7.81, 15.62, 31.25 (nominal)	Average lifespan 15.6	Average lifespan 7.81	Gonzalez-Perez *et*

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
				Reproduction 7.81 (F2)	Reproduction < 7.81	*al.*, 2021 [74]
Marine crustaceans						
Americamysis bahia	28-day exposure to neonates (< 24 hours old) under static conditions. Test item: NP	1 (ECHA, 2014)	3.9, 6.7, 9.1, 21	Growth (length) 6.7	Growth (length) 3.9	Ward and Boeri, 1991b [12]
Americamysis bahia	28-day exposure to neonates (< 24 hours old) under flow-through conditions. Test item: 4-NP	1 (ECHA, 2014)	5.94, 7.94, 9.46, 15.28, 21.06, 27.56, 49.24	Mortality 49.24 Reproduction 15.28	Mortality 27.6 Reproduction 9.5	Kuhn *et al.*, 2001 [13]
Tisbe batagliai	53-day exposure to neonates (< 24 hours old) under semi-static conditions. Test item: 4-NP	1 (ECHA, 2014)	20, 74, 300	Mortality 74	Mortality 20	Bechmann, 1999 [14]
Tigriopus japonicas	21-day reproduction study for each of two generations Test item: 4-NP	3 (ECHA, 2014)	0.01, 0.1, 1.0, and 10.0 (nominal)	1 (F0 nauplii development) 0.1 (F1 nauplii development)	0.1 (F0 nauplii development) 0.01 (F1 nauplii development)	Marcial *et al.*, 2003 [75]

[a] Indicates the source for the reliability assessment

[b] Population-relevant in black, mechanistic endpoints in red

[1] There was a problem with the supporting analysis at the highest test concentration (where an adverse effect on growth was observed.

Appendix D: Study summary for Watanabe *et al*. (2017)

Watanabe *et al*. (2017) conducted a Medaka Extended One Generation Reproduction Test (MEOGRT) with 4-NP and *O. latipes*. This was an early MEOGRT conducted as a trial of the newly adopted OECD 240 test guideline (OECD, 2015). The study met validity criteria for hatching success and survival in the controls, and for eggs per female in the reproduction phases. The temperature was 27°C and not 24 - 26°C as stipulated in the guideline but did not vary by more than 2°C. The measured test concentrations were outside of the range of ± 20% of nominal on a few occasions, most notably for the lowest test concentration. Measured concentrations of 1.27, 2.95, 9.81, 27.8 and 89.4 µg/L were used for reporting of the results, which are summarised in Table 47 (nominal concentrations: 1, 3.2, 10, 32 and 100 µg/L).

Table 47 NOEC values for each endpoint of Watanabe *et al*. (2017)

Generation / Endpoint	F0	F1 (embryo-sub adult)	F1 (adult)	F2
Hatching rate	-	27.8 (↓)	-	27.8 (↓)
Time to hatch	-	89.4	-	9.81 (↑)
Survival (2 and 4 wpf)	-	27.8	-	27.8
Survival male	89.4	-	27.8	-
Survival female	89.4	-	2.95 (2.95)	-
Eggs/female/day	89.4	-	< 1.27* (LOEC/2 = 0.64)	-
Fertilised egg/female/day	89.4	-	< 1.27* (LOEC/2 = 0.64)	-
% Fertility	89.4	-	9.81	-
Male length	89.4	< 1.27 (↓) (27.8)	89.4	-
Male weight	89.4	< 1.27 (↓) (27.8)	27.8 (↑)	-
Female length	89.4	1.27 (↓) (27.8)	9.81 (↓) (89.4)	-
Female weight	89.4	< 1.27 (↓) (27.8)	27.8 (↑)	-
HSI male	2.95 (↑)	1.27 (↑) (27.8)	1.27 (↑) (9.81)	-
HSI female	1.27 (↑)	27.8 (¯)	89.4	-
GSI male	89.4	89.4	2.95 (↑)	-
GSI female	89.4	2.95 (↓) (27.8)	9.81 (↑) (27.8)	-
SSC male	89.4	1.27 (↓)	9.81 (↓)	-
VTG male	1.27 (↑)	2.95 (↑)	9.81 (↑)	-
VTG female	89.4	9.81 (↓)	27.8 (↑)	-

Generation / Endpoint	F0	F1 (embryo-sub adult)	F1 (adult)	F2
Intersex/ sex reversal	89.4	9.81	27.8	-
Gonad phenotype	-	-	9.81	-
Testis-ova	-	-	9.81	-

- = Endpoint not determined. Direction of change in the measured endpoint are depicted by ↑ = increase or ↓ = decrease
wpf= weeks post fertilisation.
Accepted NOEC values are in black. NOEC values given in red are based on statistically significant effects either which did not follow a clear concentration response, or where the effect was < 5 % of the dilution water control. For these endpoints an alternative NOEC is given in green which is based on statistically significant effects which also follow a concentration response and where the effect is > 5 %.
* Result based on using the Jonckheere-Terpstra test (p<0.05) as recommended in OECD TG 240. Using repeated measures with Dunnett's the NOEC is 2.95 µg/L.

The outcomes from the study were as follows:

- There was a 33, 25, and 25% female mortality in F1 adults at 9.81, 27.8, and 89.4 µg/L NP, respectively compared with 4% in the control, which does not follow a clear concentration response. There was no effect on hatching survival (NOEC ≥ 27.8 µg/L) in the F2 generation. It is not possible to rule out that these effects were test substance related, but it reflects the variability in the test system.

- There were significant effects on growth of F1 sub-adults at all test concentrations, but the data was variable and did not follow a concentration response and/or was considered not to be biologically relevant (i.e. < 5% response).

- There was a significant reduction in eggs/female/ day in the F1 adult reproduction phase at all test concentrations although without following a clear concentration response (12, 11, 37, 35 and 82% reduction at 1.27, 2.95, 9.81, 27.8 and 89.4 µg/L respectively compared with the control). This was statistically significant at all test concentrations using the Jonckheere-Terpstra test (p < 0.05) (guideline recommended). The NOEC was therefore < 1.27 g/L. Using repeated measures with Dunnett's the NOEC was 2.95 µg/L. The NOEC of 2.95 µg/L was used in the ECHA registration dossier.

- There was also a significant reduction in the number of fertile eggs/female/day which is concentration dependant (13, 18, 37, 43 and 100% reduction in total fertilised eggs/day/female in the F1 adult reproduction phase at 1.27, 2.95, 9.81, 27.8 and 89.4 µg/L respectively compared with the control. It is statistically significant at all test concentrations using the Jonckheere-Terpstra test (p < 0.05) (guideline recommended). The NOEC was therefore < 1.27 µg/L. Using repeated measures with Dunnett's the NOEC is 1.27 µg/L. The NOEC of 2.95 µg/L was used in the ECHA registration dossier.

- For % fertility the NOEC was 9.81 µg/L in the F1 adults based on Jonckheere-Terpstra test ($p < 0.05$), or NOEC = 2.95 µg/L using the repeated measures ANOVA followed by Dunnett's test ($p<0.05$).

- There was a < 5% reduction in male body length of F1 sub-adult males at 1.27, 2.95, 9.81, and 27.8 µg/L. At 89.4 µg/L there was a 12% decrease in body length of F1 sub-adult males and therefore the biologically relevant NOEC = 27.8 µg/L. There was no effect on body length for F1 adult fish (12 - 15 weeks post fertilisation) but these were a selected group of fish used for pair breeding.

- There was a significant reduction in body weight of F1 sub-adult males at 1.27 (- 16%), 2.95 (- 19%), 9.81 (- 18%), and 27.8 (- 12%) µg/L, but as the effect did not follow a concentration response it was not considered to be test substance related. At 89.4 µg/L there was a 29% reduction in body weight which was considered relevant (NOEC = 27.8 µg/L). There was a significant increase in body weight for adult male fish (12 - 15 weeks post fertilisation) at 89.4 µg/L but these were a selected group of fish used for pair breeding.

- There was a significant reduction in total length of F1 sub-adult females at 2.95 (- 4%), 9.81 (- 9%), and 27.8 (- 7%) µg/L compared with the control but as the effect did not follow a concentration response it was not considered to be test substance related. At 89.4 µg/L there was a 14% reduction in body length which was considered relevant (NOEC = 27.8 µg/L). A < 5% reduction in body length of F1 adult females was not considered biologically relevant.

- There was a significant reduction in weight of F1 sub-adult females at 1.27 (- 12%), 2.95 (- 18%), 9.81 (- 29%), 27.8 (- 21%) compared with the control but as the effect did not follow a concentration response it was not considered to test substance related. At 89.4 µg/L there was a

- 42% reduction in body length which is considered relevant (NOEC = 27.8 µg/L). There was a significant increase in body weight for adult female fish (12 - 15 weeks post fertilisation) at 89.4 µg/L but these were a selected group of fish used for pair breeding. The authors suggest that this could be due to retained eggs.

- Gonadosomatic index (GSI) and hepatosomatic index (HSI) did not always follow a concentration response. These endpoints are known to be variable and may potentially have been influenced by the variability in body weight.

- Endpoints for VTG, SSC, gonad phenotype and presence of testis-ova, which are indicators for endocrine activity, all followed a clear concentration dependant response (with the exception of a decrease in VTG in sub-adult F1 females at 27.8

(12%) µg/L and 89.4 (99%) µg/L, concentrations which were potentially compromised by systemic toxicity).

The following are considered the reliable NOEC values from this study:

The NOEC for development (hatching) is 9.81 µg/L based on an increase in time to hatch in the F2 generation.

Reduced fish growth at low test concentrations (≤ 27.8 µg/L) was not convincing due to the lack of concentration response, and a lack of evidence for a corresponding effect in the adult fish. It is our assessment that the NOEC for growth was 27.8 µg/L.

The NOEC for reproduction is < 1.27 µg/L based on endpoints for eggs/female/day and the number of fertile eggs/female/day (both based on the using the Jonckheere-Terpstra test). No clear NOEC is derived for these endpoints (unbounded study), however, according to ECHA (2008) if the LOEC is a > 10 and < 20% effect the NOEC can be calculated as LOEC/2. Since the effects on egg production are 12% reduction (based on eggs per female per day) and there is a 13% reduction (based on fertilised eggs per female per day) then it is justifiable to use a NOEC of 0.64 µg/L.

The NOEC for biomarker endpoints is 1.27 µg/L based on male VTG induction in F0 and male SSC in F1 sub-adults.

Appendix E: Octylphenol dataset

Table 48 Summary of Level 3 fish and amphibian studies for octylphenol

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Oryzias latipes	Fish short term reproduction assay Test item: 4-tert-octylphenol	1 Test guideline study meeting validity criteria. Results are transparent and well reported (this report)	25.3, 82.3, 250	VTG induction male 82.3	VTG induction male 25.3	Kawashima *et al.*, 2021 [31]
Oryzias latipes	Similar to FSTRA (exposed males mated with unexposed females) Test item: 4-tert-octylphenol	1 (ECHA, 2011)	20, 41, 74, 230	Fecundity 20 Embryo abnormalities 20 VTG induction male 20	Fecundity < 20 Embryo abnormalities < 20 VTG induction male < 20	Gronen *et al.*, 1999 [54]
Danio rerio	FSTRA (Lab 1) Test item: 4-tert-octylphenol	2 Test validation study with no mortality control and acceptable spawning (this report)	3.9, 17.5, 167.1	VTG induction male 167.1	VTG induction male 17.5	US EPA, 2007 [67]
Danio rerio	FSTRA (Lab 2) Test item: 4-tert-octylphenol	2 Test validation study with no mortality and acceptable spawning (this report)	2.4, 25, 269	Number of eggs 269 VTG induction male 269	Number of eggs 25 VTG induction male 25	US EPA, 2007 [67]
Danio rerio	Similar to FSTRA Test item: 4-tert-octylphenol	1 (ECHA, 2011)	12.5, 25, 50, 100	Fertility / fecundity > 100 VTG induction male > 100	Fertility / fecundity > 100 VTG induction male > 100	Van den Belt *et al.*, 2001 [58]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
				(measured 5 days after exposure period) Female histology 25	(measured 5 days after exposure period) Female histology 12.5	
Pimephales promelas	FSTRA (Lab 4) Test item: 4-tert-octylphenol	2 Test validation study with no control mortality but low spawning however results comparable to other laboratories (this report)	2.35, 24.3, 205.6	Secondary sex characteristics 205.6 VTG induction male 205.6	Secondary sex characteristics 24.3 VTG induction male 24.3	US EPA, 2007 [67]
Pimephales promelas	FSTRA (Lab 3) Test item: 4-tert-octylphenol	2 Test validation study with no control mortality and acceptable spawning (this report)	2.35, 24.3, 205.6	Survival 205.6 VTG induction female 205.6	Survival 24.3 VTG induction male 24.3	US EPA, 2007 [67]
Cyprinodon variegatus	Similar to FSTRA (exposed males mated with unexposed females) Test item: 4-tert-octylphenol	1 (ECHA, 2011)	11.5, 33.6, 68.1	Egg viability 68.1 Male histology 68.1 VTG induction male 11.5	Egg viability 33.6 Male histology 33.6 VTG induction male < 11.5	Karels *et al*., 2003 [55]
Oncorhynchus mykiss	Similar to OECD TG 230 (21-day ED screening) Test item: 4-tert-octylphenol	1 (ECHA, 2011)	1, 10, 100	VTG induction male 10	VTG induction male 1	Routledge *et al*., 1998 [56]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Rutilus rutilus	Similar to OECD TG 230 (21-day ED screening) Test item: 4-tert-octylphenol	1 (ECHA 2011)	1, 10, 100	VTG induction male 100	VTG induction male 10	Routledge *et al.*,1998 [56]
Zoarces viviparous	Similar to FSTRA (assay for male reproductive state, no mating) Test item: 4-tert-octylphenol	2 (ECHA, 2011)	9, 35, 63	VTG induction male 35 Male histology 35	VTG induction male 9 Male histology 9	Rasmussen *et al.*, 2005 [57]
Poecilia reticulata	Similar to FSTRA (exposed males mated with unexposed females, 60-day exposure) Test item: 4-tert-octylphenol	1 (ECHA, 2011)	100, 300, 900	Male histology 100 Male secondary sex characteristics 300	Male histology < 100 Male secondary sex characteristics 100	Toft and Baatrup, 2001 [59]
Pomatoschistus minutes	28-day exposure to immature males Test item: 4-tert-octylphenol	2 (ECHA, 2011)	3, 20, 31, 101	VTG induction male 31	VTG induction male 20	Robinson *et al.*, 2004 [60]
Xenopus laevis	Amphibian metamorphosis assay Test item: 4-tert-octylphenol	2 Guideline study which meets all relevant validity criteria. Results are transparent and well reported. The dosing of test solutions was not	5.43, 13.5, 39.8, and 103	Accelerated development at 39.8 HLL and HLL:SVL 5.43, but calculations for	HLL or HLL:SVL < 5.43 Accelerated development 13.5	Kamel *et al.*, 2022 [14]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		maintained within ± 20% nominal especially at low concentrations (this report)		HLL:SVL were not stage specific (NF < 60 v NF > 60) which may have impacted the result No effect on thyroid histopathology		
Ambystoma barbouri	37-day embryo exposure Test item: Octylphenol (test substance identity not clear)	2 (ECHA, 2011)	5, 50, 500	Time to hatch 500 Larval survival 500 Snout-vent length 500	Time to hatch 50 Larval survival 50 Snout-vent length 50	Rohr *et al.*, 2003 [61]
Rana chensinensis	Short term reproduction assay style study. Male only exposure. Test item: 4-tert-octylphenol	3 No analytical verification, wild caught frogs. Non guideline study with relatively low replication (this report)	2.06, 20.6 and 206 (nominal)	Reduced spawning and hatching 206 Increased TUNEL index (measure for sperm DNA fragmentation) 20.6 Decreased GSI (at Day 10 but not at Day 30) 20.6	Reduced spawning and hatching 20.6 Increased TUNEL index 2.06 Decreased GSI (at Day 10 but not at Day 30) 2.06	Li *et al.*, 2021 [72]
Rana chensinensis	40-day development study starting	3 The eggs were field collected and therefore the overall quality is not	1.9, 20, 190	Reduced body length (Day 20) 20	Reduced body length (Day 20) 1.9	Xie *et al.*, 2019 [76]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	with tadpoles GS 20 Test item: 4-tert-octylphenol	known or the number of parent frogs. There was no solvent control. The solvent concentration varied between treatments. The density of tadpoles was high in a semi-static test but without monitoring for dissolved oxygen or iodine concentrations. There was no reporting of survival which may have affected development rates, or information on expected development rates in the control. The food was lettuce, but it was not confirmed if this was organic (this report)		Reduced body length (Day 40) 1.9 Reduced body weight (Day 40) 190 Delayed development 20 Thyroid histopathology 20	Reduced body length (Day 40) < 1.9 Reduced body weight (Day 40) 20 Delayed development 1.9 Thyroid histopathology 1.9	
Rana pipiens	Tadpole development (10 days) Test item: 4-tert-octylphenol	2 (ECHA, 2011)	0.2, 200 (nominal)	Pre-metamorphic development, body weight, day at hind-limb emergence (HLE) > 200	Pre-metamorphic development, body weight, day at hind-limb emergence (HLE) > 200	Crump *et al.*, 2002 [73]
Rana pipiens	Tadpole development started with stage	2 (ECHA, 2011)	50, 100, 150, 200, 500, 1 000, 1 500,	LC_{50} (stage 26) = 293	Body weight (stage 26) 200	Hogan *et al.*, 2006 [74]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	26 and stage 36 (2 weeks) Test item: 4-tert-octylphenol		2 000 (nominal)	Body weight (stage 26) 500 LC_{50} (stage 36) = 577.9 Increased body weight 50	Increased body weight (Stage 36) < 50	
Rana sylvatica	Tadpole development started with stage 26 (2 weeks) Test item: 4-tert-octylphenol	2 (ECHA, 2011)	50, 100, 150, 200, 500, 1 000, 1 500, 2 000 (nominal)	LC_{50} (stage 26) = 153 Body weight (stage 26) 500	Body weight (stage 26) 200	Hogan *et al.*, 2006 [74]

