Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

24-Aug-2012

English - Or. English

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

Cancels & replaces the same document of 26 July 2012

GUIDANCE DOCUMENT ON STANDARDISED TEST GUIDELINES FOR EVALUATING CHEMICALS FOR ENDOCRINE DISRUPTION

Series on Testing and Assessment

No. 150

JT03325479

ENV/JM/MONO(2012)22
Unclassified

English - Or. English

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 150

GUIDANCE DOCUMENT ON STANDARDISED TEST GUIDELINES FOR EVALUATING CHEMICALS FOR ENDOCRINE DISRUPTION

IOMC
INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS
A cooperative agreement among FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris 2012

Also published in the Series on Testing and Assessment:

No. 1, *Guidance Document for the Development of OECD Guidelines for Testing of Chemicals (1993; reformatted 1995, most recently revised 2009)*

No. 2, *Detailed Review Paper on Biodegradability Testing (1995)*

No. 3, *Guidance Document for Aquatic Effects Assessment (1995)*

No. 4, *Report of the OECD Workshop on Environmental Hazard/Risk Assessment (1995)*

No. 5, *Report of the SETAC/OECD Workshop on Avian Toxicity Testing (1996)*

No. 6, *Report of the Final Ring-test of the Daphnia magna Reproduction Test (1997)*

No. 7, *Guidance Document on Direct Phototransformation of Chemicals in Water (1997)*

No. 8, *Report of the OECD Workshop on Sharing Information about New Industrial Chemicals Assessment (1997)*

No. 9, *Guidance Document for the Conduct of Studies of Occupational Exposure to Pesticides during Agricultural Application (1997)*

No. 10, *Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data (1998)*

No. 11, *Detailed Review Paper on Aquatic Testing Methods for Pesticides and industrial Chemicals (1998)*

No. 12, *Detailed Review Document on Classification Systems for Germ Cell Mutagenicity in OECD Member Countries (1998)*

No. 13, *Detailed Review Document on Classification Systems for Sensitising Substances in OECD Member Countries 1998)*

No. 14, *Detailed Review Document on Classification Systems for Eye Irritation/Corrosion in OECD Member Countries (1998)*

No. 15, *Detailed Review Document on Classification Systems for Reproductive Toxicity in OECD Member Countries (1998)*

No. 16, *Detailed Review Document on Classification Systems for Skin Irritation/Corrosion in OECD Member Countries (1998)*

No. 17, *Environmental Exposure Assessment Strategies for Existing Industrial Chemicals in OECD Member Countries (1999)*

No. 18, *Report of the OECD Workshop on Improving the Use of Monitoring Data in the Exposure Assessment of Industrial Chemicals (2000)*

No. 19, *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (1999)*

No. 20, *Revised Draft Guidance Document for Neurotoxicity Testing (2004)*

No. 21, *Detailed Review Paper: Appraisal of Test Methods for Sex Hormone Disrupting Chemicals (2000)*

No. 22, *Guidance Document for the Performance of Out-door Monolith Lysimeter Studies (2000)*

No. 23, *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures (2000)*

No. 24, *Guidance Document on Acute Oral Toxicity Testing (2001)*

No. 25, *Detailed Review Document on Hazard Classification Systems for Specifics Target Organ Systemic Toxicity Repeated Exposure in OECD Member Countries (2001)*

No. 26, *Revised Analysis of Responses Received from Member Countries to the Questionnaire on Regulatory Acute Toxicity Data Needs (2001)*

No 27, *Guidance Document on the Use of the Harmonised System for the Classification of Chemicals which are Hazardous for the Aquatic Environment (2001)*

No 28, *Guidance Document for the Conduct of Skin Absorption Studies (2004)*

No 29, *Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (2001)*

No 30, *Detailed Review Document on Hazard Classification Systems for Mixtures (2001)*

No 31, *Detailed Review Paper on Non-Genotoxic Carcinogens Detection: The Performance of In-Vitro Cell Transformation Assays (2007)*

No. 32, *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies (2000)*

No. 33, *Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures (2001)*

No. 34, *Guidance Document on the Development, Validation and Regulatory Acceptance of New and Updated Internationally Acceptable Test Methods in Hazard Assessment (2005)*

No. 35, *Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies (2002)*

No. 36, *Report of the OECD/UNEP Workshop on the Use of Multimedia Models for Estimating Overall Environmental Persistence and Long Range Transport in the Context of PBTS/POPS Assessment (2002)*

No. 37, *Detailed Review Document on Classification Systems for Substances which Pose an Aspiration Hazard (2002)*

No. 38, *Detailed Background Review of the Uterotrophic Assay Summary of the Available Literature in Support of the Project of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to Standardise and Validate the Uterotrophic Assay (2003)*

No. 39, *Guidance Document on Acute Inhalation Toxicity Testing (2009)*

No. 40, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures which Cause Respiratory Tract Irritation and Corrosion (2003)*

No. 41, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures which in Contact with Water Release Toxic Gases (2003)*

No. 42, *Guidance Document on Reporting Summary Information on Environmental, Occupational and Consumer Exposure (2003)*

No. 43, *Guidance Document on Mammalian Reproductive Toxicity Testing and Assessment (2008)*

No. 44, *Description of Selected Key Generic Terms Used in Chemical Hazard/Risk Assessment (2003)*

No. 45, *Guidance Document on the Use of Multimedia Models for Estimating Overall Environmental Persistence and Long-range Transport (2004)*

No. 46, *Detailed Review Paper on Amphibian Metamorphosis Assay for the Detection of Thyroid Active Substances (2004)*

No. 47, *Detailed Review Paper on Fish Screening Assays for the Detection of Endocrine Active Substances (2004)*

No. 48, *New Chemical Assessment Comparisons and Implications for Work Sharing (2004)*

No. 49, *Report from the Expert Group on (Quantitative) Structure-Activity Relationships [(Q)SARs] on the Principles for the Validation of (Q)SARs (2004)*

No. 50, *Report of the OECD/IPCS Workshop on Toxicogenomics (2005)*

No. 51, *Approaches to Exposure Assessment in OECD Member Countries: Report from the Policy Dialogue on Exposure Assessment in June 2005 (2006)*

No. 52, *Comparison of Emission Estimation Methods Used in Pollutant Release and Transfer Registers (PRTRs) and Emission Scenario Documents (ESDs): Case Study of Pulp and Paper and Textile Sectors (2006)*

No. 53, *Guidance Document on Simulated Freshwater Lentic Field Tests (Outdoor Microcosms and Mesocosms) (2006)*

No. 54, *Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application (2006)*

No. 55, *Detailed Review Paper on Aquatic Arthropods in Life Cycle Toxicity Tests with an Emphasis on Developmental, Reproductive and Endocrine Disruptive Effects (2006)*

No. 56, *Guidance Document on the Breakdown of Organic Matter in Litter Bags (2006)*

No. 57, *Detailed Review Paper on Thyroid Hormone Disruption Assays (2006)*

No. 58, *Report on the Regulatory Uses and Applications in OECD Member Countries of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models in the Assessment of New and Existing Chemicals (2006)*

No. 59, *Report of the Validation of the Updated Test Guideline 407: Repeat Dose 28-Day Oral Toxicity Study in Laboratory Rats (2006)*

No. 60, *Report of the Initial Work towards the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine Active Substances (Phase 1A) (2006)*

No. 61, *Report of the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine Active Substances (Phase 1B) (2006)*

No. 62, *Final OECD Report of the Initial Work towards the Validation of the Rat Hershberger Assay: Phase-1, Androgenic Response to Testosterone Propionate, and Anti-Androgenic Effects of Flutamide (2006)*

No. 63, *Guidance Document on the Definition of Residue (2006)*

No. 64, *Guidance Document on Overview of Residue Chemistry Studies (2006)*

No. 65, *OECD Report of the Initial Work towards the Validation of the Rodent Uterotrophic Assay - Phase 1 (2006)*

No. 66, *OECD Report of the Validation of the Rodent Uterotrophic Bioassay: Phase 2. Testing of Potent and Weak Oestrogen Agonists by Multiple Laboratories (2006)*

No. 67, *Additional Data Supporting the Test Guideline on the Uterotrophic Bioassay in rodents (2007)*

No. 68, *Summary Report of the Uterotrophic Bioassay Peer Review Panel, including Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow up of this Report (2006)*

No. 69, *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models (2007)*

No. 70, *Report on the Preparation of GHS Implementation by the OECD Countries (2007)*

No. 71, *Guidance Document on the Uterotrophic Bioassay - Procedure to Test for Antioestrogenicity (2007)*

No. 72, *Guidance Document on Pesticide Residue Analytical Methods (2007)*

No. 73, *Report of the Validation of the Rat Hershberger Assay: Phase 3: Coded Testing of Androgen Agonists, Androgen*

Antagonists and Negative Reference Chemicals by Multiple Laboratories. Surgical Castrate Model Protocol (2007)

No. 74, *Detailed Review Paper for Avian Two-generation Toxicity Testing (2007)*

No. 75, *Guidance Document on the Honey Bee (Apis Mellifera L.) Brood test Under Semi-field Conditions (2007)*

No. 76, *Final Report of the Validation of the Amphibian Metamorphosis Assay for the Detection of Thyroid Active Substances: Phase 1 - Optimisation of the Test Protocol (2007)*

No. 77, *Final Report of the Validation of the Amphibian Metamorphosis Assay: Phase 2 - Multi-chemical Interlaboratory Study (2007)*

No. 78, *Final Report of the Validation of the 21-day Fish Screening Assay for the Detection of Endocrine Active Substances, Phase 2: Testing Negative Substances (2007)*

No. 79, *Validation Report of the Full Life-cycle Test with the Harpacticoid Copepods Nitocra Spinipes and Amphiascus Tenuiremis and the Calanoid Copepod Acartia Tonsa - Phase 1 (2007)*