[a] Indicates the source for the reliability assessment
[b] Population-relevant in black, mechanistic endpoints in red

Table 49 Summary of Level 4 studies for octylphenol in fish and amphibians

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Oryzias latipes	Fish sexual development test. 60-day exposure. Flow-through exposure. Test item: 4-tert-octylphenol	1 (ECHA, 2011)	6.94, 11.4, 23.7, 48.1, 94	Sex ratio phenotype gonads 48.1 VTG 11.4 Testis-ova induction 23.7	Mortality > 94 Growth > 94 Hatchability > 94 Sex ratio phenotype gonads 23.7 VTG 6.94 Testis-ova induction 11.4	Seki *et al.*, 2003 [1]
Oryzias latipes	Fish sexual development test. 100-day exposure. Semi-static exposure. Test item: 4-tert-octylphenol	3 No analytical verification, one test concentration (this report)	100 (nominal)	Growth 100 Testis-ova induction > 100	Growth < 100 Testis-ova induction > 100	Gray *et al.*, 1999a [2]
Oryzias latipes	Fish sexual development test. 6 month exposure. Semi-static exposure. Test item: 4-tert-octylphenol	3 (ECHA, 2011) No analytical verification (this report)	10, 25, 50 (nominal)	Sex ratio > 50 Testis-ova induction 50	Sex ratio > 50 Testis-ova induction 25	Gray *et al.*, 1999b [3]
Oryzias latipes	Fish sexual development test. OECD validation Lab 4. Test item: 4-tert-octylphenol	3 Test validation study. Sex ratio outside test acceptability criteria with < 20% female (this report)	11.2, 31.7, 104.7	Sex ratio 11.2 VTG 104.7	Sex ratio < 11.2 VTG 31.7	OECD, 2012 [66]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Oryzias latipes	Fish sexual development test. OECD validation Lab 5. Test item: 4-tert-octylphenol	2 Test validation study generally meeting acceptability criteria (this report)	12.1, 30.6, 89.6	Sex ratio 12.1 VTG 12.1	Sex ratio < 12.1 VTG < 12.1	OECD, 2012 [66]
Oryzias latipes	Fish sexual development test. OECD validation Lab 9. Test item: 4-tert-octylphenol	2 Test validation study generally meeting acceptability criteria (this report)	6.2, 12.3, 23.6, 50.4, 100.6	Sex ratio 50.4 VTG 12.3	Sex ratio 23.5 VTG 6.2	OECD, 2012 [66]
Danio rerio	Fish sexual development test. OECD validation Lab 1. Test item: 4-tert-octylphenol	2 Test validation study generally meeting acceptability criteria (this report)	13.8, 40.6, 73.1	Sex ratio 13.8 VTG 40.6	Sex ratio < 13.8 VTG 13.8	OECD, 2012 [66]
Danio rerio	Fish sexual development test. OECD validation Lab 2. Test item: 4-tert-octylphenol	2 Test validation study generally meeting acceptability criteria (this report)	5.7, 17.6, 42.5	Sex ratio 17.6 VTG 42.5	Sex ratio 5.7 VTG 17.6	OECD, 2012 [66]
Danio rerio	Fish sexual development test. OECD validation Lab 4. Test item: 4-tert-octylphenol	2 Test validation study generally meeting acceptability criteria (this report)	9.5, 26, 91.5, 298.1	Sex ratio 26 VTG 26	Sex ratio 9.5 VTG 9.5	OECD, 2012 [66]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Gasterostreus aculeatus	Fish sexual development test. OECD validation Lab 6. Test item: 4-tert-octylphenol	2 Test validation study generally meeting acceptability criteria, hatching was 77% rather than required 80% (this report)	12.2, 22.2, 66.9	Sex ratio > 66.9 Hatching 66.9 VTG 66.9	Sex ratio > 66.9 Hatching 22.2 VTG 22.2	OECD, 2012 [66]
Gasterostreus aculeatus	Fish sexual development test. OECD validation Lab 8. Test item: 4-tert-octylphenol	3 Test validation study. Low hatching success and survival in solvent control (61.7 and 67.9% respectively). Test guideline recommends hatching of > 80% and post-hatch survival > 70% (this report)	41.9, 130.6, 488.9	Sex ratio 41.9 Survival 130.6 VTG 41.9	Sex ratio < 41.9 Survival 41.9 VTG < 41.9	OECD, 2012 [66]
Oncorhynchus mykiss	Fish early life stage test. 60-day exposure. Flow-through exposure. Test item: 4-tert-octylphenol	4 (ECHA, 2011) 2 (EA, 2005)	6.1, 11, 22, 51, 91	Growth 11	Growth 6.1	ABC labs, 1984 (cited in EA, 2005; ECHA, 2011) [6]
Oncorhynchus mykiss	Fish early life stage test. 22- and 35-day exposure. Flow-through exposure. Test item: 4-tert-octylphenol	3 No analytical verification (this report)	1.0, 10, 30, 50 (nominal)	Growth 1.0 (the effect did not follow a concentration response) Gonado-somatic index > 50	Growth < 1.0 Gonado-somatic index > 50	Ashfield *et al.*, 1998 [7]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Zoarces viviparous	Exposure of pregnant adult females. 35-day exposure. Flow-through exposure. Test item: 4-tert-octylphenol	2 (ECHA, 2011) 3 (Fish were wild caught and not fed during the exposure and therefore were under additional stress during the test. There was no indication for the health of the fish at the start of the test or survival during the test. There were only two test concentrations and there were effects reported at the lowest test concentration (unbounded result). Although there was supporting chemical analysis the concentrations were not maintained within ± 20% nominal. There was also wide temperature variation (10.5 - 13.5°C) and no reporting of other relevant physchem parameters (this report)	14, 65	Reduced length and weight of embryos 14 VTG 14 Testis-ova induction 65	Reduced length and weight of embryos < 14 VTG < 14 Testis-ova induction 14	Rasmussen *et al.*, 2002 [8]
Poecilia reticulata	Fish early life stage test. 70-day	2 (ECHA, 2011) 3 Single test concentration (EA, 2023)	26	Gonad histology 26	Gonad histopathology < 26	Kinnberg *et al.*, 2003 [10]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	exposure. Flow-through exposure. Test item: 4-tert-octylphenol					
Poecilia reticulata	Fish early life stage test. 90-day exposure. Flow-through exposure. Test item: 4-tert-octylphenol	2 (ECHA, 2011) 3 The experiment was based on two tests which were combined for analysis of the results. No dilution water control only a solvent control (acetone, final concentration not reported). The statistics were non-standard (i.e. treatments were normalised to their respective control and then compared with the pooled control. There is no indication of whether the controls from the two experiments were significantly different). Lack of concentration response for some endpoints (e.g. GSI and sperm count probably related to combining the experiments). Wide	1.7, 11.7, 149, 200	Increased growth of males 149 Reduced body colouration 200 Increased gonopodium length 149 Posturing behaviour 149 Female GSI 149 Sperm count 149	Increased growth of males 11.7 Reduced body colouration 149 Increased gonopodium length 11.7 Posturing behaviour 11.7 Female GSI 11.7 Sperm count 11.7	Toft and Baatrup, (2003) [11]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		spacing of some treatments e.g. > 10 between the LOEC and NOEC for growth (this report)				
Xenopus laevis	Larval amphibian growth and development assay	3 (ECHA, 2011)	2.1, 21 (nominal)	Sex ratio (shift towards female) 2.1 Development > 21 Growth > 21	Sex ratio (shift towards female) < 2.1 Development > 21 Growth > 21	Kloas *et al.*, 1999 [75]
Xenopus laevis	Larval amphibian growth and development assay (OECD TG 241) Test item: 4-tert-octylphenol	2 Test guideline study which met validity criteria. Test solutions were dosed via a saturator column and were analytically verified. Growth data is reported in the supplemental information. Due to technical differences two replicates at the top test concentration were lost before the end of the test which may have impacted the sensitivity of the study (this report)	6.3, 12.6, 25.0, 54.2	Growth > 54.2 Development > 54.2 Sex reversal > 54.2 Reduced serum thyroxine 25 Thyroid histopathology 25 Gonad histopathology (male) 6.3	Growth > 54.2 Development > 54.2 Sex reversal > 54.2 Reduced serum thyroxine 12.6 Thyroid histopathology 12.6 Gonad histopathology (male) < 6.3	Haselman *et al.*, 2015 [65]
Xenopus laevis	Larval amphibian growth and	1 Regulatory study to a standard test guideline with supporting analytical	4.97, 10.3, 19.5, 38	Growth > 38 Development > 38	Growth > 38 Development > 38	ABC labs, 2012 [78]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	development assay (OECD TG 241) Test item: 4-tert-octylphenol	verification. There were minor deviations in dosing of the test substance and in temperature during the study, but these are transparently reported and discussed (this report)		Sex ratio > 38 Reduced serum thyroxine > 38 VTG > 38	Sex ratio > 38 Reduced serum thyroxine > 38 VTG > 38	
Xenopus tropicalis	Larval amphibian growth and development style assay Test item: 4-tert-octylphenol	2 This was an early version of the LAGDA and therefore experimental design may not be optimal. There were some issues with the physicochemical parameters, but these were monitored and adjusted where necessary and there was no excessive mortality during the test (this report)	1.2, 3.5, 10, 36	Growth and development > 36 Phenotypic sex ratio > 36 Sex ratio based on gross examination of the gonad tissue 10 (no effect at 36) VTG 36 Oviduct development in males 36	Growth and development > 36 Sex ratio (percent female) > 36 Sex ratio based on gross examination of the gonad tissue 36 VTG 10 Oviduct development in males 10	Porter *et al.*, 2011 [77]
Rana pipiens	Newly hatched tadpoles (stage 25), 10 day exposure Test item: 4-tert-octylphenol	2 (ECHA, 2011) 3 Only two test concentrations with effects at the lowest dose (this report)	0.02, 2 (nominal)	Delayed development 0.02 Increased body weight (stage 29 only) 0.02	Delayed development < 0.02 Increased body weight (stage 29 only) < 0.02	Croteau *et al.*, 2009

[a] Indicates the source for the reliability assessment

[b] Population-relevant in black, mechanistic endpoints in red

Table 50 Summary of Level 5 lifecycle studies for OP in fish

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Oryzias latipes	Medaka Extended One Generation Reproduction Test (MEOGRT) (OECD TG 240) conducted by United States Environmental Protection Agency Test item: 4-tert-octylphenol	2 Validation study for the MEOGRT test. The study does not meet the control performance criteria for reproduction according to OECD 240 for all generations but if generations are combined the criteria are met. There was no reporting of physicochemical parameters during the test (this report)	6.2, 12.5, 25.1, 51.1, 102.3	Fecundity 102.3 Fertility 51.1 Survival > 102.3 Hatching 102.3 Increased growth (weight) 51 Secondary sex characteristics 6.2 Increase in *VTG* gene expression 25.1 Gonad histopathology 12.5 Liver histopathology 51.1	Fecundity 51.1 Fertility 25.1 Survival > 102.3 Hatching 51.1 Increased growth (weight) 25.1 Secondary sex characteristics < 6.2 Increase in *VTG* gene expression 12.5 Gonad histopathology 6.2 Liver histopathology 12.5	Flynn *et al*., 2017 [64]
Oryzias latipes	Medaka Extended One Generation Reproduction Test (MEOGRT) (OECD TG 240) conducted by Japan's National Institute for Environmental Studies Test item: 4-tert-octylphenol	3 Well conducted study but with some issues with reporting (it was not always possible to match the LOEC values and percent change differences to the data summaries partly because statistically significant effects were not reported). There was no analytical verification	6.25, 12.5, 25, 50, 100 µg/L (nominal)	Fecundity 100 Fertility 100 Survival 100 Hatching 25 Increased growth (length) 12.5 Secondary sex characteristics 6.25 VTG induction 6.25	Fecundity 50 Fertility 50 Survival 50 Hatching 12.5 Increased growth (length) 6.25 Secondary sex characteristics < 6.25 VTG induction < 6.25	Flynn *et al*., 2017 [64]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		of test solutions. (this report)				
Oryzias latipes	Fish full life cycle test. Full life cycle exposure. Flow-through exposure. Test item: 4-tert-octylphenol	2 (ECHA, 2022) 3 The study is assessed as not reliable for PNEC derivation. There are uncertainties regarding the purity of the test chemical, and test design. It is unclear on how many replicates (if any) are used for the developmental stage of the study and the number of fish per treatment. The numbers of fish indicated for sex ratio determination do not suggest equal number of fish per treatment (even accounting for mortalities). The maximum number of fish for sex ratio determination was 59 which is low compared with 120 fish used in OECD standard studies (e.g. minimum of 120	2, 20, 50	Reduced growth 50 Offspring mortality 2 (did not follow a concentration response) Sex ratio 2 (did not follow a concentration response)	Reduced growth 20 Offspring mortality < 2 Sex ratio < 2 Testis-ova induction > 50	Knörr and Braunbeck, 2002 [4]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		eggs per treatment for the FSDT). The statistics were not robust (only application of non-parametric approaches with no indication for whether this was the most appropriate approach). There is wide spacing of treatments, and yet endpoints for mortality and sex ratio do not follow a clear concentrations response. The 28.6% mortality at the lowest test concentration of 2 µg/L octylphenol at up to 7 days post hatch is not seen in other more contemporary studies suggesting that there were methodological issues associated with this test (this report)				
Oryzias latipes	Fish full life cycle test. Full life cycle exposure. Flow-through exposure.	4 (ECHA, 2011)	Not reported	Fertility/fecundity 82.3 Hatching success > 82.3	Fertility/fecundity 30.4 Hatching success > 82.3	Ministry of the Environment of Japan, 2002 [5]

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	Test item: 4-tert-octylphenol			Testis-ova induction 30.4 VTG 9.9	Testis-ova induction 9.9 VTG < 9.9	
Danio rerio	Fish full life cycle test. Full life cycle exposure. Flow-through exposure. Test item: 4-tert-octylphenol	1 (ECHA, 2011)	1.2, 3.2, 12, 35	Fertilisation 35 Offspring growth 35 Eggs per female 35 Time-to-first spawn 35 Sex-ratio > 35	Fertilisation 12 Offspring growth 12 Eggs per female 12 Time-to-first spawn 12 Sex ratio 35	Wenzel *et al.*, 2001 [9]

[a] Indicates the source for the reliability assessment
[b] Population-relevant in black, mechanistic endpoints in red

Table 51 Summary of Level 4 and 5 studies for octylphenol in other aquatic species

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Daphnia magna	Similar to OECD TG 211. 21-day exposure to juveniles. Test item: 4-tert-octylphenol	4 (ECHA, 2011)	37, 62, 120, 230, 510	Reproduction 120 Body length 120	Reproduction 62 Body length 62	IUCLID, 2000 [12]
Daphnia magna	Similar to OECD TG 211. 21-day exposure to juveniles. Test item: 4-tert-octylphenol	2 (ECHA, 2022)	10, 30, 100, 300, 1 000	Reproduction 100	Reproduction 30	Huls, 1992 (ECHA registration dossier, 2022) [39]
Potamopyrgus antipodarum	OECD 242. Reproduction study. Test item 4-tert-octylphenol	2 Development of a test guideline study with the test repeated over different seasons. Six tests in total one without analytical verification. Worst case result from the six studies used for PNEC derivation (this report)	0.5, 1.7, 6.2, 21, 72 (arithmetic mean of 5 studies)	Increase in embryos 1.7[c] Decrease in embryos 72[d]	Increase in embryos 0.5 Decrease in embryos 21	UBA, 2017

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Potamopyrgus antipodarum	Adult exposure, 9 weeks	3 (ECHA, 2011)	1, 5, 25, 100 (nominal)	Increase in embryos 5 Decrease in embryos 100	Increase in embryos 1 Decrease in embryos 25	Jobling *et al*., 2003 [71]
Nucella lapillus	Adult exposure, 2-3 months	3 (ECHA, 2011)	1, 10, 100 (nominal)	-	Increase in relative numbers of oocytes and in length of capsule gland and weight of female pallial glands, increase in length of penis and the prostate gland < 1	Oehlmann *et al*., 2000 [70]
Tigriopus japonicas	Similar to OECD TG 211. 21-day semi-static exposure to < 24 hour old *nauplii*. Test item: 4-tert-octylphenol	2 (ECHA, 2011) 3 No analytical verification (this report)	0.01, 0.1, 1, 10 (nominal)	Time to first copepodid stage (F1 generation) 0.01 Time to sexual development (F1 generation) 10	Time to first copepodid stage (F1 generation) < 0.01 Time to sexual development (F1 generation) 1	Marcial *et al*., 2003 [13]

[a] Indicates the source for the reliability assessment
[b] Population-relevant in black, mechanistic endpoints in red
[c] Increase in reproduction observed in 3 of 6 studies (one study had no analytical verification). Population relevance of this endpoint is unclear.
[d] Decrease in reproduction observed in 2 of 6 studies (one study had no analytical verification)

Appendix F: EE2 dataset

Table 52 Summary of amphibian studies

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Xenopus tropicalis	Modified amphibian metamorphosis assay (tadpoles exposed from Nieuwkoop and Faber stage 46 - 66). 70 tadpoles/treatment. Flow-through exposure.	Reliable (JRC, 2022) 2 (this report)	177.8,17.7, 1.77 (equivalent to 0.06, 0.6, 6 nM)	Sex ratio 17.7 Time to metamorphosis 17.7 Survival > 177.8	Sex ratio 1.77 Time to metamorphosis 1.77 Survival 177.8	Pettersson and Berg, 2007
Rana temporaria	Modified amphibian metamorphosis assay (tadpoles exposed from Gosner stage 22 - 46). 70 tadpoles/treatment. Flow-through exposure.	Reliable (JRC, 2022) 3 Egg masses collected from the wild, so may have been exposed to the test substance (or affected by viruses/ parasites, etc.). Only two concentrations tested (this report)	2.07, 26.6 (equivalent to 0.07, 0.9 nM)	Sex ratio one month post-metamorphosis 26.6 Survival > 26.6	Sex ratio one month post-metamorphosis 2.07 Survival 26.6	Pettersson and Berg, 2007

[a] Indicates the source for the reliability assessment

[b] Population-relevant endpoints highlighted in black

Table 53 Summary of Level 4 studies for EE2 in fish

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Cnesterodon decemmaculatus	90-day exposure of newborns. Semi-static exposure.	2 (this report)	32.3, 99.1, 296.2	Growth > 296.2 Survival 99.1 Sex ratio 32.3 Secondary sex characteristics 32.3	Growth 296.2 Survival 32.3 Sex ratio < 32.3 Secondary sex characteristics < 32.3	Young *et al.*, 2020
Danio rerio	Modified Fish Sex Development Test (FSDT). Larvae exposed from 43 - 71 days post fertilisation (dpf) and then held in clean water until 190 dpf. Only results from experiment 2 reported. Semi-static exposure.	1 (Matthiessen *et al.*, 2017) 2 (this report)	1.67, 3, 10	Fertility 3 Sex ratio 1.67	Fertility 1.67 Sex ratio < 1.67	Maack and Segner, 2004
Oncorhynchus mykiss	Adult male fish exposed for 62-days. Flow-through exposure.	Reliable (JRC, 2022) 2 Three test concentrations but mortality at the top concentration. An effect on fertility at the lowest test concentration was > 20% and therefore did not meet the criteria for	15.6, 131, 750	Fertility 15.6 Survival 750 Testis histopathology > 750 Gonadosomatic index 131 Plasma 11-ketotestosterone 131	Fertility < 15.6 Survival 131 Testis histopathology 750 Gonadosomatic index 15.6 Plasma 11-ketotestosterone 15.6	Schultz *et al.*, 2003

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		deriving a LOEC (this report)				
Oncorhynchus mykiss	Male fry exposed from 60-days post fertilisation for 76 days. 10 ng/L below analytical limit of detection. Semi-static exposure.	1 (Matthiessen *et al.*, 2017) 2 (this report)	10 (nominal), 80, 1 620, 9 880	Sex ratio 80 Survival 1 620 Growth 1 620 Testis histopathology 10 Testis-ova induction 80 Gonadal 11-ketotestosterone↓ 9 880 Gonadal oestradiol and testosterone > 9 880	Sex ratio 10 Survival 80 Growth 80 Testis histopathology < 10 Testis-ova induction 10 Gonadal 11-ketotestosterone 1 620 Gonadal oestradiol and testosterone 9 880	Depiereux *et al.*, 2014
Oncorhynchus mykiss	Adult males exposed from 857 days post fertilisation for 56 days, to assess effects on F1. Flow-through exposure.	1 (Matthiessen *et al.*, 2017) 2 (this report)	0.8, 8.6, 65	Survival (F1) 0.8 Plasma 11-ketotestosterone (F0) 65 Plasma luteinizing hormone (F0) 65	Survival (F1) < 0.8 Plasma 11-ketotestosterone (F0) 8.6 Plasma luteinizing hormone (F0) 8.6	Brown *et al.*, 2007
Oryzias latipes	60-day exposure of hatched embryos and then held in clean water for 6 weeks. Semi-static exposure.	1 (Matthiessen *et al.*, 2017) Reliable (JRC, 2022) 3 No analytical verification. Semi-static renewal with a 75% water change every 3 days,	1, 10, 100 (all nominal)	Survival 100 Sex ratio 100 Growth 100 Reproduction 10 Testis histopathology 100	Survival 10 Sex ratio 10 Growth 10 Reproduction 1 Testis histopathology 10	Scholz and Gutzeit, 2000

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		limited reporting of methods (this report)		Gonadosomatic index 10 Testis aromatase induction 10	Gonadosomatic index 1 Testis aromatase induction 1	
Rutilus rutilus	Modified Fish Sex Development Test (FSDT). Newly fertilised eggs exposed for 112 days. Flow-through exposure.	Reliable (JRC, 2022) 3 Two test concentrations, limited reporting of methods (this report)	0.3, 4	Sex ratio 0.3 Survival 4 Growth 4 Head aromatase induction 0.3 Whole body VTG induction 4	Sex ratio < 0.3 Survival 0.3 Growth 0.3 Head aromatase induction < 0.3 Whole body VTG induction 0.3	Lange *et al.*, 2008

[a] Indicates the source for the reliability assessment

[b] Population-relevant endpoints highlighted in black and non-population endpoints highlighted in red

Table 54 Summary of Level 5 lifecycle studies for EE2 in fish

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
Cyprinodon variegatus	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity Test; FLCTT). Partial LCTT initiated with sub-adults: 0 - 43, 0 - 59 or 0 - 73 days. F1 held in clean water until 7days post fertilisation. Spawning trials not undertaken at 328 or 723 ng/L, due to toxicity. Flow-through exposure.	Reliable (JRC, 2022) 2 (this report)	0.25, 1.73, 18.1, 117, 328, 723	Reproduction 117 Hatch (F1) 117 Survival 328 Growth > 328 Testis-ova induction 18.1 Gonad histopathology 18.1	Reproduction 18.1 Hatch (F1) 18.1 Survival 117 Growth 328 Testis-ova induction 1.73 Gonad histopathology 1.73	Zillioux *et al.*, 2001
Danio rerio	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity Test). Initiated with spawning adults. F0 exposed for 40 days, F1 exposed until 210 days post fertilisation (hpf) and F2 terminated at 100 hpf. Only continuous exposure results	1 (Matthiessen *et al.*, 2017) Reliable (JRC, 2022b) 2 (this report)	0.5, 4.5, 50 (nominal)	Fertility (F1) 0.5 Reproduction (F1) 4.5 Sex ratio (F1) 4.5 Growth (F0) > 50 Plasma VTG induction (F0) 0.5 Plasma 11-ketotestosterone (F0) 0.5	Fertility (F1) < 0.5 Reproduction (F1). 0.5 Sex ratio (F1) 0.5 Growth (F0) 50 Plasma VTG induction (F0) < 0.5 Plasma 11-ketotestosterone (F0) < 0.5	Nash *et al.*, 2004

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	reported. Results from recovery studies not reported. Analysis not determined at 50 ng/L as study terminated at day 10 due to complete F0 reproductive failure. F1: complete reproductive failure at 4.5 ng/L. Flow-through exposure.					
Danio rerio	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity Test). Initiated with embryos F0 and F1 terminated at approximately 35 days post hatch. F0: complete reproductive failure at 10 ng/L. Significant decrease in F0 78 days survival at 10 ng/L not considered biologically significant in this report (91.3%	1 (Matthiessen *et al.*, 2017) Reliable (JRC, 2022) 2 (this report)	F0: 0.05, 0.3, 1.1, 10 F1: 0.1, 0.3, 2.0	Growth (F0) 1.1 Growth (F1) 0.3 Reproduction (F1) 1.1 Reproduction (F1) 2.0 Fertility (F0) 1.1	Growth (F0) 0.3 Growth (F1) 0.1 Reproduction (F1). 0.3 Reproduction (F1) 0.3 Fertility (F0) 0.3	Wenzel *et al.*, 2001; Segner *et al.*, 2003

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L) LOEC	NOEC	Reference
	versus 94.4% in control) F1: virtually complete reproduction failure at 2.0 ng/L, so unable to assess fertility at 2.0 ng/L. Flow-through exposure.					
Danio rerio	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity Test, FLCTT). Initiated with embryos. FLCTT: F0 exposure from 0 - 177 days post fertilisation and F1 terminated at approximately 162 dpf (9.3 ng/L concentration depurated from 177 - 285 dpf). F0: complete reproductive failure at 9.3 ng/L.	1 (Matthiessen *et al*., 2017) Reliable (JRC, 2022) 2 (this report)	F0: 0.05, 0.31, 1.1, 9.3 F1: 0.09, 0.36, 2.0, NA.	Growth (F0) 1.1 Growth (F1) 2.0 Reproduction (F0) 1.1 Reproduction (F1) 2.0 Fertility (F0) 1.1 Fertility (F1) 2.0 Survival (F0) 9.3	Growth (F0) 0.31 Growth (F1) 0.36 Reproduction (F0) 0.31 Reproduction (F1) 0.36 Fertility (F0) 0.31 Fertility (F1) 0.36 Survival (F0) 1.1	Schäfers *et al*., 2007
Danio rerio	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity). Initiated with embryos	1 (Matthiessen *et al*., 2017) Reliable (JRC, 2022) 2 (this report)	0.19, 0.24, 1	Growth 0.19 Mortality (F1) 0.19 Reproduction > 23	Growth < 0.19 Mortality (F1) < 0.19 Reproduction 23	Soares *et al*., 2009