No. 80, *Guidance on Grouping of Chemicals (2007)*

No. 81, *Summary Report of the Validation Peer Review for the Updated Test Guideline 407, and Agreement of the Working Group of National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2007)*

No. 82, *Guidance Document on Amphibian Thyroid Histology (2007)*

No. 83, *Summary Report of the Peer Review Panel on the Stably Transfected Transcriptional Activation Assay for Detecting Estrogenic Activity of Chemicals, and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2007)*

No. 84, *Report on the Workshop on the Application of the GHS Classification Criteria to HPV Chemicals, 5-6 July, Bern Switzerland (2007)*

No. 85, *Report of the Validation Peer Review for the Hershberger Bioassay, and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2007)*

No. 86, *Report of the OECD Validation of the Rodent Hershberger Bioassay: Phase 2: Testing of Androgen Agonists, Androgen Antagonists and a 5 α-Reductase Inhibitor in Dose Response Studies by Multiple Laboratories (2008)*

No. 87, *Report of the Ring Test and Statistical Analysis of Performance of the Guidance on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (Transformation/ Dissolution Protocol) (2008)*

No. 88, *Workshop on Integrated Approaches to Testing and Assessment (2008)*

No. 89, *Retrospective Performance Assessment of the Test Guideline 426 on Developmental Neurotoxicity (2008)*

No.90, *Background Review Document on the Rodent Hershberger Bioassay (2008)*

No. 91, *Report of the Validation of the Amphibian Metamorphosis Assay (Phase 3) (2008)*

No. 92, *Report of the Validation Peer Review for the Amphibian Metamorphosis Assay and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2008)*

No. 93, *Report of the Validation of an Enhancement of OECD TG 211: Daphnia Magna Reproduction Test (2008)*

No. 94, *Report of the Validation Peer Review for the 21-Day Fish Endocrine Screening Assay and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-Up of this Report (2008)*

No. 95, *Detailed Review Paper on Fish Life-Cycle Tests (2008)*

No. 96, *Guidance Document on Magnitude of Pesticide Residues in Processed Commodities (2008)*

No. 97, *Detailed Review Paper on the use of Metabolising Systems for In Vitro Testing of Endocrine Disruptors (2008)*

No. 98, *Considerations Regarding Applicability of the Guidance on Transformation/Dissolution of Metals Compounds in Aqueous Media (Transformation/Dissolution Protocol) (2008)*

No. 99, *Comparison between OECD Test Guidelines and ISO Standards in the Areas of Ecotoxicology and Health Effects (2008)*

No. 100, *Report of the Second Survey on Available Omics Tools (2009)*

No. 101, *Report of the Workshop on Structural Alerts for the OECD (Q)SAR Application Toolbox, 15-16 May 2008, Utrecht, the Netherlands (2009)*

No. 102, *Guidance Document for using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping of Chemicals (2009)*

No. 103, *Detailed Review Paper on Transgenic Rodent Mutation Assays (2009)*

No. 104, *Performance Assessment: Comparsion of 403 and CxT Protocols via Simulation and for Selected Real Data Sets (2009)*

No. 105, *Report on Biostatistical Performance Assessment of the draft TG 436: Acute Toxic Class Testing Method for Acute Inhalation Toxicity (2009)*

No. 106, *Guidance Document for Histologic Evaluation of Endocrine and Reproductive Test in Rodents (2009)*

No. 107, *Preservative Treated Wood to the Environment for Wood Held in Storage after Treatment and for Wooden Commodities that are not Covered and are not in Contact with Ground. (2009)*

No. 108, *Report of the Validation of the Hershberger Bioassay (weanling model) (2009)*

No. 109, *Literature Review on the 21-Day Fish Assay and the Fish Short-Term Reproduction Assay (2009)*

No. 110, *Report of the Validation Peer Review for the Weanling Hershberger Bioassay and Agreement of the Working Group of National Coordinators of the Test Guidelines Programme on the Follow-up of this Report (2009)*

No. 111, *Report of the Expert Consultation to Evaluate an Estrogen Receptor Binding Affinity Model for Hazard Identification (2009)*

No. 112, *The 2007 OECD List of High Production Volume Chemicals (2009)*

No. 113, *Report of the Focus Session on Current and Forthcoming Approaches for Chemical Safety and Animal Welfare (2010)*

No. 114, *Performance Assessment of Different Cytotoxic and Cytostatic Measures for the In Vitro Micronucleus Test (MNVIT): Summary of results in the collaborative trial (2010)*

No. 115, *Guidance Document on the Weanling Hershberger Bioassay in Rats: A Short-term Screening Assay for (Anti) Androgenic Properties (2009)*

No. 116, *Guidance Document 116 on the Conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451, 452 and 453 – 2nd Edition (2011)*

No. 117, *Guidance Document 117 on the Current Implementation of Internal Triggers in Test Guideline 443 for an Extended One Generation Reproductive Toxicity Study, in the United States and Canada (2011)*

No. 118, *Workshop Report on OECD Countries' Activities Regarding Testing, Assessment and Management of Endocrine Disrupters Part I and Part II (2010)*

No. 119, *Classification and Labelling of chemicals according to the UN Globally Harmonized System: Outcome of the Analysis of Classification of Selected Chemicals Listed in Annex III of the Rotterdam Convention (2010)*

No. 120, *Part 1: Report of the Expert Consultation on Scientific and Regulatory Evaluation of Organic Chemistry Mechanism-based Structural Alerts for the Identification of DNA Binding Chemicals (2010)*

No. 120, *Part 2: Report of the Expert Consultation on Scientific and Regulatory Evaluation of Organic Chemistry Mechanism-based Structural Alerts for the Identification of DNA Binding Chemicals (2010)*

No. 121, *Detailed Review Paper (DRP) on Molluscs Life-cycle Toxicity Testing (2010)*

No. 122, *Guidance Document on the Determination of the Toxicity of a Test Chemical to the Dung Beetle Aphodius Constans (2010)*

No. 123, *Guidance Document on the Diagnosis of Endocrine-related Histopathology in Fish Gonads (2010)*

No. 124, *Guidance for the Derivation of an Acute Reference Dose (2010)*

No. 125, *Guidance Document on Histopathology for Inhalation Toxicity Studies, Supporting TG 412 (Subacute Inhalation Toxicity: 28-Day) and TG 413 (Subchronic Inhalation Toxicity: 90-Day) (2010)*

No. 126, S*hort Guidance on the Threshold Approach for Acute Fish Toxicity (2010)*

No. 127, *Peer Review Report of the Validation of the 21-Day Androgenised Female Stickleback Screening Assay (2010)*

No. 128, *Validation Report of the 21-Day Androgenised Female Stickleback Screening Assay (2010)*

No. 129, *Guidance Document on Using Cytotoxicity Tests to Estimate Starting Doses for Acute Oral Systemic Toxicity Tests(2010)*

No. 131, *Report of the Test Method Validation of Avian Acute Oral Toxicity Test (OECD Test Guideline 223) (2010)*

No. 132, *Report of the Multi-Laboratory Validation of the H295R Steroidogenesis Assay to Identify Modulators (2010)*

No.133, *Peer Review Report for the H295R Cell-Based Assay for Steroidogenesis (2010)*

No.134, *Report of the Validation of a Soil Bioaccumulation Test with Terrestrial Oligochaetes by an International ring test (2010)*

No.135, *Detailed Review Paper on Environmental Endocrine Disruptor Screening: The use of Estrogen and Androgen Receptor Binding and Transactivation Assays in Fish (2010)*

No. 136, *Validation Report of the Chironomid Full Life-Cycle Toxicity Test (2010)*

No. 137, *Explanatory Background Document to the OECD Test Guideline on In Vitro Skin Irritation Testing (2010)*

No. 138, *Report of the Workshop on Using Mechanistic Information in Forming Chemical Categories (2011)*

No. 139, *Report of the Expert Consultation on Scientific and Regulatory Evaluation of Organic Chemistry Mechanism Based Structural Alerts for the Identification of Protein-binding Chemicals (2011)*

No. 140, *Report of the WHO/OECD/ILSI (Hesi) Workshop on Risk Assessment of Combined Exposures to Multiple Chemicals (2011)*

No. 141, *Report of the Phase 1 of the Validation of the Fish Sexual Development Test for the Detection of Endocrine Active Substances (2011)*

No. 142, *Report of the Phase 2 of the Validation of the Fish Sexual Development Test for the Detection of Endocrine Active Substances (2011)*

No. 143, *Peer Review Report for the Validation of the Fish Sexual Development Test and Agreement of the Working Group of National Co-ordinators of the Test Guideline Programme on the Follow-up of the Peer Review (2011)*

No. 144, *Validation Report for the Acute Chironomid Assay (2011)*

No. 145, *Transgenic Rodent Somatic and Germ Cell Gene Mutation Assay: Retrospective Performance Assessment (2011)*

No. 146, *Syrian Hamster Embryonic (SHE) Cell PH 6.7 Cell Transformation Assay Prevalidation Study Report (2012)*

No. 147, *Syrian Hamster Embryonic (SHE) Cell PH 7.0 Cell Transformation Assay Prevalidation Study Report (2012)*

No. 148, *Guidance Document on the Androgenised Female Stickleback Screen (2011)*

No. 149, *Validation Report of the Balb/c 3T3 Cell Transformation Assay (2012)*

No. 150, *Guidance Document on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption (2012)*

No. 152, *Case Study: Assessment of an Extended Chemical Category, the Short-chain Methacrylates, Targeted on Bioaccumulation (2011)*

No. 153, *Guidance Document for the Derivation of an Acute Reference Concentration (Arfc) (2011)*

No. 154, *Validation Report: Part 1 – Validation of Efficacy Methods for Antimicrobials used on Hard Surfaces (2011)*

No. 154, *Validation Report: Part 2 – Validation of Efficacy Methods for Antimicrobials used on Hard Surfaces (2011)*