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	F0 exposure 8 months. F1 generation exposure for approx. 3.3 days. Flow-through exposure.			Gonadosomatic index > 1 Gonad histopathology > 1	Gonadosomatic index 1 Gonad histopathology 1	
Danio rerio	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity). Exposure 1 - 124 days post fertilisation. Fecundity/fertility assessed in one-off spawning events during courtship behaviour assessments. No spawning at 5.58 ng/L. Flow-through exposure.	2 (Matthiessen *et al*., 2017) 2 (this report)	0.05 (nominal), 0.5 (nominal), 5.58	Sex ratio 0.5 Reproduction 5.58 Fertility > 0.5 Secondary sex characteristics 0.05	Sex ratio 0.05 Reproduction 0.5 Fertility 0.5 Secondary sex characteristics < 0.05	Larsen *et al*., 2008
Fundulus heteroclitus	Sixty-one week exposure from eggs. Semi-static exposure.	2 (Matthiessen *et al*., 2017) 3 No reliable analytical verification and semi-static test design. No dilution water control. Concentration of solvent not reported. Lack of reporting of	0.1, 1, 10, 100 (all nominal)	Reproduction 100 Hatch 100 Growth 100 Secondary sex characteristics 100 Liver VTG induction 100 Gonad histopathology 100	Reproduction 10 Hatch 10 Growth 10 Secondary sex characteristics 10 Liver VTG induction 10	Peters *et al*., 2010

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
		some results and physchem parameters (this report)			Gonad histopathology 100	
Gobiocypris rarus	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity). Initiated with embryos. No F2 generation, as complete spawning failure in F1 at all concentrations. Flow-through exposure.	1 (Matthiessen *et al.*, 2017) Reliable (JRC, 2022) 2 (this report)	0.18, 0.91, 3.61, 13.6, 55.8	Sex ratio (F0, F1) 0.18 Fertility/reproduction (F0, F1) 0.18 Survival (F1) 0.18 Growth (F1) 0.18 Testes-ova induction (F0, F1) 0.18 Pasma VTG induction (F0, F1) 0.18 Gonadosomatic index (F0, F1) 0.18	Sex ratio (F0, F1) < 0.18 Fertility/reproduction (F0, F1) < 0.18 Survival (F1) < 0.18 Growth(F1) < 0.18 Testes-ova induction (F0, F1) < 0.18 Plasma VTG induction F0, F1) < 0.18 Gonadosomatic index (F0, F1) < 0.18	Zha *et al.*, 2008
Oryzias latipes	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity). Initiated with embryos. After four to six months of exposure males and females mated with non-exposed fish.	2 or 3 (Matthiessen *et al.*, 2017) Reliable (JRC, 2022) 3 No analytical verification, no dilution water control, limited methodology (this report)	0.2, 2, 10 (all nominal)	Fertility 10 Testis-ova induction 2	Fertility 2 Testis-ova induction 0.2	Balch *et al.*, 2004

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	Semi-static exposure.					
Pimephales promelas	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity Test). Initiated with embryos. F0, F1 terminated at 28 days post hatch. F0: reproductive pairs could only be set up from 0.16 and 0.76 ng/L exposures, as no phenotypic males were observed at ≥ 2.8 ng/L. Spawning substrates were added to the 2.8 and 12 treatments for group spawning. Flow-through exposure.	1 (Matthiessen *et al*., 2017) Reliable (JRC, 2022) 2 (this report)	0.16, 0.76, 2.8, 12, 47	Hatch (F0) >4 7 Hatch (F1) > 0.76 Adult survival (F0) > 0.76 Survival (F1) > 0.76 Growth (F0) 2.8 Growth (F1) 0.16 Reproduction (F0) > 0.76 Sex ratio (F0) 2.8 Gonad histopathology (F1) 2.8 Testes-ova induction (F0) 2.8 Plasma VTG induction (F0) 12	Hatch (F0) 47 Hatch (F1) 0.76 Adult survival (F0) 0.76 Survival (F1) 0.76 Growth (F0) 0.76 Growth (F1) < 0.16 Reproduction (F0) 0.76 Sex ratio (F0) 0.76 Gonad histopathology (F1) 0.76 Testes-ova induction (F0) 0.76 Plasma VTG induction (F0) 2.8	Länge *et al*., 2001
Pimephales promelas	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity Test). Initiated with embryos F0, F1 terminated after fertilisation assessed. F0: spawning tiles were only added to	2 (Matthiessen *et al*., 2017) Reliable (JRC, 2022) 2 (this report)	0.32 (nominal), 0.96 (nominal), 3.54, 9.55, 22.7	Sex ratio 0.32 Fertility 0.32 Reproduction 3.54 Growth (male) > 0.96 Growth (female) 22.7 Hatch, larval survival > 22.7 Gonadosomatic index 3.54	Sex ratio < 0.32 Fertility < 0.32 Reproduction 0.96 Growth (male) 0.96 Growth (female) 9.55	Parrott and Blunt, 2005

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	tanks containing fish exposed to 0.32 and 0.96 ng/L, as no phenotypic males were observed at ≥ 3.54 ng/L. Radio-imunnoassay limit of detection 0.74-1.5 ng/L). Flow-through exposure.			Hepatosomatic index 9.55 Secondary sex characteristics 9.55	Hatch, larval survival 22.7 Gonadosomatic index 0.96 Hepatosomatic index 3.54 Secondary sex characteristics 3.54	
Poecilia reticulata	Newborns exposed for 108 days. Endpoints only assessed in males. Assume fish from Kristensen *et al.*, 2005. Flow-through exposure.	Reliable (JRC, 2022) 2 (this report)	11, 44, 112	Growth 112 Sex ratio 112 Survival > 112 Secondary sex characteristics 112 Gonad histopathology 44	Growth 44 Sex ratio 44 Survival 112 Secondary sex characteristics 44 Gonad histopathology 11	Nielsen and Baatrup, 2006
Poecilia reticulata	Modified OPPTS 850.1500 (Fish Life Cycle Toxicity). Newborns exposed for 108 days and reproduction behaviour determined with unexposed males and females.	1 (Matthiessen *et al.*, 2017) 2 (this report)	10.5, 44, 112	Growth 112 Reproduction 112 Sex ratio 112 Survival > 112 Secondary sex characteristics 112 Colouration > 112 Sperm count > 112	Growth 44 Reproduction 44 Sex ratio 44 Survival 112 Secondary sex characteristics 44 Colouration 112 Sperm count 112	Kristensen *et al.*, 2005

Taxonomic group / species	Study type	Klimisch score[a]	Measured test concentration (µg/L)	Results[b] (µg/L)		Reference
				LOEC	NOEC	
	No spawning at 112 ng/L. Flow-through exposure.					
Promatoschistus minutus	Immature fish exposed for 7 months, in which EE2 was used as a positive control. Flow-through exposure.	3 (Matthiessen *et al.* 2017) Reliable (JRC 2022b) 3 Single test concentration and no analytical verification (this report)	6 (nominal)	Reproduction 6 Fertility 6 Growth 6 Secondary sex characteristics 6 Liver VTG mRNA induction 6	Reproduction < 6 Fertility < 6 Growth < 6 Secondary sex characteristics < 6 Liver VTG mRNA induction < 6	Robinson *et al.*, 2003

[a] Indicates the source for the reliability assessment

[b] Population-relevant endpoints highlighted in black, non-population endpoints highlighted in red.

Appendix G: Alternative EE2 HC5

G.1 Alternative EE2 probabilistic assessment

An HC_5 for EE2 was derived in Section 61 using only the lowest NOEC values for each species, from studies considered reliable and relevant. The resulting HC_5 of 0.029 ng/L was derived using only six NOEC values, and had wide confidence limits (0.00372 – 0.974 ng/L). The test designs of the studies including wide spacing of test concentrations, appear to contribute to the uncertainties in the SSD.

For two of the species there was more than one relevant and reliable fish study with EE2, which were not included in a SSD which is based on one result for each species (ECHA, 2008). Nevertheless, the additional studies were all valid fish life cycle or partial life cycle tests which could be used to improve confidence in the HC_5 estimate. A HC_5 was therefore derived using all reliable EE2 NOEC values, as shown in Table 55. An additional two values for *D. rerio* (Wenzel *et al*., 2001 and Schäfers *et al*., 2007) and one additional value for *On. mykiss* (Depiereux *et al*., 2014) were included in the SSD.

Table 55 All reliable values selected for deriving the updated HC_5 for EE2

Species	Endpoint (NOEC)	Value selected for HC5 (ng/L)	Reference
Danio rerio	Sex ratio F1 growth F0 growth/reproduction/fertility	0.05 0.1 0.31	Larsen *et al*., 2008 Wenzel *et al*., 2001 Schäfers *et al*., 2007
Oncorhynchus mykiss	Juvenile mortality Sex ratio	0.4 (est.) 10	Brown *et al*., 2007 Depiereux *et al*., 2014
Pimephales promelas	Reproduction	0.76	Länge *et al*., 2001
Xenopus tropicalis	Sex ratio	1.77	Pettersson and Berg, 2007
Cyprinodon variegatus	Reproduction	18.1	Zilliouz *et al*., 2001
Poecilia reticulata	Reproduction/sex ratio	44	Nielsen & Baartrup, 2006; Kristensen *et al*., 2005

The SSD constructed in ssdtools using this dataset is shown in Figure 9. The alternative “model average” HC_5 for EE2 is 0.029 ng/L, with confidence limits of 0.00698 ng/L to 0.375 ng/L.

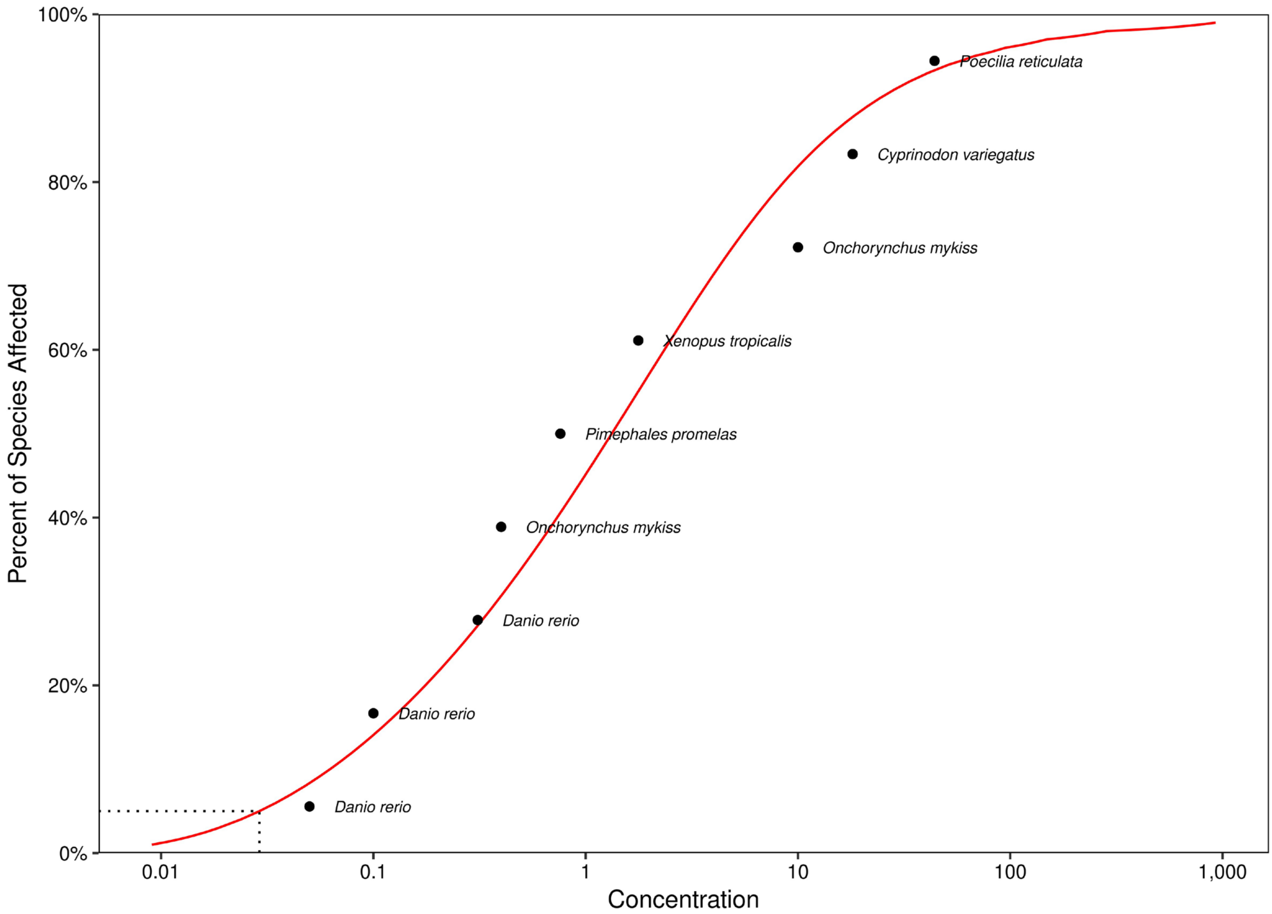


Figure 9 Updated SSD for fish and amphibian using all 9 reliable EE2 data (model average curve from ssdtools)

The inclusion of the additional NOEC values did improve the confidence limits but only slightly. The SSD was then constructed removing the two insensitive results for *C. variegatus* and *P. reticulata*. The SSD constructed in ssdtools using this dataset is shown in Figure 10. The alternative "model average" HC_5 for EE2 is 0.032 ng/L, with confidence limits of 0.0086-0.253 ng/L. Although this further improves confidence around the HC_5 value, the result is based on only four species. It is therefore recommended to use the deterministic PNEC for EE2.