No. 155, *Peer Review for the Validation of the Modified Skin Irritation Test Method using LabyCyte EPI-MODEL24; Additional Studies; and Agreement of the Working Group of National Coordinators on the Follow-up to the Peer Review* (2011)

No. 156, *Guidance Notes on Dermal Absorption* (2011)

No. 157, *Validation Report Phase 1 for the Zebrafish Embryo Toxicity Test (2011)*

No. 158, *Report of Progress on the Interlaboratory Validation of the OECD Harpacticoid Copepod Development and Reproduction Test (2011)*

No. 159, *Validation Report for the Skin Irritation Test Method using Labcyte Epi-Model24 (2011)*

No. 160, *Guidance Document on the Bovine Corneal Opacity and Permeability (BCOP) and Isolated Chicken Eye (ICE) Test Methods: Collection of Tissues for Histological Evaluation and Collection of Data on Non-Severe Irritants (2011)*

No. 161, *Peer Review Report for the Validation of the Stably Transfected Transcriptional Activation Assay for the Detection of Androgenic and Anti-Androgenic Activity of Chemicals (2011)*

No. 165, *Guidance Document on Crop Field Trials (2011)*

No. 166, *SIDS Initial Assessment Profiles agreed in the course of the OECD HPV Chemicals Programme from 1993 to 2013 (2012)*

No. 167, *Crosswalk of Harmonized U.S. - Canada Industrial Function and Consumer and Commercial Product Categories with EU Chemical Product and Article Categories (2012)*

No. 168, *The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins Binding to Proteins (2012)*

No. 169, *Use of the AOP to Develop Chemical Categories and Integrated Assessment and Testing Approaches (2012)*

No. 170, *Guidance Document for Demonstrating Efficacy of Pool and Spa Disinfectants and Field Testing*

ABOUT THE OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 34 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in ten different series: Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides and Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; and Safety of Manufactured Nanomaterials. More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (www.oecd.org/ehs/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organisations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. UNDP is an observer. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France

Fax: (33-1) 44 30 61 80

E-mail: ehscont@oecd.org

FOREWORD

This guidance document was developed as a follow-up to the workshop on OECD countries' activities regarding testing, assessment and management of endocrine disrupters, which was held in Copenhagen (Denmark) on 22-24 September 2010 (see document No. 118 published in the Series on Testing and Assessment).

In 2010, the OECD Secretariat presented the objectives and a draft outline of the document at the meeting of the Working Group of National Coordinators of the Test Guidelines Programme (WNT). The document was then developed by two consultants in close cooperation with an advisory group on testing and assessment of endocrine disrupters (EDTA AG). In November 2010, comments were requested from the WNT, the EDTA AG, the Task Force on Hazard Assessment and experts involved in the assessment of chemicals. The EDTA AG addressed the comments from the WNT at a meeting held in April 2011, and a progress report was presented to the WNT at its 2011 meeting. In May 2011, comments were requested from the WNT on the changes made to the draft Guidance Document. In parallel to the finalisation of the draft Guidance document, three case studies were then developed by the consultants to evaluate whether the conclusions and next steps recommended in the draft guidance document are sensible and helpful when assessed in light of comprehensive datasets. The draft guidance document and the three case studies were reviewed again and revised at a meeting of the EDTA AG in December 2011.

The draft guidance document was approved by the WNT at its meeting held in April 2012. It was declassified by the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology on 26 July 2012. This document is published under the responsibility of the Joint Meeting.

CONTRIBUTORS

The following people contributed to the writing of this Guidance Document: Sneha Satya (Australia), Birgit Mertens (Belgium), Mike Wade (Canada), Ulla Hass, Henrik Holbech, Magnus Løfstedt, Anne Marie Vinggaard, Henrik Tyle, Pia Juul Nielsen, Marie Louise Holmer and Sofie Christiansen (Denmark), Jukka Ahtiainen and Petteri Talasniemi (Finland), Jean-Marc Porcher, Rija Samsera and Claire Beausoleil (France), Anja Kehrer, Bernd Schaefer, Frauke Stock, Tanja Juffernholz, Marike Kolossa-Gehring, Karen-ildico Hirsch-Ernst, Vera Ritz, and Susanne Walter-Rohde (Germany), Zoltan Marcsek (Hungary), Albertio Mantovani and Cinzia La Rocca (Italy), Taisen Iguchi, Shigeki Miyachi, Kunihiko Yamazaki, Chisumi Eto, Akihiro Hirose, Atsushi Ono, Jun Kanno and Norihisa Tatarazako (Japan), Jong-Ryool Kim and Jisung Ryu (Korea), Betty Hakkert, ZhiChao Dang, Theo Traas, Theo Vermeire, Wim Mennes, Aldert Piersma, and Emiel Rorije (Netherlands), Boguslaw Baranski (Poland), Yvonne Adersson, Agneta Ohlsson and Ing-Marie Olsson (Sweden), José Maria Navas (Spain), Walter Lichtensteiger (Switzerland), Philippa Edwards, Mike Roberts, Ovnair Sepai, Susy Brescia, Ian Indans, Christina Lye, Daniel Merckel, and David Lovell (United Kingdom), Vicki Dellarco, Leslie Touart and Gary Timm (United States), Peter Korytar, Sharon Munn, François Busquet, Elise Grignard and Miriam Jacobs (European Union), Yasuyoshi Okuno, Remi Bars, Nina Hallmark, Masatoshi Kumamoto, Gernot Klotz, Rick Becker and James Wheeler (Business and Industry Advisory Committee), Gwynne Lyons (Environmental NGO), Kate Willett and Kristy Sullivan (International Council on Animal Protection in OECD Programmes), Jenny Odum and Peter Matthiessen (OECD consultants).

Table of Contents

A. Introduction and Background to the Guidance

A.1 The OECD initiated a high-priority activity in 1998 to revise existing and to develop new Test Guidelines for the screening and testing of endocrine disrupting chemicals. Since then a number of potential assays have been developed into Test Guidelines and others are in development. The screens and tests are contained within the "OECD Conceptual Framework for the Screening and Testing of Endocrine Disrupting Chemicals" (CF) which was modified and updated by the EDTA Advisory Group (AG) in 2011. The original and revised versions of the CF are shown in Annex 1. A workshop on "OECD Countries Activities Regarding Testing, Assessment and Management of Endocrine Disruptors" was held in Copenhagen on 22-24 September 2009 (OECD, 2010b). One output from this workshop was a recommendation that a Guidance Document (GD) on the assessment of chemicals for endocrine disruption should be developed by the EDTA AG. This was supported by the EDTA AG at its meeting on 17-18 May 2010. The objectives and scope of the GD were defined such that the document would be a tool to support regulatory authorities by helping to interpret assay results and suggesting possible additional studies for reducing uncertainty. The guidance should not prejudge or constrain what regulatory actions may be taken by a member country and should not suggest a testing strategy. The guidance should also support but not duplicate other GDs *e.g.* guidance on hazard assessment. It should be noted that the use of many of these tests for determination of toxicity due to endocrine disruption (hazard and risk assessment) for mammals and wildlife is rather new, and therefore the guidance given is considered to be subject to changes based on new evidence. The guidance is intended to be a "living" document that will be updated as the science in this area evolves.

A.2 In the context of this document, an endocrine disrupter (ED) has been defined according to WHO (2002), *i.e.*

"An ED 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."

WHO (2002) also defines the term "potential ED" such that "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."

It is acknowledged that many other definitions exist (*e.g.*Weybridge Conference, 1996) but the WHO (2002) definition has been used as a working definition for this document because it covers both human health and wildlife populations. This definition is widely used but not universally accepted.

For the purposes of this document, we have operationally defined the term "possible ED" to mean a chemical that is able to alter the functioning of the endocrine system but for which information about possible adverse consequences of that alteration in an intact organism is uncertain.

A.1 Objectives

The objectives of the Guidance Document are:

A.3 To support regulatory authorities' decisions on the hazard of specific chemicals and toxicologically-relevant metabolites when they receive test results from a Test Guideline (TG) or draft TG for the screening/testing of chemicals for endocrine disrupting properties. The contexts for these decisions will vary, depending on local legislation and practice, so the advice is worded in such a way as to permit flexible interpretation.

A.4 To provide guidance on how to interpret the outcome of individual tests and how to increase evidence on whether or not a substance may be an ED. Testing strategies or guidance on interpretation from a suite of tests are not given

A.5 Hazard assessment methods in this document are arranged in a two step process, with the intention of minimising animal testing globally through application of the 3Rs (Replace, Reduce and Refine the use of laboratory animals in testing):

- Use of a harmonised framework for assessing test results together with existing information on likely or known hazards should avoid unnecessary animal testing.
- Recommendation of a test method that may be performed if regulatory authorities need more evidence. The test method is defined precisely to facilitate the Mutual Acceptance of Data and to avoid unnecessary duplication of testing. The recommended test method will utilise non-animal tests where possible although a few alternative scenarios are considered depending upon existing information.

A.2 General Approach

A.6 The general approach taken by this GD is primarily to consider the possible results that might be obtained from each ED-responsive assay[1], and to provide guidance about how these results might be interpreted in the light of data that may or may not already be available from other *in vitro* or *in vivo* assays. The nature, quantity and quality of the existing and new data in each of these scenarios should be evaluated in a weight-of-evidence approach (for example see USEPA, 2011; Borgert et al, 2011), and there is generally no single 'right' answer. Use of other technologies (for example gene extression analysis or"omics" data) may help in understanding the link between endocrine-related mechanisms and apical effects in a weight-of-evidence approach.This GD should therefore be used flexibly in the light of local regulatory needs. The key questions addressed concern likely mechanisms of endocrine action and any resulting apical effects that can be attributed to such action. Given the widely agreed definition of endocrine disrupting chemicals (WHO, 2002), the advice only suggests that a chemical is an ED if an adverse *in vivo* effect can be plausibly linked to an endocrine mode of action.