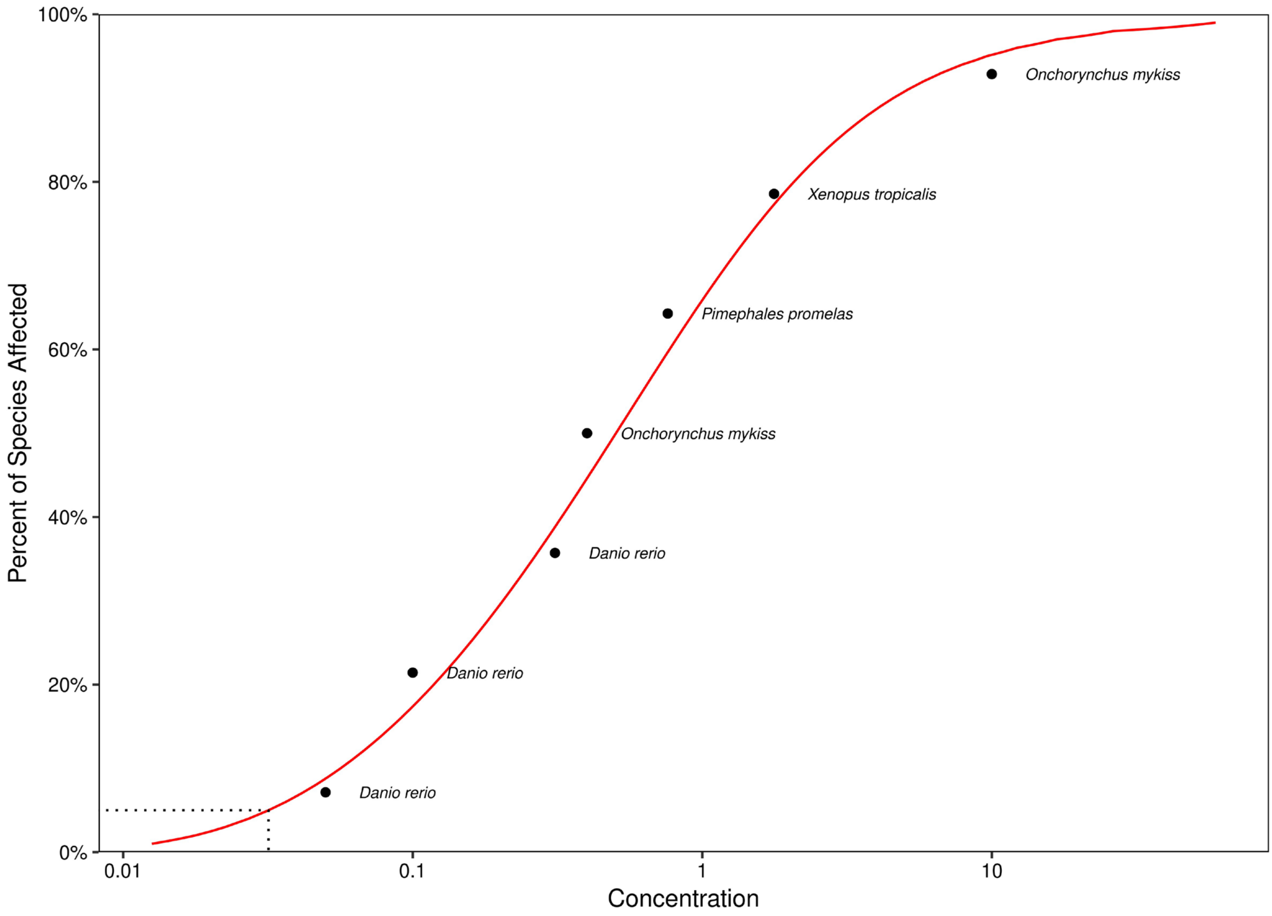


Figure 10 Updated SSD for fish and amphibian using 7 reliable EE2 data (model average curve from ssdtools), and excluding two studies where sensitive life stages were not exposed.

Appendix H: Endocrine activity for substances included in the review reported in CompTox

Table 56 Endocrine activity data for the validation substances based on data reported on the CompTox Dashboard

Substance	ER agonist[1]	ER ant-agonist	AR agonist	AR ant-agonist	Aromatase	Steroidogenesis
EE2	1.0	Inactive	Inactive	1.0	Inactive	â 11-deoxycortisol (AC_{50} = μM) â corticosterone (AC_{50} = μM) â cortisol (AC_{50} = μM) â deoxycorticosterone (AC_{50} = μM) á 17α-hydroxypregnenolone (AC_{50} = μM) â 17α-hydroxyprogesterone (AC_{50} = μM) â progesterone (AC_{50} = μM) Inactive for other hormones
Nonylphenol	0.435	Inactive	Inactive	0.227	AC_{50} = 55.0 μM; cytotox. AC_{50} = 49.0 μM. Inactive	Not reported
Octylphenol	0.393	Inactive	0.0035	0.229	AC_{50} = 20.4 μM[b,c]; cytotox. AC_{50} = 85.1 μM. active	Not reported
3-Benzylidene camphor (3-BC)	No data	-	-	-	-	-
Methyl paraben	0.0068	Inactive	Inactive	Inactive	Inactive	á progesterone (AC_{50} = 52.5 μM) á 17α-hydroxyprogesterone (AC_{50} = 41.7 μM) Inactive for other hormones

Substance	ER agonist[1]	ER ant-agonist	AR agonist	AR ant-agonist	Aromatase	Steroidogenesis
Propyl paraben	0.205	Inactive	Inactive	Inactive	Inactive	â 11-deoxycortisol (AC_{50} = 33.9 µM) â androstenedione (AC_{50} = 91.2 µM)[a] â corticosterone (AC_{50} = 34.7 µM)[b] â cortisol (AC_{50} = 69.2 µM)[a] â deoxycorticosterone (AC_{50} = 23.4 µM) á progesterone (AC_{50} = 6.46 µM) á 17α-hydroxyprogesterone (AC_{50} = 8.51 µM) á 17α-hydroxypregnenolone (AC_{50} = 21.9 µM) â testosterone (AC_{50} = 93.3 µM)[a,b] Inactive for other hormones
Fenbuconazole	Inactive	Inactive	Inactive	0.129	AC_{50} = 109.7 µM; cytotox = inactive	â cortisol (AC_{50} = 0.389 µM) á 17α-hydroxyprogesterone (AC_{50} = 0.331 µM) á progesterone (AC_{50} = 0.575 µM)[a] Inactive for other hormones
Cyproconazole	0.0268	Inactive	Inactive	Inactive	AC_{50} = 33.88 µM; cytotox = inactive	â 11-deoxycortisol (AC_{50} = 3.31 µM)[a,b] â androstenedione (AC_{50} = 5.01 µM) â cortisol (AC_{50} = 3.31 µM) â oestradiol (AC_{50} = 4.68 µM) â oestrone (AC_{50} = 3.39 µM) â 17α-hydroxypregnenolone (AC_{50} = 3.98 µM) â 17α-hydroxyprogesterone (AC_{50} = 3.16 µM) â progesterone (AC_{50} = 1.86 µM) Inactive for other hormones
Etridiazole	Inactive	Inactive	Inactive	Inactive	Inactive	á oestrone (AC_{50} = 39.8 µM)[b] Inactive for other hormones
Propiconazole	0.0000437	Inactive	Inactive	0.167	AC_{50} = 29.5 µM; cytotox AC_{50} = 74.1 µM[a,c]	â corticosterone (AC_{50} = 0.331 µM) â cortisol (AC_{50} = 0.295 µM)

Substance	ER agonist[1]	ER ant-agonist	AR agonist	AR ant-agonist	Aromatase	Steroidogenesis
						Inactive for other hormones
Prochloraz	Inactive	Inactive	Inactive	0.295	AC_{50} = 1.51 μM; cytotox AC_{50} = 77.6 μM[a,b,c]	â 11-deoxycortisol (AC_{50} = 0.347 μM) â androstenedione (AC_{50} = 0.174 μM) á corticosterone (AC_{50} = 0.0224 μM) â cortisol (AC_{50} = 0.302 μM) á 11-deoxycorticosterone (AC_{50} = 0.0501 μM) â oestradiol (AC_{50} = 0.200 μM) â oestrone (AC_{50} = 0.282 μM) â 17α-hydroxypregnenolone (AC_{50} = 0.123 μM) á progesterone (AC_{50} = 0.174 μM) â testosterone (AC_{50} = 0.174 μM) Inactive for other hormones
Vinclozolin	Inactive	Inactive	Inactive	0.416	AC_{50} = 57.5 μM[a]; cytotox AC_{50} = 66.1 μM[a]	â 11-deoxycortisol (AC_{50} = 28.2 μM) â androstenedione (AC_{50} = 35.5 μM) â 11-deoxycorticosterone (AC_{50} = 14.5 μM) â 17α-hydroxyprogesterone (AC_{50} = 23.4 μM) â progesterone (AC_{50} = 1.13 μM) â testosterone (AC_{50} = 28.2 μM) Inactive for other hormones

All data as reported on the US EPA CompTox dashboard
[1] Oestrogen receptor agonist activity based on the ToxCast Bioactivity model result
[2] Oestrogen receptor antagonist activity based on the ToxCast Bioactivity model result
[3] Androgen receptor agonist activity based on the ToxCast Bioactivity model result
[4] Androgen receptor antagonist activity based on the ToxCast Bioactivity model result
[5] AC_{50} result for aromatase inhibition (TOX21_Aromatase_Inhibition) and cytotoxicity (TOX21_Aromatase_Inhibition_viability) in the associated assay.
[6] Steroidogenesis effects (AC_{50}) based on the result for hormone responses in H295R cell line (up or down).
Data flags: [a] only the highest above baseline, [b] borderline activity, [c] Less than 50% efficacy.

Appendix I: Decision tree for applying AF to account for ED properties adapted from James *et al*., (2023)

The decision tree for applying additional AF to account for ED properties adapted from James *et al*., (2023) is given in Figure 11. For this report the validation substances are assumed to be confirmed ED (although for some of the substances a regulatory decision on the ED properties may not have been confirmed). The adapted decision tree therefore focuses on confirmed EDCs. Further information on data availability requirements for application of AF to unconfirmed EDCs can be found in James *et al*. (2023).

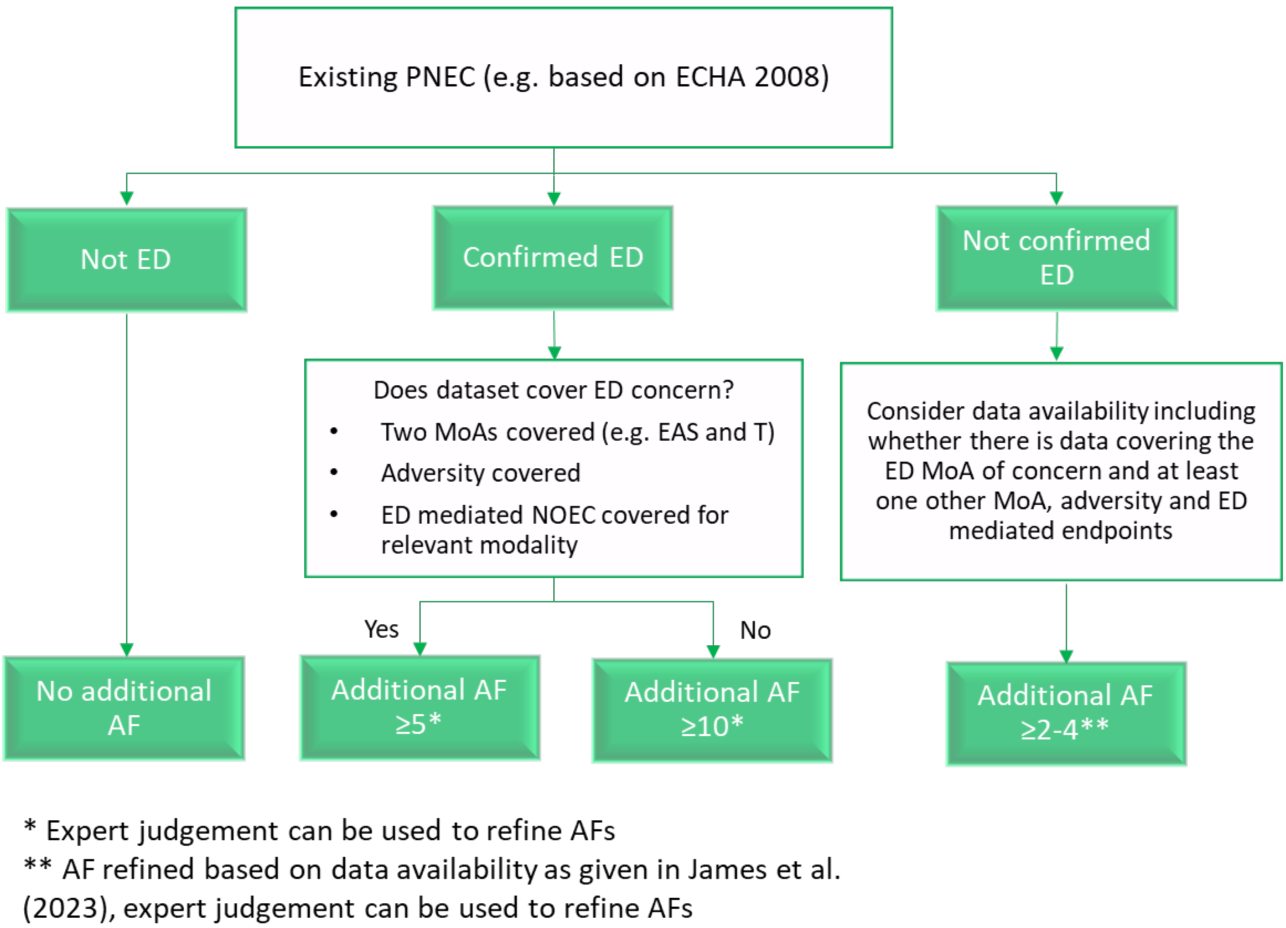
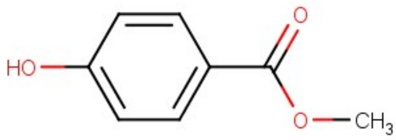


Figure 11. A proposed decision tree for applying AF to account for ED properties adapted from James *et al*. (2023)

Appendix J: Validation substance data summaries

J.1. Methyl paraben

This section focuses on methyl paraben (CAS: 99-76-3; EC: 202-785-7). Methyl paraben has the molecular formula $C_8H_8O_3$, and the structural formula is shown below.



The physicochemical properties of methyl paraben are shown in Table 57, based on data from IUCLID (2022).

Table 57 Physicochemical properties of methyl paraben (CAS: 99-76-3) based on data from IUCLID (2022)

Property	Value
Molecular weight	152.15
Density	1.3775 g/cm^3 at 20 °C
Vapour pressure	0.000028 Pa at 20°C
Partition coefficient	log Kow: 1.66 - 1.98
Water solubility	1 880 mg/L at 20 °C

Aquatic toxicity data for methyl paraben is available from the substance dossier in the ECHA Dissemination Portal (IUCLID, 2022b). The aquatic dataset includes two fish sexual development tests (Level 4) and one chronic aquatic invertebrate assay (Level 4) which are potentially reliable and relevant for PNEC derivation (Table 58). Chronic algae data is also reported. The lowest chronic effect level was reported in a fish sexual development study conducted according to OECD TG 234 (Level 4) with *Danio rerio*, which determined a NOEC of 24 µg/L for survival (IUCLID, 2022b). There is also a result from a FSTRA with the Japanese medaka (*Oryzias latipes*) conducted as part of the Ministry of the Environment of Japan's fourth program on endocrine disrupting effects of chemical substances (Kawashima *et al*., 2021). A PNEC was deterministically derived by applying an AF of 10 (ECHA, 2008, Table R.10-4), resulting in a PNEC of 2.4 µg/L (EA, 2020b).

Table 58 Summary of reliable data relevant for chronic PNEC derivation for methyl paraben (IUCLID, 2022b)

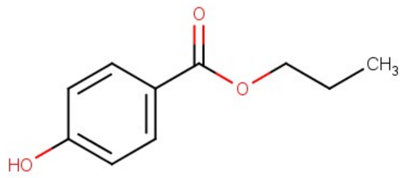
Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
Algae and aquatic plants	*Raphidocelis subcapitata*	Growth inhibition (ISO 8692)	NOEC: 20 000 (based on 72-hour growth rate) EC_{10}: 31 000 (based on 72-hour growth rate)	IUCLID, 2022b
Aquatic invertebrates	*Daphnia magna*	Reproduction (OECD TG 211)	NOEC: 200 (based on reproduction)	IUCLID, 2022b
Fish	*Oryzias latipes*	FSTRA (OECD TG 229)	NOEC: 357 (male VTG induction) NOEC: 1 900 (reduction in number of fertilised eggs and total number of eggs)	Kawashima *et al*., 2021
Fish	*Danio rerio*	2 FSDT (OECD TG 234)	FSDT 2 (definitive study) NOEC: 7.96 (based on female VTG) NOEC: 24 (based on total length at 35 days post-fertilisation [dpf], total survival and wet weight of all fish at 70 dpf) NOEC: 78.4 (based on hatching	IUCLID, 2022b

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
			success and female weight at 70 dpf) FSDT 1 (repeated because fish did not achieve sexual maturity within the test) NOEC: ≥ 250 (based on post-hatch survival at 35 dpf, survival between 35 and 70 dpf, male and female length and total length of all fish at 70 dpf, male weight at 70 dpf, sex ratio, % male, % female, % undifferentiated, male VTG, undifferentiated VTG)	
	Danio rerio	FSDT (OECD TG 234)	NOEC: 110 (based on post-hatch survival at 35 dpf and total survival at 63 dpf) NOEC: 2 780 (based on total length at 35 dpf and wet weight of all fish at 63 dpf) NOEC: ≥ 9 640 (based on hatching success and total	IUCLID, 2022b

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
			length of all fish at 63 dpf)	

J.2. Propyl paraben

This section focuses on propyl paraben (CAS: 94-13-3; EC: 202-307-7). Propyl paraben has the molecular formula $C_{10}H_{12}O_3$, and the structural formula is shown below.



The physicochemical properties of propyl paraben are shown in Table 59, based on data from IUCLID (2021b).

Table 59 Physicochemical properties of propyl paraben (CAS: 94-13-3) based on data from IUCLID (2021b)

Property	Value
Molecular weight	180.20
Density	1.287 g/cm³ at 20 °C
Vapour pressure	0.00034 Pa at 20°C
Partition coefficient	2.34 – 3.04 (average: 2.80)
Water solubility	500 mg/L at 25°C

Aquatic toxicity data for propyl paraben is available from the substance dossier in the ECHA Dissemination Portal (IUCLID, 2021b). The aquatic dataset includes one fish sexual development test (Level 4) and one chronic aquatic invertebrate assay (Level 4) which are potentially reliable and relevant for PNEC derivation (Table 60). Chronic algae data is also reported. The lowest chronic effect level was reported in a fish sexual development study conducted according to OECD TG 234 (Level 4) with *Danio rerio*, which determined a NOEC of 165 µg/L for sex ratio and proportion of females (IUCLID, 2021b). A PNEC can be deterministically derived by applying an AF of 10 (ECHA, 2008, Table R.10-4), resulting in a PNEC of 16.5 µg/L. A deterministic PNEC of 5 µg/L was previously derived by applying an AF of 50 on the *Daphnia* reproduction NOEC of 250 µg/L as no chronic fish studies were available at that time (EA, 2020b). There is also a result from a FSTRA with the Japanese medaka (*Oryzias latipes*) conducted as part of the Ministry of the Environment of Japan's fourth program on endocrine disrupting effects of chemical substances (Kawashima *et al*., 2021).

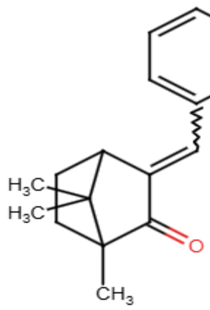
Table 60 Summary of reliable data relevant for chronic PNEC derivation for propyl paraben (IUCLID, 2021b)

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
Algae and aquatic plants	*Raphidocelis subcapitata*	Growth inhibition (OECD TG 201)	NOEC: 2 100 (based on 72-hour growth rate and yield) EC_{10}: 5 900 (based on 72-hour growth rate) EC_{10}: 3 500 (based on 72-hour yield)	IUCLID, 2021b
Aquatic invertebrates	*Daphnia magna*	Reproduction (OECD TG 211)	NOEC: 250 (based on mobility, length, reproduction and development) EC_{10}: 160 (based on reproduction) LC_{10}: 440 (based on mobility) EC_{10}: 580 (based on development rate) EC_{10}: 690 (based on length)	IUCLID, 2021b
Fish	*Oryzias latipes*	FSTRA (OECD TG 229)	LOEC: 311 (male VTG induction) NOEC: 311 (reduction in number of fertilised eggs)	Kawashima *et al.*, 2021
Fish	*Danio rerio*	FSDT (OECD TG 234)	NOEC: 165 (based on sex ratio and % females) NOEC: ≥ 518 (based on post-hatch survival at 35 and 70 days	IUCLID, 2021b

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
			post-fertilisation [dpf], % males and female VTG content)	

J.3. 3-Benzylidene camphor (3-BC)

This is a summary of the available data for 3-Benzylidene camphor (3-BC) (CAS: 15087-24-8; EC: 239-139-9). 3-BC has the molecular formula $C_{17}H_{20}O$, and the structural formula is shown below.



The physicochemical properties of 3-BC are shown in Table 61, based on data from ECHA (2016).

Table 61 Physicochemical properties of 3-BC (CAS: 15087-24-8) based on data from ECHA (2016)

Property	Value
Molecular weight	240.35
Density	1.079 g/cm^3 at 20°C (calculated)
Vapour pressure	4.45 x 10^{-3} Pa at 25°C (calculated)
Partition coefficient	Log K_{OW}: 5.37 (calculated)
Water solubility	0.6893 mg/L at 25 °C (calculated)

3-BC is not registered under REACH and therefore there is no data for this substance in the ECHA Dissemination Portal. However, a SVHC support document for environmental

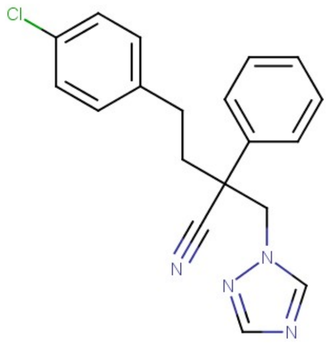
ED properties is available and contains aquatic ecotoxicity data to support PNEC derivation (ECHA, 2016). The aquatic dataset includes two fish ED screening studies (Level 3) and one AMA (Level 3) which are reliable and potentially relevant for PNEC derivation (Table 62). Chronic algae data was not reported. The lowest chronic effect level was reported in a FSTRA study (Level 3) conducted with *Pimephales promelas*, which determined a NOEC of 33 µg/L based on reproduction, male VTG induction and male secondary sex characteristics (Kunz *et al*., 2006). Although the NOEC for testis and ovary histology is lower, this endpoint is not relevant for PNEC derivation. A tentative PNEC was deterministically derived by applying an AF of 100 (ECHA, 2008, Table R.10-4), resulting in a PNEC of 0.33 µg/L (EA, 2021). A FSTRA is a mechanistic study used in the absence of other information however it would not usually be suitable for PNEC derivation because the relatively short exposure period is considered insufficient for assessing reproductive effects.

Table 62 Summary of reliable data relevant for chronic PNEC derivation for 3-BC (ECHA, 2016)

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
Algae and aquatic plants	None reported			ECHA, 2016
Amphibian	*Xenopus laevis* (frog)	AMA (OECD TG 231)	NOEC: ≥ 50 (based on metamorphosis rate, length and sex ratio)	ECHA, 2016
Fish	*Pimephales promelas*	FSTRA (OECD TG 229)	NOEC: 0.5 (based on testis and ovarian histology) NOEC: 33 (based on reproduction, male VTG induction and male secondary sex characteristics)	ECHA, 2016
	Pimephales promelas	14-day ED screening	NOEC: 100 (based on growth and VTG induction)	ECHA, 2016

J.4. Fenbuconazole

This is a data summary for fenbuconazole (CAS: 114369-43-6; EC: 406-140-2). Fenbuconazole has the molecular formula $C_{19}H_{17}ClN_4$, and the structural formula is shown below.



The physicochemical properties of fenbuconazole are shown in Table 63, based on data from IUCLID (2021c).

Table 63 Physicochemical properties of fenbuconazole (CAS: 114369-43-6) based on data from IUCLID (2021c)

Property	Value
Molecular weight	336.8
Density	1.28 at 20°C
Vapour pressure	3.4×10^{-7} Pa at 25°C
Partition coefficient	Log K_{OW}: 3.79
Water solubility	2.47 mg/L at 20°C

Aquatic toxicity data for fenbuconazole is available from an EU Draft Assessment Report (DAR) (EFSA, 2010a). The aquatic dataset includes one fish ED screening assay (Level 3), one fish early life stage study (Level 4), one fish full life cycle study (Level 5) and two aquatic invertebrate reproduction assays (Level 4) which are reliable and relevant for PNEC derivation (Table 64). Chronic algae data is available in the substance dossier in the ECHA Dissemination Portal (IUCLID, 2021c). The lowest chronic effect level was reported in the FLCTT study (Level 5) conducted with *Pimephales promelas*, which determined a NOEC of 23 µg/L based on reproduction (number of spawns); biomarkers and histology was not assessed in this study. The PNEC was deterministically derived by applying an AF of 10 (ECHA, 2008, Table R.10-4), resulting in a PNEC of 2.3 µg/L (EA, 2023).

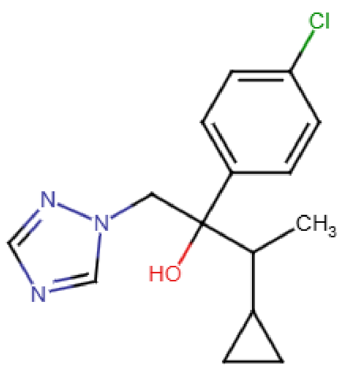
Table 64 Summary of reliable data relevant for chronic PNEC derivation for fenbuconazole (EFSA, 2010a, IUCLID, 2021c)

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
Algae and aquatic plants	*Selenastrum capricornutum*	Growth inhibition (OECD TG 201)	NOEC: 25 (based on 72-hour growth rate) NOEC: 77 (based on 96-hour growth rate) NOEC: 290 (based on 120-hour growth rate)	IUCLID, 2021c
	Desmodesmus subspicatus	Growth inhibition (OECD TG 201)	NOEC: 125 (based on 96-hour growth rate)	IUCLID, 2021c
	Pseudokirchneriella subcapitata	Growth inhibition (OECD TG 201)	NOEC: 324 (based on 72-hour growth rate)	IUCLID, 2021c
	Lemna gibba	Growth inhibition (OECD TG 221)	NOEC: 237 (based on biomass, growth rate, dry weight and frond number)	IUCLID, 2021c
Aquatic invertebrates	*Daphnia magna*	Reproduction (OECD TG 211)	NOEC: 78 (based on reproductive capacity, immobility and adult length)	EFSA, 2010a
	Chironomus riparius	31-day reproduction (OECD TG 218)	NOEC: 1 730 (development and emergence)	EFSA, 2010a
Fish	*Oncorhynchus mykiss*	21-day prolonged toxicity	NOEC: 320 (based on body weight and length)	EFSA, 2010a

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
		(OECD TG 204)		
	Pimephales promelas	FELS (OECD TG 210)	NOEC: 82 (based on fry wet weight)	EFSA, 2010a
	Pimephales promelas	FLCTT (US EPA OPPTS 850.1500)	NOEC: 23 (based on number of spawns)	EFSA, 2010a

J.5. Cyproconazole

This is a data summary for cyproconazole (CAS: 94361-06-5, 94361-07-6, 113096-99-4; EC: 619-020-1). Cyproconazole has the molecular formula $C_{15}H_{18}ClN_3O$, and the structural formula is shown below.



The physicochemical properties of fenbuconazole are shown in Table 65, based on data from EFSA (2010b).

Table 65 Physicochemical properties of cyproconazole based on data from EFSA (2010b)

Property	Value
Molecular weight	291.8
Vapour pressure	2.6 x 10^{-5} Pa at 25 °C
Partition coefficient	Log K_{OW}: 3.09
Water solubility	93 mg/L at 22 °C

Aquatic toxicity data for cyproconazole is available from a European Union Draft Assessment Report (EU DAR) (EFSA, 2010b) and in the ECHA registration portal (ECHA, 2023). The aquatic dataset includes one 21-day fish assay (Level 3), one fish short-term reproduction assay (Level 3), two fish early life stage studies (Level 4), two fish full life cycle studies (Level 5) and three aquatic invertebrate reproduction assays (Level 4) which are reliable and relevant for PNEC derivation (Table 66). Chronic algal toxicity was also reported. The lowest chronic effect level relevant for PNEC derivation was reported in an algal growth test, with a NOEC of 21 µg/L. The PNEC was deterministically derived by applying an AF of 10 (ECHA, 2008, Table R.10-4), resulting in a PNEC of 2.1 µg/L, and is the PNEC used in the ECHA registration dossier.

For the preparation of the second priority substance watch list under the Water Framework Directive, the JRC identified two PNECs for cyproconazole in freshwater (JRC, 2016). The lowest PNEC was 0.6 µg/L, which was deterministically derived by applying an AF of 10 (ECHA, 2008, Table R.10-4) on a NOEC of 6 µg/L for *Raphidocelis subcapitata* (INERIS, 2011). This algae study is 120 hours duration and therefore does not follow the standard test design. The result is therefore not used as part of the present assessment. The dataset used for this PNEC derivation is also summarised in Table 66.