A.7 Secondly, this document provides advice on the next step in testing (if any) which might be appropriate for a regulatory authority to take, given the various data scenarios. It should be noted that it has only been possible to cover the most likely scenarios. Advice on further testing which may be needed to assist in deciding if a chemical is an ED is generally limited to a single next step, and this GD therefore does not present an entire hazard testing strategy for possible EDs.

A.8 The key advice for each assay is given in the form of a table which lists a series of scenarios for combinations of different assay results and varying backgrounds of existing data, and provides advice on interpretation and further testing which may be considered in each scenario. However, each table should be read in conjunction with the preceding text that explains issues related to the assay and for which there is insufficient space in the tabular format. Once again, it is important to note that these tables (so-called 'building blocks') are purely advisory, so individual regulatory authorities are not in any way bound to follow the advice. This is all the more important given that the guidelines for testing for endocrine disruption are relatively new and the field will probably develop further.

[1] ED-responsive assays are those *in vitro* or *in vivo* assays whose endpoints are known to respond positively to EDs and/or possible EDs.

A.3 Scope and Limitations

A.3.1 Assays and Endocrine Modalities Covered

A.9 The scope of the main section of the GD is limited to providing guidance on how to interpret results from assays included in the OECD Conceptual Framework (CF) for testing and assessment of EDs (see Annex 1). As the field of endocrine disruption is still developing, the CF will be subject to periodic revisions. In fact, during the writing of the GD, the CF has been revised. The assays discussed are those included in the **original** CF plus some additional assays that were considered relevant to assessment of EDs. The CF as revised in 2011 now includes all of these assays but some other assays were also added to the CF that are not included in this GD. Guidance is provided on the endpoints for the assays discussed, with respect to the endocrine modalities listed below. This is followed by guidance on how to increase evidence that a chemical is/is not an ED based on the result from the assay under consideration and other existing relevant information. Various scenarios are considered and the guidance suggests different considerations and the next test that may be performed in a single step.

A.10 Detailed guidance is given for the most relevant assays in the **original** CF from the perspective of ED identification, while for the other assays, a more limited guidance is provided (Annex 2). The GD is limited to endocrine mechanisms and hazard assessment. Information on chemical exposure (*e.g.* on use, volume, fate, levels, duration and route) is not considered.

A.11 The GD only covers the same endocrine modalities as the CF, *i.e.*:

- Estrogen receptor mediated
- Androgen receptor mediated
- Thyroid hormone mediated
- Steroidogenesis interference

A.12 Although the assays in this guidance are applicable to most types of EDs which are currently known (*i.e.* those operating via estrogen/ androgen/ thyroid/ steroidogenesis – EATS - modalities), it should be recognised that the assays may not be responsive to certain poorly-understood chemical types or modes of action. For example, it is unlikely that EDs that damage the corticosteroid system of wildlife species will be covered (Trenzado *et al*, 2003) although the adrenals are examined in many mammalian assays, therefore providing an alert. Some EDs may have epigenetic effects (although such effects are not confined to EDs). Such potential effects have been reviewed and discussed *inter alia* by Anway and Skinner (2006) and Crews and McLachlan (2006). In essence, an epigenetic effect is a change in phenotype or gene expression, inherited over rounds of cell division and sometimes transgenerationally, caused by mechanisms other than alterations in gene sequence (*e.g.* histone modifications, DNA methylation, RNAi mediated gene silencing). It has been suggested that epigenetic changes may result in transgenerational phenotypic effects and it is currently unclear whether the long-term assays available for testing possible EDs (*e.g.* fish, avian and rodent lifecycle tests) would reveal the full range of potential epigenetic responses. For example, Brown *et al*. (2009) failed to observe heritable reproductive defects in the offspring of male rainbow trout exposed to a strong estrogen. The field of epigenetics is currently being reviewed, and has been published as a draft OECD Detailed Review Paper on the "State of the Science on Novel In Vitro and In Vivo Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors" (OECD 2011a)

A.3.2 Scope of Assessment and Restriction to Single Assays

A.13 The GD does not present a testing strategy as it is restricted to a single step when further testing is recommended or proposed for consideration. It only recommends the most appropriate assay that

could be performed if countries need more evidence to support a management decision. The proposed guidance is not meant to encourage animal testing. It encourages the maximal use of all existing information consistent with OECD's Integrated Approaches to Testing and Assessment (OECD, 2008).

A.14 The level of confidence about whether or not a compound impacts endocrine function will increase with combined lines of pertinent evidence from multiple studies and endpoints across taxa, and which encompass different life stage effects and a range of doses. The amount of evidence needed to decide whether a substance is an ED in a regulatory context will depend on different authorities' policies/frameworks and the regulatory decision context. For example, results from a particular test or building block may suffice when making a decision for priority setting but may not be adequate for more predictive hazard or risk assessment.

Guidance is not given on the conduct of weight of evidence evaluations, risk assessments or the relevance for human health of results from the assays considered. Some guidance for this is provided in OECD (2008, 2010b) and WHO (2007). It is acknowledged that some mechanisms of action in rodents may not be relevant for humans *e.g* increased TSH and thyroid hyperplasia leading to the induction of thyroid tumours in rats, but the human relevance of specific mechanisms are not discussed.

Furthermore the guidance does not consider exposure, however this should be included when deciding whether further testing is needed in order to avoid unnecessary animal tests. This may be particularly relevant to wildlife where the environmental risk assessment should compare the sensitivity of all species and further testing should be limited to the concerned group of organisms driving the risk assessment. Lastly, as in any evaluation, it is essential that the degree of confidence and uncertainty be communicated in the characterization of the conclusions.

A.3.3 Rationale for Assay Inclusion

A.15 Detailed guidance is provided in the main part of this document on the validated and/or widely-accepted assays[2] in the **original** CF, these are listed in Part A of Table A.1. The terms 'validation' and 'validated assays' are used as defined in the OECD GD on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment, No. 34 (OECD, 2005) (see also Glossary). Validation may have been conducted by OECD or other organisations (*e.g.* ICCVAM). Note that the word 'assay' is used here to be consistent with the terminology used in the CF and describes a "test method" as defined in OECD (2005) *i.e.* "a test method is an experimental system that can be used to obtain a range of information from chemical properties through the adverse effects of a substance. The term 'test method' may be used interchangeably with 'assay' for ecotoxicity as well as for human health studies". The word 'screen' is used in this document to describe *in vitro* or *in vivo* assays which provide information on an endocrine disruption mechanism, but not generally information on adverse effects, for use in hazard or risk assessment. However, some regulatory authorities may wish to use positive screening tests for preliminary risk assessments. On the other hand, the word 'test' covers *in vivo* assays which can provide evidence to support a conclusion that a chemical is an ED that can cause adverse effects in an intact organism. An example of a screen would be the estrogen binding assay which only measures receptor binding activity *in vitro*, whereas an example of a test would be the medaka multi-generation test which measures reproductive success in intact fish. 'Screen' and 'test' are also broadly defined in OECD (2005) but here the word 'test' is used more precisely, see the glossary for all terms.

A.16 Assays providing information on potential interaction with endocrine systems, but which have not yet completed validation or are test guidelines that are not primarily designed for testing

[2] These are assays which have been validated at the national or international level, especially as OECD TGs.

specifically for endocrine disruption, are listed in Part B of Table A.1. All of these assays are now included in the **revised** CF. Limited guidance for them is given in Annex 2. These assays (*e.g.* OECD TG 408 Repeated Dose 90-Day Oral Toxicity Study in Rodents and OECD TG 451-3 Combined Chronic Toxicity/Carcinogenicity Studies) contain relevant endocrine endpoints (*e.g.* weights and histopathology of sex organs), and are used as such for REACH (OECD TG 408) and pesticide dossier evaluation, for example. OECD TG 453 (Combined Chronic Toxicity/Carcinogenicity Studies) provides information on carcinogenicity in endocrine tissues and is therefore very important for endocrine assessment of chemicals. OECD TG 421 (Reproduction/Developmental Toxicity Screening Test) and OECD TG 422 (Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test) provide information on reproduction in addition to effects on endocrine organs and are also used for REACH, but as they are not validated for endocrine outcomes, they are included in Annex 2.

A.3.4 Rationale for Assay Exclusion

A.17 Assays mentioned in the **original or revised** CF but not covered in this document are listed in Part C of Table A.1. Guidance for these assays has generally been omitted either because there is insufficient experience in their use (*e.g.* invertebrate lifecycle assays and *in vitro* assays for determining disruption of thyroid function), or because they are thought not to offer significant advantages over existing tests (*e.g.* fish hepatocyte vitellogenin assay).

A.18 *In vitro* screening assays for disruption of thyroid function have not been validated and the Detailed Review Paper on Thyroid Hormone Disruption Assays (OECD, 2006a) concluded: "The complicated nature of the thyroid system, makes development of an *in vitro* battery of assays to detect thyroid disruption unlikely in the near future. The conclusion is based on two facts: the *in vitro* assays available need further development before they can be validated, and the number of *in vitro* assays required to encompass every potential point of disruption in the thyroid system would be too great for a manageable assay battery. Furthermore, *in vitro* assays alone would not detect interactions within the thyroid system in response to toxicants. However, recommendations were made on *in vitro* assays that could be developed and utilised for high throughput screens in the near future". No guidance has therefore been written at present. There is, however, use of these assays in research and therefore data may be available and could be considered as "existing data" when evaluating the results of the assays considered in this guidance. The OECD Validation Management Group for non-animals tests is currently discussing the availability of thyroid assays..