Table 66 Summary of reliable data relevant for chronic PNEC derivation for cyproconazole (EFSA, 2010b; INERIS, 2011)

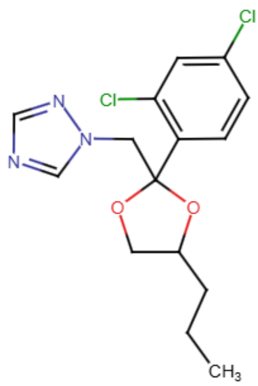
Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
Algae and aquatic plants	*Desmodesmus subspicatus*	Growth inhibition (OECD TG 201)	NOEC: 21 (based on 96-hour biomass)	EFSA, 2010b; INERIS, 2011
	Chlorella vulgaris	Growth inhibition (OECD TG 201)	NOEC: 392 (based on 72-hour growth)	EFSA, 2010b; INERIS, 2011
	Pseudokirchneriella subcapitata	Growth inhibition (OECD TG 201)	NOEC: 6 (based on 120-hour abundance, non-standard test design)	INERIS, 2011
	Lemna gibba	Growth inhibition (OECD TG 221)	NOEC: 12 500 (based on biomass and dry weight)	EFSA, 2010b; INERIS, 2011
Aquatic invertebrates	*Daphnia magna*	Reproduction (OECD TG 211)	NOEC: 290 (based on reproduction)	EFSA, 2010b; INERIS, 2011
	Daphnia magna	Reproduction (OECD TG 211)	NOEC: 23 (based on reproduction)	EFSA, 2010b
	Chironomus riparius	Emergence and development (OECD TG 218)	NOEC: 5 000 (based on development rate)	EFSA, 2010b; INERIS, 2011
	Daphnia magna	Reproduction (OECD TG 211)	NOEC: 19 (further details not publicly available)	INERIS, 2011

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
	Daphnia magna	Reproduction (OECD TG 211)	NOEC: 19 000 (further details not publicly available)	INERIS, 2011
Fish	*Oncorhynchus mykiss*	21-day prolonged toxicity (OECD TG 204)	NOEC: 650 (based on clinical effects (loss of coordination)) NOEC: 5 040 (based on body weight and length)	EFSA, 2010b; INERIS, 2011
	Pimephales promelas	FSTRA (OECD TG 229)	NOEC: ≥ 2 000 (based on reproduction, male and female SSC, histology and VTG)	EFSA, 2010b; INERIS, 2011; IUCLID, 2024
	Oncorhynchus mykiss	FELS (OECD TG 210)	NOEC: < 160 (based on fry growth, study conducted in 1993)	EFSA, 2010b; INERIS, 2011
	Oncorhynchus mykiss	FELS (OECD TG 210)	NOEC: ≥ 305 (based on hatching success, survival and larval growth, study conducted in 2006)	EFSA, 2010b; INERIS, 2011
	Pimephales promelas	FLCTT (US EPA OPPTS 850.1500)	NOEC: 125 (based on egg production) NOEC: 500 (based on VTG inhibition)	EFSA, 2010b; INERIS, 2011

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
	Pimephales promelas	FLCTT (US EPA OPPTS 850.1500)	NOEC: 510 (based on number of spawns per female) NOEC: ≥ 950 (based on F0 and F1 survival, length, wet weight, VTG and sex ratio) NOEC: 950 (based on number of spawns per female)	EFSA, 2010b; INERIS, 2011

J.6. Propiconazole

This is a data summary for propiconazole (CAS: 60207-90-1; EC: 262-104-4). Propiconazole has the molecular formula $C_{15}H_{17}Cl_2N_3O_2$, and the structural formula is shown below.



The physicochemical properties of propiconazole are shown in Table 67, based on data from EFSA (2017).

Table 67 Physicochemical properties of propiconazole (CAS: 60207-90-1) based on data from EFSA (2017)

Property	Value
Molecular weight	342.2
Vapour pressure	5.6×10^{-5} Pa at 25 °C
Partition coefficient	Log K_{OW}: 3.72
Water solubility	150 mg/L at 20 °C

Aquatic toxicity data for propiconazole is available from an EU DAR (EFSA, 2017) and the SETAC Pellston Workshop® review (Matthiessen *et al*., 2017). The aquatic dataset in the DAR includes two fish short-term reproduction assays (Level 3), one amphibian metamorphosis assay (Level 3), one fish early life stage study (Level 4), two fish full life cycle studies (Level 5) and four aquatic invertebrate reproduction assays (Level 4) which are reliable and relevant for PNEC derivation (Table 68). Chronic algal toxicity was also reported. The lowest chronic effect level reported in the EU DAR was for a FLCTT study (Level 5) conducted with *Cyprinodon variegatus*, which determined a NOEC of 68 µg/L based on reproduction.

For the preparation of the third priority substance watch list under the Water Framework Directive, the JRC report different NOECs for propiconazole in freshwater (JRC, 2020). There were two NOECs for algae (16 µg/L for *Desmodesmus subspicatus* and 0.95 µg/L for an unspecified species), a NOEC of 60 µg/L for crustaceans and a NOEC of 5.8 µg/L for fish. The NOEC for fish is assumed to be the result for VTG inhibition from Skolness *et al*. (2013), however this endpoint is not population-relevant. The JRC derived a PNEC of 0.095 µg/L based on the algal NOEC of 0.95 µg/L citing Zhou *et al*. (2019) as the source. Zhou *et al*. (2019), is not however the primary source and therefore the results cannot be verified. Similarly, the result for *D. subspicatus* cannot be verified. The JRC PNEC will therefore not be used as the basis for the assessment.

The standard PNEC for propiconazole is therefore based on the NOEC of 68 µg/L from the fish life cycle test with *C. variegatus* with an AF of 10 applied (PNEC = 6.8 µg/L). Although this is with an estuarine species.

Table 68 Summary of reliable data relevant for chronic PNEC derivation for propiconazole (EFSA, 2017; Matthiessen *et al*., 2017; JRC, 2020)

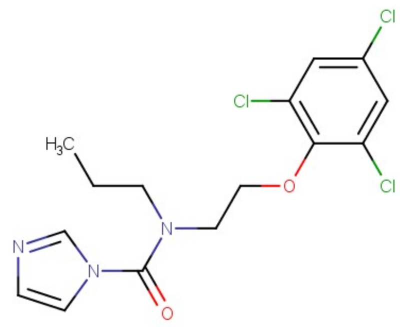
Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
Algae and aquatic plants	*Pseudokirchneriella subcapitata*	Growth inhibition (OECD TG 201)	NOEC: 460 (based on 72-hour growth rate) NOEC: 130 (based on 72-hour biomass and yield)	EFSA, 2017
	Desmodesmus subspicatus	Not reported	NOEC: 16	JRC, 2020
	Not reported	Not reported	NOEC: 0.95 (further details not available)	JRC, 2020; Zhou *et al*., 2019[1]
Aquatic invertebrates	*Daphnia magna*	Reproduction (OECD TG 211)	NOEC: 310 (based on reproduction and mortality)	EFSA, 2017; Matthiessen *et al*., 2017
	Daphnia magna	Reproduction (OECD TG 211)	NOEC: 370 (based on growth [adult length]) EC_{10}: 350 (based on reproduction)	EFSA, 2017; Matthiessen *et al*., 2017
	Daphnia magna	Reproduction (OECD TG 211)	NOEC: 60 (based on development)	JRC, 2020; Zhou *et al*., 2019[1]
	Chironomus riparius	28-day emergence and development (OECD TG 218/219)	NOEC: 4 000 (based on development rate)	EFSA, 2017
	Americamysis bahia	28-day reproduction	NOEC: 114 (based on reproduction)	EFSA, 2017

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
			NOEC: 205 (based on reproduction and mortality)	Matthiessen *et al.*, 2017
Amphibians	*Xenopus laevis*	AMA (OECD TG 231)	NOEC: 56 (based on snout-to-vent length and body weight) NOEC: ≥ 570 (based on hind-limb length, development and thyroid histology)	EFSA, 2017; Matthiessen *et al.*, 2017
Fish	*Pimephales promelas*	FSTRA (OECD TG 229)	NOEC: 120 (based on reproduction, female body weight and length, gonadosomatic index, female VTG inhibition, female histopathology and male nuptial tubercle score but also effects on female survival (general toxicity)) NOEC: ≥ 1 000 (based on survival, male body weight and length, male histopathology and male VTG induction)	EFSA, 2017; Matthiessen *et al.*, 2017
	Pimephales promelas	FSTRA (OECD TG 229)	NOEC: 5.8 (based on female VTG induction) NOEC: 53 (based on fecundity and gonadosomatic index)	Skolness *et al.*, 2013
	Pimephales promelas	FELS (OECD TG 210)	NOEC: 430 (based on survival)	EFSA, 2017

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
			EC_{10}: 380 (based on wet weight)	
	Pimephales promelas	FLCTT (US EPA OPPTS 850.1500)	NOEC: 7.8 (based on increased F1 male nuptial tubercle score) NOEC: 21 (based on F1 male interstitial cell hyperplasia and hypertrophy) NOEC: 188 (based on F0 and F1 reproduction and F1 female VTG inhibition) NOEC: ≥ 558 (based on F0 length and body weight; F1 hatching success, survival, growth, sex ratio and male VTG induction; and F2 hatching success and survival) EC_{10}: 119 (based on F0 and F1 egg production)	EFSA, 2017; Matthiessen *et al.*, 2017
	Cyprinodon variegatus	FLCTT (US EPA OPPTS 850.1500)	NOEC: 68 (based on F0 reproduction) NOEC: 150 (based on F1 hatching success) NOEC: 290 (based on F1 survival) NOEC: ≥ 550 (based on F0 hatching, survival and growth)	EFSA, 2017; Matthiessen *et al.*, 2017

J.7. Prochloraz

This is a data summary for prochloraz (CAS: 67747-09-5; EC: 266-994-5). Prochloraz has the molecular formula $C_{15}H_{16}Cl_3N_3O_2$, and the structural formula is shown below.



The physicochemical properties of prochloraz are shown in Table 69, based on data from EFSA (2011).

Table 69 Physicochemical properties of prochloraz (CAS: 67747-09-5) based on data from EFSA (2011)

Property	Value
Molecular weight	376.7
Vapour pressure	1.5 x 10^{-4} Pa at 25 °C
Partition coefficient	Log K_{OW}: 3.50 – 3.53 at 25°C; 4.3 – 4.4 at 20 – 25 °C
Water solubility	34.4 mg/L at 25 °C

Aquatic toxicity data for prochloraz is available from an EU DAR (EFSA, 2011). The aquatic dataset includes one fish early life stage study (Level 4), one fish full life cycle study (Level 5) and one aquatic invertebrate reproduction assay (Level 4) which are reliable and relevant for PNEC derivation (Table 70). The lowest chronic effect level was reported for algae, with a NOEC of 3.2 µg/L. Four additional fish short-term reproduction assays (OECD, 2006; Onishi *et al*., 2021), a FELS test (EFSA, 2011), five FSDT from OECD method validation (Holbech *et al*., 2012) and a MEOGRT (Flynn *et al*., 2017) are available in the literature. In the FSTRA (OECD CF Level 3) the lowest NOEC is from a study with *Oryzias latipes*, which was 20 µg/L based on spawning activity (OECD, 2006). The NOEC of 48.5 µg/L for reduced growth in the FELS test with *P. promelas* (EFSA, 2011), was lower than the NOECs of 96 – 106 µg/L with the same species, based on sex ratio in an FSDT (Holbech *et al*., 2012). In the MEOGRT the NOEC based on fecundity in the F2 generation was 17.5 µg/L. There was also a NOEC of 5.3 µg/L based on reduced growth of the F1 sub-adult, however the biological relevance of this was unclear because the change was small and was not apparent in the adult fish.

For the preparation of the third priority substance watch list under the Water Framework Directive, the JRC identified three PNECs for prochloraz in freshwater (JRC, 2020). The lowest PNEC was reported to be 0.161 µg/L, which was derived from acute data (Zhou *et al*., 2019). The other two PNECs were derived by EU Member States (France: 10 µg/L, Netherlands: 1.3 µg/L) but there is no publicly available information on the dataset used.

Table 70 Summary of reliable data relevant for chronic PNEC derivation for prochloraz (EFSA, 2011)

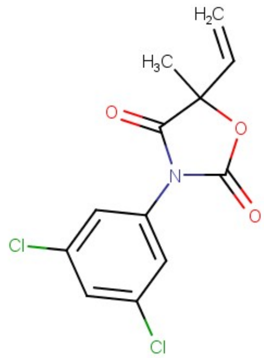
Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
Algae and aquatic plants	*Desmodesmus subspicatus*	OECD TG 201	NOEC: 3.2	EFSA DAR, 2007
Aquatic invertebrates	*Daphnia magna*	Reproduction (OECD TG 211)	NOEC: 22.2 (based on reproduction)	EFSA, 2011
Fish	*Oncorhynchus mykiss*	28-day assay	NOEC: 180 (based on behavioural abnormalities)	EFSA, 2011
	Pimephales promelas	FSTRA (OECD TG 229)	NOEC: 100 (based on female VTG)	OECD, 2006
	Oryzias latipes	FSTRA (OECD TG 229)	NOEC: 20 (based on spawning activity) NOEC: 100 (based on female VTG)	OECD, 2006
	Danio rerio	FSTRA (OECD TG 229)	NOEC: 100 (based on female VTG)	OECD, 2006
	Oryzias latipes	FSTRA (OECD TG 229)	NOEC: 20.6 (based on female VTG and fertility rate)	Onishi *et al*., 2021

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
			NOEC: ≥ 44.9 (based on mortality, length, weight, GSI, HSI, male VTG, male and female SSC, number of total eggs, number of fertilised eggs)	
	Pimephales promelas	FELS (OECD TG 210)	NOEC: 48.5 (further details not publicly available)	EFSA, 2011
	Pimephales promelas	FSDT (OECD TG 234)	NOEC: 96 – 106 (sex ratio based on 2 labs in FSDT validation) NOEC: < 29 – 31 (female VTG based on 2 labs in FSDT validation)	Holbech *et al.*, 2012
	Danio rerio	FSDT (OECD TG 234)	NOEC: 48 - < 60 (sex ratio based on 3 labs in FSDT validation) NOEC: (15 – 60 female VTG based on 3 labs in FSDT validation)	Holbech *et al.*, 2012
	Oryzias latipes	MEOGRT (OECD TG 240)	NOEC: 17.5 (fecundity) 5.3 (decreased growth in sub adults although biological relevance unclear because the effect was small and not observed in the adult, VTG in females and SSC in males)	Flynn *et al.*, 2017

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
	Pimephales promelas	FLCTT (US EPA OPPTS 850.1500)	NOEC: 24.9 (based on F1 egg production) NOEC: ≥ 191 (based on F1 and F2 sex ratio and male and female VTG)	EFSA, 2011

J.8. Vinclozolin

This is a data summary for vinclozolin (CAS: 50471-44-8; EC: 256-599-6). Vinclozolin has the molecular formula $C_{12}H_9Cl_2NO_3$, and the structural formula is shown below.



The physicochemical properties of vinclozolin are shown in Table 71, based on data from PubChem[1].

Table 71 Physicochemical properties of vinclozolin (CAS: 50471-44-8) based on data from PubChem[2]

Property	Value
Molecular weight	286.11
Vapour pressure	1.6×10^{-5} Pa at 20 °C
Partition coefficient	Log K_{OW}: 3.10
Water solubility	2.6 mg/L at 20 °C

A summary of aquatic ED data for vinclozolin is available from the SETAC Pellston Workshop® (Matthiessen *et al*., 2017). The aquatic dataset includes one fish full life cycle study (Level 5), three fish multigenerational assays (Level 5) and five aquatic invertebrate reproduction assay (Level 4) which are reliable and relevant for PNEC derivation (Table 72). Ten additional studies are available in the literature covering fish short-term reproduction assays (FSTRA) (Nakamura *et al*., 2014; Onishi *et al*., 2021; Martinović *et al*.,

[1] https://pubchem.ncbi.nlm.nih.gov/compound/Vinclozolin (accessed January 2023)

[2] https://pubchem.ncbi.nlm.nih.gov/compound/Vinclozolin (accessed January 2023)

2008; Makynen *et al*., 2000), androgenised female stickleback screening (AFSS) assays (Katsiadaki *et al*., 2006; Jolly *et al*., 2009; Pottinger *et al*., 2013; OECD, 2022), rapid androgen disruption activity reporter (RADAR) assays (OECD, 2022) and a medaka extended one-generation reproduction test (MEOGRT) (Flynn *et al*., 2017). Mechanistic data indicates anti-androgenic activity from IC_{50} = 4.25 µg/L in an AFSS, a NOEC of 33.3 µg/L in the RADAR, and a NOEC of 137 µg/L in a FSTRA. A regulatory assessment for vinclozolin is not available. Matthiessen *et al*. (2017) conclude that population-relevant effects of vinclozolin (LOEC) typically occur at 150 µg/L and above.

In the absence of risk assessment data the MEOGRT study (Level 5), which determined a NOEC of 17.3 µg/L based on F1 and F2 larvae survival (Flynn *et al*., 2017) as a conservative assessment although it should be noted that these effects were not observed in three Medaka Multigeneration Tests (US EPA, 2013). However, a reduced number of male secondary sex characteristics were consistently reported in the medaka multigeneration studies at test concentrations below 17.3 µg/L. A PNEC can be deterministically derived by applying an AF of 10, resulting in a PNEC of 1.7 µg/L.

Table 72 Summary of reliable data relevant for chronic PNEC derivation for vinclozolin (Matthiessen *et al*., 2017 and various scientific publications)

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
Algae and aquatic plants	*Pseudokirchneriella subcapitata*	120-day growth	NOEC: 1 020	PPDB Database[3]
	Lemna gibba	120-day growth	NOEC: 900	
Aquatic invertebrates	*Lymnaea stagnalis*	Snail reproduction	NOEC: ≥ 0.0922 (based on parental survival, growth and reproduction)	Giusti *et al*., 2014, cited in Matthiessen *et al*., 2017
	Lymnaea stagnalis	Mollusc partial life cycle (21 days)	NOEC: < 0.025 (based on mean cumulated fecundity of young adults) NOEC: 25 (based on cumulated fecundity of adult snails) NOEC: 250 (based on adult mortality) NOEC: ≥ 2 500 (based on adult fertility)	Ducrot *et al*., (2010) cited in Matthiessen *et al*., 2017 (Klimisch 3)
	Americamysis bahia	Multigeneration reproduction	NOEC: 120 (based on F0 sex ratio)	Matthiessen *et al*., 2017; US EPA test method validation

[3] Pesticide Properties DataBase (PPDB) accessed January 2023 http://sitem.herts.ac.uk/aeru/ppdb/en/Reports/680.htm

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
	Daphnia magna	Reproduction	NOEC: 790 (based on parental length, weight and reproduction)	Matthiessen *et al*., 2017
	Amphiascus tenuiremis	Reproduction	NOEC: 100 (based on survival, growth and reproduction)	Matthiessen *et al*., 2017, US EPA test method validation
Fish	*Oryzias latipes*	FSTRA (OECD TG 229)	NOEC: 189.3 (based on fertility and ↓ female VTG) NOEC: ≥ 606 (based on survival, growth, fecundity, male VTG, male and female SSC, gonadosomatic index [GSI], gonadal histology and hepatosomatic index [HSI])	Nakamura *et al*., 2014
	Oryzias latipes	FSTRA (OECD TG 229)	NOEC: 137 (based on fertility and ↓ female VTG) NOEC: ≥ 453 (based on survival, growth, fecundity, male VTG, male and female SSC)	Onishi *et al*., 2021
	Pimephales promelas	FSTRA (pre-TG)	NOEC: < 60 (based on fecundity and ↑ female growth) NOEC: 60 (based on ↓ male SSC, ↑ male GSI and ↑ female VTG)	Martinović *et al*., 2008 (Klimisch 3)

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
			NOEC: 255 (based on female gonadal histology [↑ atresia]) NOEC: ≥ 450 (based on survival, male growth, fertility, female SSC, male gonadal histology, female GSI and male VTG and testosterone)	
	Pimephales promelas	FSTRA (pre-TG)	NOEC: 178 (based on ↓ female GSI and delayed gonadal development and ↑ male E2) NOEC: > 710 (based on survival, male GSI and gonadal histology, male testosterone and female testosterone and E2)	Makynen *et al*., 2000
	Oryzias latipes	RADAR validation with 5 labs	NOEC: 33.3 (2 labs), 100 (3 labs)	OECD, 2022
	Gasterosteus aculeatus	AFSS	NOEC: < 25 (based on ↓ spiggin)	Katsiadaki *et al*., 2006
	Gasterosteus aculeatus	AFSS	NOEC: 25 (based on ↓ spiggin)	Jolly *et al*., 2009
	Gasterosteus aculeatus	AFSS	IC10: 4.25 (based on ↓ spiggin)	Pottinger *et al*., 2013
	Gasterosteus aculeatus	AFSS	NOEC: 25 (based on ↓ spiggin)	OECD, 2022

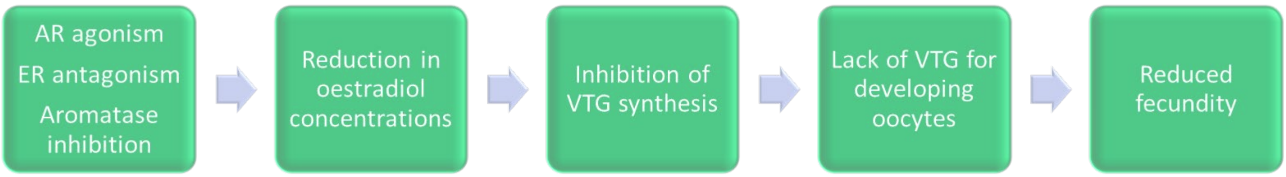
Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
	Pimephales promelas	FLCTT (EPA)	NOEC: 50 (based on F0 fecundity and delayed gonadal development and F1 survival, ↓ F1 length and F1 sex ratio) NOEC: ≥ 150 (based on F0 survival, male and female growth, fecundity and sex ratio)	Matthiessen *et al.*, 2017
	Oryzias talipes	MMT lifecycle validation	NOEC: ≥ 360 (based on F1 fecundity, fertility and female VTG)	Matthiessen *et al.*, 2017
	Oryzias latipes	MMT lifecycle validation	NOEC: < 9 (based on F2 male secondary sex characteristics [SSC]) NOEC: ≥ 332.5 (based on F1 fecundity and fertility)	Matthiessen *et al.*, 2017
	Oryzias latipes	MMT lifecycle validation	NOEC: < 5 (based on F2 male [phenotypic] SSC) NOEC: 13 (based on F1 male [phenotypic] SSC) NOEC: ≥ 197 (based on F0 and F1 fecundity and fertility)	US EPA, 2013; Matthiessen *et al.*, 2017
	Oryzias latipes	MEOGRT (OECD TG 241)	NOEC: < 17.3 (based on ↓ F1 male SSC) NOEC: 17.3 (based on F1 and F2 larvae survival, ↑ F0 female	Flynn *et al.*, 2017

Trophic level	Species	Study type	Chronic effect level (µg/L)	Reference
			SSC and ↓ F2 male SSC) NOEC: 33.2 (based on ↑ female SSC [F1]) NOEC: 69.5 (based on ↑ F1 and F2 female weight and ↑ F2 male VTG) NOEC: 136.3 (based on F0 and F1 fecundity and fertility) NOEC: ≥ 253.3 (based on F0 and F1 adult survival, male weight and length, male and female gonadal histology, hatching success, F1 sex ratio, F1 and F2 female VTG and F1 male VTG)	

Appendix K Adverse outcome pathways

Adverse Outcome Pathways (AOPs) provide a framework to describe the likely toxicological consequence of a biochemical interaction based on a known mode of action (MoA) (Ankley *et al.*, 2020). The AOP usually starts with a molecular initiating event (MiE) (for example interaction with a steroid receptor) and then progresses through a number of key events, usually at higher levels of biological organisation, culminating in an adverse outcome. The adverse outcome would usually be an effect on a population-relevant endpoint (e.g. development or reproduction). AOPs can be used to organise available data to allow the determination of a biologically plausible link between an endocrine MoA and an adverse effect in an intact organism, population or sub-population.

AOPs for effects in fish are available for some EAS MoAs, for example aromatase inhibition, oestrogen receptor (ER) antagonism, and androgen receptor (AR) agonism (Ankley *et al.*, 2020; Ankley *et al.*, 2023). These mechanisms of action have different MiEs but would be expected to lead to the same adverse outcomes (reduced fecundity in female fish and skewed male sex ratio). The AOPs for these ED mechanisms of action are given in Figure 12 and Figure 13 adapted from Ankley *et al.* (2020) and Ankley *et al.* (2023).



Note: VTG = vitellogenin, an egg yolk protein

Figure 12 Common elements of the AOP for AR agonism, ER antagonism and aromatase inhibition in reproductively active female fish adapted from Ankley *et al.* (2020)

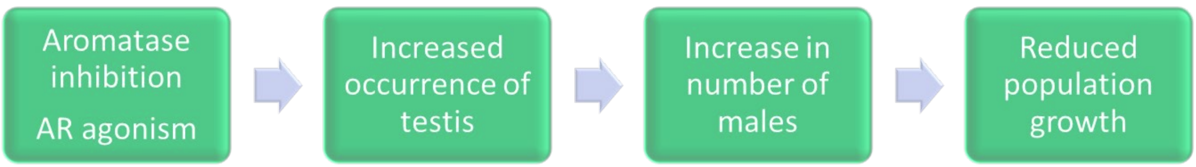
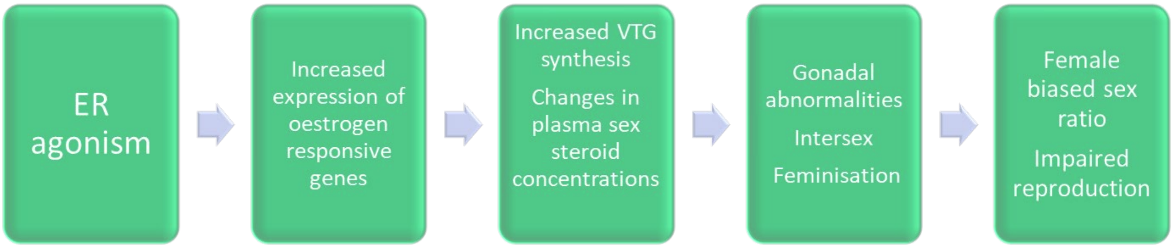


Figure 13 Common elements of the AOPs for AR agonism, and aromatase inhibition in juvenile fish adapted from Ankley *et al.* (2023)

These AOPs demonstrate that outcomes differ depending on sex and life stage.

The key events for AOPs relating to ER agonism in fish are less well defined (Figure 14) but there are clear effects on the male, based on induction of vitellogenin (VTG) and feminisation. ER agonism also affects outcomes in the female leading to reduced reproduction (Morshead *et al.*, 2023). As most studies co-expose males and females it can be unclear whether effects on reproduction are due to effects via the male or direct effects on the female (Morshead *et al.*, 2023).



Note: VTG = vitellogenin, an egg yolk protein

Figure 14 Generalised AOP for ER agonism in fish adapted from Ankley *et al*. (2010)

List of abbreviations

AC_{50}	Concentration at 50% of maximum activity
ADME	Adsorption, Distribution, Metabolism, and Elimination
AF	Assessment Factor
AfA	Applications for Authorisation
AFSS	Androgenised female stickleback screen
AICc	Akaike's Information Criterion corrected for sample size
AMA	Amphibian Metamorphosis Assay
AOP	Adverse Outcome Pathway
AR	Androgen Receptor
AUC	Area under the curve
3-BC	3-Benzylidene camphor
CRED	Criteria for Reporting and evaluating Ecotoxicity Data
DEHP	Bis(2-ethylhexyl) phthalate
DIO	Iodothyronine deiodinase
dpf	days post fertilization
op'DDT	Dichlorodiphenyltrichloroethane
EA	Environment Agency
EAS	Oestrogen, Androgen and Steroidogenesis
EC	European Commission
EC_{10}	The test concentration at which there is a 10% effect
ECHA	European Chemicals Agency
ED	Endocrine Disrupter or Endocrine Disrupting
EE2	17α-ethinyloestradiol

EFSA	European Food Safety Authority
ETNC	Environmental Threshold of No Concern
ER	Oestrogen receptor
EQS	Environmental Quality Standard
EU DAR	European Union Draft Assessment Report
FELS	Fish Early Life Stage
FLCTT	Fish Life Cycle Toxicity Test
FSDT	Fish Sexual Development Test
FSTRA	Fish Short-Term Reproduction Assay
GLP	Good Laboratory Practice
GSI	Gonadosomatic Index
HC_5	The hazardous concentrations affecting 95% of species
HLL	Hind Limb Length
HSI	Hepatosomatic Index
HSE	Health and Safety Executive
IUCLID	International Uniform ChemicaL Information Database
JRC	Joint Research Centre
LADGA	Larval Amphibian Growth and Development Assay
lnorm	Log normal
llogis	Log logistic distribution
LOEC	Lowest Observed Effect Concentration
MEOGRT	Medaka Extended One Generation Test
MIE	Molecular Initiating Event
MoA	Mechanism of Action
MMT	Medaka Multigeneration Test
NMDR	Non Monotonic Dose Response

NOEC	No Observed Effect Concentration
NP	Nonylphenol
NPEO	Nonylphenol ethoxylate
OECD	Organisation for Economic Co-operation and Development
OP	Octylphenol
PNEC	Predicted No Effect Concentration
PPAR	Peroxisome proliferator-activated receptor
PPDB	Pesticide Properties DataBase
RADAR	Rapid Androgen Disruption Activity Reporter
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals
RXR	Retinoid X receptor
SETAC	Society of Environmental Toxicology and Chemistry
SSD	Species Sensitivity Distribution
SSC	Secondary Sex Characteristics
SVHC	Substance of Very High Concern
SVL	Snout Vent Length
TBT	Tributyl tin
TPO	Thyroperoxidase
TR	Thyroid receptor
TRHR	Thyrotropin-releasing hormone receptor
TSHR	Thyroid-stimulating hormone receptor
UBA	German Federal Environment Agency
US EPA	United States Environmental Protection Agency
VTG	Vitellogenin

Would you like to find out more about us or your environment?

Then call us on

03708 506 506 (Monday to Friday, 8am to 6pm)

Email: enquiries@environment-agency.gov.uk

Or visit our website

www.gov.uk/environment-agency

incident hotline

0800 807060 **(24 hours)**

floodline

0345 988 1188 **(24 hours)**

Find out about call charges (https://www.gov.uk/call-charges)

Environment first

Are you viewing this onscreen? Please consider the environment and only print if absolutely necessary. If you are reading a paper copy, please don't forget to reuse and recycle.