A.19 The Yeast Estrogen and Yeast Androgen screens have also not been included in the guidance, although they are commonly used as *in vitro* screens in ecotoxicology (Routledge and Sumpter, 1996; Sohoni and Sumpter, 1998). They suffer from limitations such as problems with materials that have fungicidal activity or inhibit cell proliferation, solubility, permeability or transport issues across the cell wall (ICCVAM, 2003). It has also been reported that the YES assay is not sensitive for anti-estrogenic chemicals (Fang *et al*., 2000) The Detailed Review Paper on "Environmental endocrine disruptor screening: The use of estrogen and androgen receptor binding and transactivation assays in fish" (OECD, 2010d) describes these assays with the following recommendation:

"The YES/YAS-assays are recommended for further detailed evaluations primarily focusing on assays with fish steroid hormone receptors. It should be further evaluated whether such tests can provide meaningful information with special emphasis towards fish. Detailed comparisons on advantages/disadvantages to other *in vitro* assays, such as regarding yeast cell membrane permeability to certain compounds or chemical classes or other potential limitations, need to be clearly demonstrated before any further test method development or validation is performed. The sensitivity and specificity of any proposed YES/YAS assay needs to be demonstrated prior any further developments towards a Test Guideline."

The YES and YAS assays could be considered to be the forerunners of the ER and AR STTA assays and many of the possible next steps to be taken would be the same. These "building blocks" could therefore be used cautiously to provide guidance for the YES and YAS assays, but noting the limitations described above. The guidance for the ER STTA (OECD TG 455) would cover the YES assay and is given in Section C.2.1. The guidance for the AR STTA would cover the YAS assay and is given in Section Annex 2.1.

A.20 Guidance about tests that are based on the induction of proliferation, *e.g.* The E-screen where proliferation in estrogen-responding cells, particularly in the MCF-7 human breast cancer cell line, is used to detect estrogenic activity (Soto and Sonnenschein, 2001) is also not included. Proliferation assays are not recommended by ICCVAM (2003) because cell proliferation can be mediated through pathways other than those involving transcriptional activation of estrogen responsive genes. However, it should be noted that ICCVAM will complete a review of MCF-7 validation studies in 2011, so additions to the GD on this subject may be made in the future.

Table A.1. Screens and tests for which guidance is provided in this document.

Those listed under (A) are established assays which have been in wide use as validated OECD or national test guidelines. Guidance for these assays can be found within the body of the main GD. Those assays listed under (B) have not yet received full validation for endocrine outcomes, or are test guidelines that are not primarily designed for testing specifically for endocrine disruption, and guidance for these has been placed in Annex 2. Assays listed under (C) are those listed in the CF (as revised in 2011) but for which no guidance is provided. All assays have been sorted according to which level they occupy in the CF. Existing OECD test guidelines are indicated by the prefix "OECD TG".

It is important to bear in mind that the CF (see Annex 1.4) **is not a testing strategy** to be followed linearly from Level 1 through to Level 5, although in cases where little or no information is available (*i.e*. for new chemicals) it could provide guidance about where to start testing.

Conceptual Framework level	***In vitro* screens**	**Mammalian *in vivo* screens and tests**	**Wildlife *in vivo* screens and tests**
A. Validated assays for which guidance is provided in the main Guidance Document			
2	• ER Binding Assay (US EPA OPPTS 890.1250) • AR Binding Assay (US EPA OPPTS 890.1150) • OECD TG 455: Stably Transfected Human ERα Transcriptional Activation Assay (ER STTA) (including guidance for the antagonism assay – not part of OECD TG) • OECD TG 456: H295R Steroidogenesis Assay • Aromatase Assay (US EPA OPPTS 890.1200)	Nil	Nil
3	Nil	• OECD TG 440: Uterotrophic Bioassay in Rodents (UT Assay) (including OECD GD on the use of the assay to screen for anti-estrogenicity)	• OECD TG 231: Amphibian Metamorphosis Assay (AMA) • OECD TG 229: Fish Short Term

Conceptual Framework level	*In vitro* screens	Mammalian *in vivo* screens and tests	Wildlife *in vivo* screens and tests
		• OECD TG 441: Hershberger Bioassay in Rats (H Assay)	Reproduction Assay (FSTRA) • OECD TG 230: 21-Day Fish Assay • Androgenised Female Stickleback Screen (AFSS) (OECD GD 140)
4	Nil	• Pubertal Development and Thyroid Function Assay in Peripubertal Male Rats (PP male Assay) (US EPA OPPTS 890.1500) • Pubertal Development and Thyroid Function Assay in Peripubertal female Rats (PP female assay) (US EPA OPPTS 890.1450) • OECD TG 407: Repeated Dose 28-Day Oral Toxicity Study in Rodents • OECD TG 415: One-Generation Reproduction Toxicity Study (Guidance for this has been included with that for OECD TG 416)	• Fish Sexual Development Test (FSDT) (OECD TG 234) • OECD TG 206: Avian Reproduction Test
5	Nil	• OECD TG 416: Two-Generation Reproduction Toxicity Study (most recent update [adopted in 2001]) • OECD TG 443: Extended One-Generation Reproductive Toxicity Study	• Fish Lifecycle Toxicity Test (FLCTT) (US EPA OPPTS 850.1500)
B. Assays that have not yet completed validation, or not primarily designed for detection of endocrine disruption, for which limited guidance is given in Annex 2			
2	• Stably Transfected Human AR Transactivation Assay (AR	Nil	Nil

Conceptual Framework level	*In vitro* screens	Mammalian *in vivo* screens and tests	Wildlife *in vivo* screens and tests
	STTA)		
3	Nil	Nil	Nil
4	Nil	• Adult Male Assay • OECD TG 408: Repeated Dose 90-Day Oral Toxicity Study • OECD TG 451-3: Combined Chronic Toxicity/Carcinogenicity Studies • OECD TG 421 and 422: Combined 28-Day Reproductive Screening Tests •	• Larval Amphibian Growth and Development Assay (LAGDA) (draft OECD TG)
5	Nil		• Medaka Multi-Generation Test (MMGT) (draft OECD TG) • Avian Two-Generation Test (ATGT) (draft OECD TG)
C. Assays corresponding to those in the CF (original or revised) for which no guidance has been written at present			
2	• *TR binding affinity* • *AhR binding affinity* • *High-throughput pre-screens (not defined in CF)* • *Thyroid function in vitro* • *Fish hepatocyte VTG assay* • *Yeast transactivation assays (YES and YAS)* • *Proliferation-based screens*	Nil	Nil
3	Nil	• Non-receptor mediated hormone function	• *Xenopus* embryo thyroid signalling assay

Conceptual Framework level	***In vitro* screens**	**Mammalian *in vivo* screens and tests**	**Wildlife *in vivo* screens and tests**
		(not defined in CF)	(when/if TG is available)
4	Nil	• Prenatal Developmental Toxicity Study (OECD TG 414) • Developmental Neurotoxicity (OECD TG 426)	• Fish reproduction partial lifecycle test (when/if TG is available) • Mollusc partial lifecycle assays (when/if TG available) • Chironomid toxicity test (OECD TG 218-219) • *Daphnia* Reproduction Test (with male induction) (OECD TG 211) • Earthworm Reproduction Test (OECD TG 222) • Enchytraeid Reproduction Test (OECD TG 220) • Sediment Water *Lumbriculus* Toxicity Test Using Spiked Sediment (OECD TG 225) • Predatory mite reproduction test in soil (OECD TG 226) • Collembolan Reproduction Test in Soil (OECD TG 232)
5	Nil	Nil	• Mysid Life Cycle Toxicity Test (when TG is available) • Copepod Reproduction and Development Test (when TG is available) • Sediment Water Chironomid Life Cycle Toxicity Test (OECD TG 233) • Mollusc Full Lifecycle Assays (when TG is

Conceptual Framework level	***In vitro* screens**	**Mammalian *in vivo* screens and tests**	**Wildlife *in vivo* screens and tests**
			available) • *Daphnia* Multigeneration Assay (if TG is available)

B. General Guidance on Data Selection for Endocrine Assessment and Assays to be Included

B.1 The purpose of this section is to provide background information on the relevance of various types of data for supporting decisions about the endocrine disrupting properties of chemicals and other test materials (*e.g.* effluents; natural waters; contaminated foods *etc*.) in humans and vertebrate wildlife. Interpretation of results from invertebrate test guidelines is not included due to the rather poor current understanding of endocrinology in most invertebrates, and the lack of diagnostic screening endpoints with these taxonomic groups (*e.g.* OECD, 2010c). Nevertheless, non-OECD test assays, including those utilizing invertebrate species, may provide information that can be used in decision making. Furthermore, the document only deals with oestrogen-, androgen-, and thyroid-mediated endocrine disruption, and with interference with steroidogenesis. It does not cover other possible types of endocrine disruption, such as effects on the hypothalamus-pituitary-adrenal axis, on Ah receptor pathways or on the endocrine control of neural development. **The section is organised according to the OECD CF (Annex 1), as updated in 2011 with tests which were unavailable when it was first proposed (Annex 1.4)**.

B.2 It is important to bear in mind that the CF **is not a testing strategy** to be followed linearly from Level 1 through to Level 5, although in cases where little or no information is available (*i.e.* for new chemicals) it could provide ideas about where to start testing. In principle, any test can be conducted at any time in the hazard assessment process, depending on the perceived need for information. However, the data generated at various levels have a range of differing applications and implications, and must be interpreted accordingly. The purpose of this GD is therefore to assist assessors of endocrine-relevant tests with data interpretation in the light of information that may already exist, and to provide **optional** suggestions for obtaining additional data, if required, to increase confidence in conclusions on the endocrine disrupting possibilities of a particular chemical. It is clear that decisions about whether to obtain further data will be largely driven by regulatory needs which vary between jurisdictions, so advice on 'next steps which could be taken to increase evidence' is in no sense mandatory. As stated earlier, this process of data interpretation and assessment involves the need for a **weight of evidence approach** that considers both mechanistic and apical information, and it is self-evident that the more data which support a particular conclusion, the more reliable that conclusion will be.

B.3 This guidance supplements other GDs available on identification and interpretation of changes indicative of endocrine disruption such as the *GD on Mammalian Reproductive Toxicity Testing and Assessment* (OECD 2008c), the *GD for Histologic Evaluation of Endocrine and Reproductive Tests in Rodents* (OECD, 2009a) the *GD on the Diagnosis of Endocrine-related Histopathology in Fish Gonads* (OECD, 2010a) and the *Draft GD in support of The Test Guideline on the Extended One Generation Reproductive Toxicity Study* OECD (2010f).

B.4 Subsequent sections of this document will deal separately and in detail with *in vitro* mechanistic screens and *in vivo* screens and tests covering endpoints relevant for humans or vertebrate wildlife. In the context of vertebrate wildlife screens and tests, the test species are fish, amphibians and birds. General issues concerning such screens/tests are briefly considered together in this section. The distinction between screening assays used only for possible hazard detection and tests that may be used for both more comprehensive hazard detection and risk assessment is also discussed. The ability of the different assays at the different levels of the CF to detect EDs is discussed briefly here and in more detail in Section C.

B.5 It should be remembered that due to the molecular similarities of endocrine systems and receptor homologies across the vertebrates, there may be some potential for using information from non-mammalian vertebrate test assays for assessing endocrine activity in mammals (and *vice versa*), and especially for extrapolation between various *in vitro* screens. This must be tempered with the

knowledge that outcomes associated with a given endocrine modality can vary significantly across the vertebrates. The *in vitro* screens in question (although at present based largely on mammalian receptors and/or enzymes) are generally capable of providing information applicable to both humans and vertebrate wildlife (OECD, 2010d). Such extrapolation of *in vitro* information is generally qualitative (*e.g.* "Does the chemical bind to the estrogen receptor?") rather than quantitative (*e.g.* "What is the potency of the chemical in a particular taxonomic group?").

B.6 On the other hand, the purposes of the two in vivo assay types (mammalian and wildlife) are rather different. Whereas mammalian assays contribute mainly to risk assessments whose objective is to protect individual human beings, non-mammalian assays were originally intended to provide information to help predict possible impacts on wildlife populations. This in turn may affect the way in which assay data are interpreted. Nevertheless such assays may provide useful information for risk assessment across vertebrate species, including humans, because the fundamental approaches to such assessments are similar.

B.1 Considerations on the Assays Addressed

B.7 The considerations set out below are based partly on ideas proposed in Table 2 of OECD document ENV/JM/MONO(2010)2 *Conclusions based on a Nord-Utte project related to the OECD Conceptual Framework*. However, they have been augmented with information relevant for wildlife testing, and have also been amended in the light of recent scientific developments.

B.1.1 Conceptual Framework Level 1: Existing Data and Non-Test Information

B.8 It is important to emphasise that before conducting any assessment of data from an endocrine disruption screen or test, all existing information on the test chemical should be collated. Such data should ideally include physico-chemical properties, and fate and behaviour, as well as any toxicological and ecotoxicological information. However, it is recognised that all these types of information may not be available. It should also be noted that in some circumstances, regulatory decisions may be made on the basis of Level 1 data alone, without the need to proceed to any form of additional testing or screening and therefore the approaches listed below are essential for data gathering.

B.9 Data on structural analogues and from (Q)SAR models should be considered, especially if data on the chemical under consideration are scarce. At the present stage of (Q)SAR development in OECD, (Q)SAR models predicting mechanism would be used for prioritisation, ranking and hazard identification.

B.10 More advanced models, *e.g.* Mode of Action (WHO, 2007) or Adverse Outcome Pathway models (Schultz, 2010; Ankley, 2010), are in development. Some (Q)SAR models for endocrine disruption activity and reproductive toxicity effects are now becoming available (*e.g.* OECD, 2009b). The output of these models can be applicable (with caution) to interpretation of the mechanisms underlying *in vivo* results with vertebrates. Furthermore, other (Q)SAR methodologies such as categorization in the OECD (Q)SAR Toolbox can be used to identify groups of chemicals and structural alerts that are linked to *in vivo* effects, thereby elucidating possible key modes of action or mechanisms. Finally, (Q)SAR models that can predict metabolic transformation may be used in the interpretation of, *e.g*. disagreement between *in vitro* and *in vivo* results.

B.11 All existing relevant data should be maximally used (*e.g.*, structural and physic-chemical information, *in vivo* and *in vitro* testing, (Q)SAR models, computational and other non-testing assays, toxicokinetic and toxicodynamic information, category and read-across assessment methodologies) in a weight of evidence approach before entering any other level of the CF. Such existing data/knowledge may be of great value when interpreting the results of endocrine screens/tests, but before they are used, their quality must be evaluated. A quality scoring system such as that recommended by Klimisch *et al*. (1997) can be helpful in this regard. It is also important to know whether an *in vivo* endocrine disruption test has been performed at doses or concentrations which would not be expected to cause systemic toxicity that could mask endocrine effects, or which could cause misleading endocrine changes secondary to general or specific (non-endocrine) organ toxicities.

B.12 Information on metabolism and toxicokinetics is also very valuable. Any available toxicokinetic data (*e.g.* if OECD TG 417 (Toxicokinetics) has been carried out) may help with decisions about route of administration for *in vivo* studies, the relevance of metabolism for *in vitro* studies and the relevance of results from one species to another. For example, if a chemical is metabolised then the addition of metabolising systems to *in vitro* tests should be considered (see below Para B.18). Toxicokinetic studies may also provide information on bioavailability, half-lives for absorption and elimination, and clearance rates, and any nonlinear kinetics resulting from saturation of absorption, which may help with interpretation of toxicity and endocrine data. *In silico* systems are also being developed to predict metabolism, e.g. "Metapath" is a system for simulating xenobiotic metabolism being developed by the joint US, EU, Canadian and Australian project of the OECD Working Group of Pesticides.

B.13 Another important issue concerning initial data collation is the value of extrapolating data from mammalian tests when interpreting data from other vertebrates, and *vice versa*. The broad similarity of endocrine systems across the vertebrates means that such extrapolation can be of considerable value, so it is vital that mammalian toxicologists and wildlife ecotoxicologists who assess endocrine disruption-related data should not operate without reference to each other. Extrapolation of thyroid effects between mammalian and amphibian screening models has been investigated in a recent review (Pickford, 2010). Out of 41 chemicals considered, 32 had been tested in thyroid-sensitive mammalian

screens and 27 in thyroid-sensitive amphibian screens, but only one chemical was reported to exhibit thyroid activity in amphibian assays with the absence of activity in mammalian assays, while none of the chemicals that showed positive results in the mammalian assays were negative in the amphibian assays. Hence, there seems to be a good foundation for extrapolation of qualitative screening level information between these two animal groups, although it should be noted that only the Amphibian Metamorphosis Assay (AMA) is able to identify thyroid agonists and disturbance to peripheral tissue deiodination. However, as noted above, similarities of endocrine systems at the molecular level do not necessarily mean that the physiological outcomes of a given modality will be the same in all vertebrates.

B.1.2 Conceptual Framework Level 2: *In Vitro* Assays Providing Data About Selected Endocrine Mechanism(s) / Pathway(s)

B.14 Assays at this level are screening assays used for hazard detection, identification of possible mechanisms of action (MOAs), prediction of adverse outcome pathways (AOPs), priority-setting, and weight-of-evidence based judgements leading to a conclusion. It is envisaged that a battery of *in vitro* tests would be carried out wherever possible as a single test will usually only provide information on one modality. The results from a combination of tests will increase weight of evidence.

B.15 Certain types of test data might be used to derive preliminary or more advanced judgements about a test chemical. Most *in vitro* assays can also provide "potency" data, based on binding affinity or similar measures. These assays are in most cases deliberately over-responsive (compared with many *in vivo* systems) towards chemicals that bind to a receptor as they are designed to provide alerts for endocrine disruption. In other words, they will provide positives for some chemicals which give no *in vivo* responses, but are intended to minimise the risk that EDs will go undetected. It is noted that lack of metabolic systems in *in vitro* assays may lead to false negatives for chemicals which are bio-transformed to endocrine active metabolites but may potentially also lead to false positives for endocrine active chemicals which are very quickly transformed to endocrine inactive metabolites. Some cell based assays for EDs do have metabolic capability (Coombes, 2000) and it is important to establish whether or not this is the case when starting to use an assay.

B.16 Positive *in vitro* test results indicate the possibility of endocrine disruption effects *in vivo.* Current *in vitro* tests covered by the CF are largely based on mammalian systems, but their results can be used with caution to draw conclusions about possible EDs in other vertebrates, although potency and adverse consequences may differ.

B.17 *In vitro* screens can provide mechanistic data that are useful for the design of further *in vivo* studies. Again, cautious extrapolation to non-mammalian vertebrate *in vivo* tests is feasible.

B.18 *In vitro* screens are relevant for effects in humans and vertebrate wildlife because many are based on highly conserved hormone receptors or interaction with key enzymes or other key molecules involved in the regulation of hormone levels in all vertebrates. Chemicals that bind to these receptors or otherwise interfere with key processes of hormone regulation have the potential to cause effects in *in vivo* studies of both mammals and non-mammalian vertebrate wildlife, assuming concentrations that reach the target are sufficiently high (*e.g.* dependent on ADME).

B.19 Negative *in vitro* results alone cannot be used to exclude possible endocrine disruption activity because of their inherent limitations, such as inability or unknown capacity to metabolically activate toxicants. In addition, chemicals can interfere with the endocrine system in other ways than through the receptor, such as effects on the hypothalamic-pituitary-gonadal axis (HPG) that can only be detected in whole animal studies. For example, chemicals can interfere with the hormonal feedback loops in the HPG axis which could only be revealed in intact animals *e.g.* by changes in hormone levels. Each *in vitro* assay measures a certain mechanism and thus conclusions can be drawn only in the context of what the *in vitro* assay evaluates. However, negative *in vitro* effects should only be

interpreted as a tentative indication of a lack of endocrine disruption activity for the modality in question, if it can be substantiated that the compound does not undergo metabolic activation *e.g.* by the use of ADME information.

B.20 Consideration should be given to the inclusion of metabolising systems in *in vitro* screens: see OECD (2008a) (Detailed Review Paper on the Use of Metabolising Systems for *In vitro* Testing of Endocrine Disruptors. No. 97) and the publication of this (Jacobs *et al.*, 2008). It should be noted however that these systems are not applied on a regular basis with many *in vitro* assays (*e.g.* due to cytotoxicity) and are not validated. Some cell-based Level 2 assays may have limited metabolic capability and this may need to be assessed when setting up the assay. Another possible way of including metabolism is to carry out *in vitro* metabolism studies prior to the Level 2 assays. Identified metabolites or reaction mixture extracts containing metabolites could then be tested. It should be noted that *in vitro* metabolising systems may differ in some respects from *in vivo* systems, so their use still implies some uncertainty. The relative activities of different xenobiotic metabolising enzymes may differ *in vivo* and *in vitro* depending upon availability of cofactors, stability of the enzymes or loss of subcellular compartments.

B.1.3 Conceptual Framework Level 3: *In Vivo* Assays Providing Data about Selected Endocrine Mechanism(s) / Pathway(s)

B.21 Assays at Level 3 provide *in vivo* screening for *possible* endocrine disruption activity. They are designed to provide a yes/no (qualitative) answer about the ability to interact with estrogen, androgen and thyroid hormone receptor mediated modalities, or interfere with steroidogenesis. Other non-receptor processes such as inhibition of iodination of thyroid hormones are also detected. It should be noted that although Level 3 (and 4) assays do not generally expose organisms for a large proportion of their life cycle, and therefore are incapable of revealing the full spectrum of possible ED effects, experience to date suggests that they are sufficiently responsive to identify some EATS active substances.

B.22 Assays at this level are screening assays designed primarily for hazard detection and for revealing mechanisms of action, although some authorities may also use them for making regulatory decisions in some circumstances. These assays are designed to provide alerts to chemicals with possible endocrine disrupting properties, and detect alterations in endocrine-sensitive tissues. Therefore they are of deliberately high responsiveness (*e.g.* use in some cases of castrated/immature animal models without an intact HPG axis, which are therefore unable to compensate fully for endocrine perturbations., Although in the case of the ecotoxicity tests their responsiveness is nevertheless comparable with the high sensitivity of some wildlife species. The route of exposure may not be representative of the natural situation making direct extrapolation to the real world difficult, *e.g.* subcutaneous exposure in an assay when human exposure is dermal or oral.

B.23 They generally include the possibility for metabolic activation (albeit metabolism specific to rodents, fish or amphibians) of a chemical, a feature recommended for, but often absent from current *in vitro* screens.

B.24 Assays are short in duration (*e.g.* the UT and H assays generally have 3 day and 10 day dosing periods respectively whilst the AMA and fish screens employ 3 weeks dosing) and they generally only use very few (or a single) concentrations or dose levels. These assays also provide some information about the potency of a chemical *in vivo*, with respect to the magnitude of a change and the dose/concentration at which the change occurs.

B. 25 It should be noted that both the 21-day fish assay (OECD TG 230) and the fish short term reproduction assay (OECD TG 229) are *in vivo* screens that primarily give information about endocrine disruption mechanisms in adult fish. Additionally, OECD TG 229 includes apical endpoints

(*i.e.* fecundity and by direct association also fertility) which can be affected both by some endocrine disrupting chemicals and some other chemicals toxic to reproduction.

B.26 A **positive** outcome (*i.e*. a statistically significant change(s) in an ED-specific endpoint) of Level 3 assays indicates a possibility for adverse effects in the reproductive and developmental studies at Levels 4 and 5 and may in certain cases (UT assay) indicate effects in immature animals (which may be considered of concern). The specific criteria for a positive result in these assays are given in the 'building blocks' in Section C but are generally significant changes in sex organ weight (UT and H assays), development (AMA), secondary sexual characteristics, and biomarkers such as vitellogenin or spiggin (fish screens).

B.27 However, a compound found **negative** in Level 3 assays can be regarded as inactive against the specific modalities evaluated by those assays, but could still have endocrine disrupting properties mediated through other mechanisms. These may be detected by a more comprehensive Level 4 or 5 assay than those *in vivo* screening assays covered by Level 3, although it is assumed selection of Level 3 assays is generally targeted on a previously suspected mode of action.

B.28 The results from these *in vivo* screens can be used to decide if higher-tier *in vivo* tests should be performed to reduce uncertainty about certain effects of EDs *in vivo* and to gain more information about potency. They may or may not provide data which can be used with confidence in human or wildlife risk assessments because the information does not always indicate whether, or to what extent, adverse effects on apical endpoints have occurred. Also, Level 3 screens do not encompass all possible modes by which EATS systems can be affected.

B.1.4 Conceptual Framework Level 4: *In Vivo* Assays Providing Data on Adverse Effects on Endocrine-Relevant Endpoints

B.29 Assays at Level 4 can provide a more thorough assessment (in comparison with Level 3 assays) of the possible or actual endocrine disrupting effects of a chemical in developing or adult organisms because they are sensitive to more than one mode of endocrine disrupting action. A compound found to be positive indicates a possibility for adverse effects and which may require further investigation. However, if sufficient data for decision making are available, further animal testing is not necessary. At this level, assays have numerous endpoints and therefore the criteria for a positive result are more complex than at lower levels, but generally a chemically-induced, biologically significant change in an endocrine endpoint would be considered a positive result. A compound found to be negative is inactive under the specific conditions evaluated by the assay. A negative conclusion regarding endocrine disruption, however, requires combined lines of evidence because as with Level 3 assays, a compound found negative in a Level 4 assay may still have endocrine disrupting properties either mediated through mechanisms not covered by the assay or because the assay was not sufficiently sensitive. However, it is assumed that a particular assay is selected to address a specific, suspected mode of action.

B.30 This level includes assays that are not specifically designed to detect EDs but have endpoints that are highly relevant for their detection. These assays include many standard repeated dose mammalian toxicology tests *e.g.* OECD TG 407 (28-day Repeated Dose Toxicity Test) and OECD TG 408 (90-Day Repeated Dose Toxicity Test). Most of these standard toxicology tests have not been validated for detection of EDs, with the exception of the 28-day Repeated Dose Toxicity Test (OECD TG 407). This updated assay has been validated for some endocrine endpoints but the sensitivity of the assay is not sufficient to identify all EATS-mediated EDs. The validation of the assay (OECD, 2006b) showed that it identified strong and moderate EDs acting through the ER and AR; and EDs weakly and strongly affecting thyroid function. It was relatively insensitive to weak EDs acting through the ER and AR.

B.31 The reproduction/developmental screening tests OECD TG 421 and 422 are included in Level 4 as supplemental tests because they give limited but useful information on interaction with endocrine systems. EDs may be detected by effects on reproduction (gestation, gestation length, dystocia, implantation losses), genital malformations in offspring, marked feminized AGD in males, changes in histopathology of sex organs or effects on the thyroid gland. The one generation assay (OECD TG 415) is also included at this level. This assay provides a more thorough assessment of effects on reproduction and development than OECD TG 421/422 but is not as comprehensive as the reproductive studies in Level 5.

B.32 The Prenatal Developmental toxicity (OECD TG 414) and the Developmental Neurotoxicity (OECD TG 426) studies are also included in Level 4 as they involve repeated dosing of pregnant females and therefore potential exposure of the developing fetus. Both assays include some endpoints that may detect endocrine disruption (*e.g.* abnormalities of male and female genitalia).

B.33 All assays at this level include apical endpoints and are designed both for hazard and risk assessment. The use of intact animal models provides an evaluation under normal physiological conditions but the responsiveness of these assays may be lower than Level 3 assays as hormone feedback mechanisms may provide some compensation in the case of EDs. Depending upon the guideline/protocol used, the fact that a substance may interact with a hormone system in these assays does not necessarily mean that it will cause adverse effects in humans *e.g.* the results for a chemical tested in the male or female pubertal assays with only two dose levels may not provide sufficient information on adverse effects. However, for ecological systems, effects on apical endpoints at this level, such as fecundity, would be considered adverse.

B.34 Level 4 assays may provide information about the potency of a compound which may be investigated further at Level 5, although some of these assays (*e.g.* the fish sexual development test and the peripubertal assays) may test relatively few concentrations or dose levels, thus limiting the precision of the results, and hence their usefulness for setting 'safe' concentrations or doses in a risk assessment. Effects on some endpoints included in the assays can be considered as adverse apical impacts (*e.g.* major histopathologic changes in reproductive organs in rats; biased phenotypic sex ratios in developing fish) while others represent an effect on an indicator of hormonal activity for either humans or wildlife (*e.g.* changes in thyroid hormone levels or vitellogenin titres).

B.35 Level 4 tests (*e.g.* the Fish Sexual Development Test or the 28-day Repeated Dose Toxicity Test (OECD TG 407)) may also support an evaluation about whether specific endocrine-mediated effects are more or less sensitive than general toxicity. This of course only applies if the tests have sufficient statistical power, test an appropriate range of concentrations, and are conducted under conditions comparable to standard tests.t.

B.36 Some (*e.g.* the Fish Sexual Development Test or the 28-day Repeated Dose Toxicity Test (OECD TG 407)), but not all, Level 4 assays can therefore provide data on adverse effects which may be sufficient for use in hazard assessments, or in risk assessments which seek to identify 'safe' concentrations or doses. However, most do not provide more comprehensive information about possible endocrine disrupting effects such as those obtainable from lifecycle experiments (Level 5).

B.1.5 Conceptual Framework 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

B.37 The developmental and reproductive toxicity studies at Level 5 provide data on adverse effects and are especially useful for risk assessment as they add to the weight of evidence concerning the potential for impacts in humans and vertebrate wildlife, and provide data on dose/concentration-response. The effects observed in reproductive tests with rodents, and in partial or full lifecycle toxicity studies with fish, amphibians and birds, may be due to endocrine disruption or other

mechanisms, but the effect or pattern of effects, *e.g.* decreased anogenital distance and malformations of reproductive organs in male rats, may indicate that effects mediated via impact on the endocrine system are involved. Some of these tests may also include measurement of endpoints which are indicative of endocrine disruption activity (*e.g.* altered sex ratio in the fish lifecycle test, alteration of puberty onset in mammalian multigeneration tests).

B.38 Among the current OECD Test Guidelines for mammalian reproductive toxicity, exposure during all vulnerable periods of development is performed in the two-generation reproductive toxicity study design (OECD TG 416). This was updated in 2001 with endocrine disruption sensitive endpoints such as, VO, PPS, estrous cyclicity, evaluation of primordial follicle counts, AGD (triggered by sex ratio in F1) *etc*. This study provides a wealth of information, particularly if combined with data from long-term repeat dosing studies *e.g* the 90-d repeated dose test (OECD TG 408) where the histopathology of the thyroid and mammary gland and possibly hormone data could be available. However older reproductive toxicity studies that lack sensitive endpoints (*e.g.* onset of puberty) cannot fully exclude the possibility that chemicals testing negative may still be EDs. The updated OECD TG 416 does not include some endocrine disruption-related sensitive endpoints such as nipple retention. Late effects becoming manifest after weaning of the animals are partly covered in young adults, especially in relation to reproductive function and developmental neurotoxicity, but other potentially important late effects such as premature reproductive senescence (Cooper *et al*, 2007) are also not assessed. Effects becoming manifest during ageing are not included in any current guidelines for reproductive toxicity but are being reviewed by OECD. It is recognised that at the present time level 5 assays do not cover all endocrine outcomes and this review should address these gaps.

B.39 A number of enhancements of the OECD test guidelines for reproductive toxicity in rodents for the detection of effects of EDs are in development. The new extended one generation reproductive toxicity study) EOGRTS) (OECD TG 443) includes more endpoints sensitive to endocrine disruption than OECD TG 416 and, as it also uses reduced animal numbers, it is expected that it will often replace OECD TG 416 for mammalian reproductive toxicity testing. Endpoints sensitive to endocrine disruption, not specified in OECD TG 416, include areola/nipple retention, anogenital distance at birth, measurement of thyroid hormones and TSH levels. Effects on the developing nervous and immune systems are also assessed. These systems may also be sensitive to endocrine influences. This test is also expected to have greater sensitivity than OECD TG 416 as it requires an increased number of pups to be examined. In summary, 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

B.40 Thus, one and two generation studies conducted before the inclusion of sensitive endocrine endpoints (*e.g.* sexual maturation) by themselves may not be considered adequate for demonstrating the probable absence of endocrine disrupting activity although they still provide much valuable data (mainly restricted to fertility and effects on reproductive organs).

B.41 In contrast, fish and bird single- or multi-generation lifecycle tests (some of which are currently being validated and have not yet been developed into OECD test guidelines) include evaluation of exposure of many endocrine disruption-sensitive processes, and thus there is a higher level of confidence about negative tests. The degree of confidence will nevertheless still be constrained by the statistical power of the test and the ability to control study conditions across multiple generations. This particularly applies to the multi-generation test with medaka which is under development as a draft OECD test guideline. This test covers *inter alia* the possibility of detecting effects caused by the maternal transfer to offspring of bioaccumulative EDs.

.

B.2 Endpoints in the Various Assays of the Conceptual Framework

B.42 In order to facilitate the interpretation of hazard data derived from screens and tests in the Revised Conceptual Framework, the following table (Table B.1) presents a listing of possible endpoints and their applicability for identifying endocrine disrupting mechanisms and/or effects resulting from the four modalities under consideration (*i.e.* estrogen-mediated activity, androgen-mediated activity, thyroid-related activity and steroidogenesis disruption related-activity). Endpoints for those assays that have not yet received full validation for endocrine outcomes, or are test guidelines that are not primarily designed for testing specifically for endocrine disruption are listed in Annex 2 (Table Annex 2).

B.43 Where possible, the direction of change is indicated for the endpoints. The data from validation studies on the assays has been used to guide the changes as much as possible, although in some cases it has not been possible to generalise and in other cases extrapolations have been made across similar endpoints in different studies *e.g.* OECD TG 416 has not been validated for thyroid-related activities but it is reasonable to suppose that thyroid changes in OECD TG 416 would be similar to those seen in the OECD TG 407 and the pubertal assays. In all cases the direction of change is illustrative and not all possibilities are given, *e.g.* for steroidogenesis disruption, only inhibition of steroidogenic enzymes is illustrated reflecting the chemicals used in validation studies whereas in theory induction may be possible. Specific chemicals may also differ in the endpoints affected and the direction of change. Table B.1 also lists those endpoints which are not directly linked to endocrine disruption-related mechanisms.

B.44 The endpoints listed are those specified in the guideline (either OECD or OPPTS), or those most commonly used, for methods for which no guidelines are available. Other endpoints may be added, particularly changes in titres of hormones such as estradiol, testosterone, LH, FSH *etc*, are frequently added to OECD TG 407, OECD TG 412 for example.

B.45 However, it should be noted that several assays with wildlife species (especially the larval amphibian growth and development assay, the avian reproduction test, and the fish lifecycle tests) and the CF Level 4 and 5 mammalian assays are not solely designed to detect the effects of endocrine disrupters, but they are expected to be sensitive to many such chemicals, as well as to other reproductively toxic materials. Furthermore, most of these assays with wildlife species are still in development, so a full description of their reactions to the types of EDs under consideration here cannot yet be given.

Table B.1. Endpoints relevant for endocrine disruption modalities in Test Guidelines and other endocrine disruption-sensitive assays (in the revised Conceptual Framework) for which guidance with interpretation of data have been developed.

Probable direction of change is indicated where possible**.**

Note that for many assays, individual endpoints may not in themselves be diagnostic of an endocrine disruption modality. Such diagnosis often relies on a combination of endpoints or assays in a weight of evidence assessment.

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
***In vitro* screens**							
ER Binding Assay (US EPA OPPTS 890.1250) [Table C.2.1]	Displacement of ligand from receptor. Binding cannot distinguish between agonism or antagonism		Nil	Nil	Nil	Nil	Nil
AR Binding Assay (US EPA OPPTS 890.1150) [Table C.2.2]	Nil	Nil	Displacement of ligand from receptor. Binding cannot distinguish between agonism or antagonism		Nil	Nil	Nil

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
OECD TG 455: Stably transfected hERα transcriptional activation assay (ER STTA) (including guidance for the antagonism assay) [Table C.2.3]	Activation of reporter gene linked to ER	Inhibition of activation of reporter gene linked to ER Note: this is not addressed in OECD TG 455 as the antagonist assay is currently In validation	Nil	Nil	Nil	Nil	Activators of the Ah receptor may inhibit activation of reporter gene linked to ER through crosstalk at the DNA level
OECD TG 456: H295R steroidogenesis assay [Table C.2.4]	Nil	Nil	Nil	Nil	Nil	Inhibition and induction of estradiol and testosterone synthesis	Nil
Aromatase Assay (US EPA OPPTS 890.1200) [Table C.2.5]	Nil	Nil	Nil	Nil	Nil	Inhibition of aromatase (CYP 19) activity	Nil

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
Wildlife *in vivo* screens and tests							
OECD TG 229: Fish short-term reproduction assay (FSTRA) [Table C.3.1]	VTG induction in males Depression of male 2^{o} sex characteristics in fathead minnow or medaka Specific gonad histopathologic findings as listed in OECD (2010a)***	VTG depression in females (assuming no systemic toxicity) Specific gonad histopathologic findings as listed in OECD (2010a)***	Induction of male 2^{o} sex characteristics in female fathead minnow or medaka Specific gonad histopathologic findings as listed in OECD (2010a)***	Depression of male 2^{o} sex characteristics in fathead minnow or medaka Specific gonad histopathologic findings as listed in OECD (2010a)***	Nil	Possible effects on:- VTG depression in females (assuming no systemic toxicity) Gonad histo-pathology (*e.g.* Leydig cell hyperplasia – see OECD 2010a)***	Fecundity depression Certain histopathologic findings not related to endocrine activity Behaviour

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
OECD TG 230: 21-Day Fish Assay [Table C.3.2]	VTG induction in males Depression of male 2^o sex characteristics in fathead minnow or medaka	VTG depression in females (assuming no systemic toxicity)	Induction of male 2^o sex characteristics in female fathead minnow or medaka	Depression of male 2^o sex characteristics in fathead minnow or medaka	Nil	Possible effects on:- VTG depression in females (assuming no systemic toxicity)	Behaviour Certain histopathologic findings (if measured – see OECD 2010a)
Androgenised female stickleback screen (AFSS) (GD 140) [Table C.3.3]	Nil	Nil	Spiggin induction	Spiggin depression	Nil	Nil	Nil

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
Fish sexual development test (FSDT) (OECD TG 234) [Table C.3.4]	Female-biased phenotypic sex ratio* VTG induction in males and females Specific gonad histopathologic findings (optional) as listed in OECD (2010a)***	Male-biased phenotypic sex ratio* Increase in sexually undifferentiated fish. VTG depression in females Specific gonad histopathologic findings (optional) as listed in OECD (2010a)***	Male-biased phenotypic sex ratio* VTG depression in males and females Specific gonad histopathologic findings (optional) as listed in OECD (2010a)***	Induction of intersex fish VTG induction in females Female-biased phenotypic sex ratio* Specific gonad histopathologic findings (optional) as listed in OECD (2010a)***	Nil	Possible effects on:- Male-biased phenotypic sex ratio* VTG depression in males and females	Body length Body weight Morphological abnormalities Certain histopathologic findings not related to endocrine activity

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
Fish Lifecycle Toxicity Test (FLCTT) (US EPA OPPTS 850.1500, possibly with endocrine-sensitive additions). Note: No endpoints specific to a particular EATS modality are included at present but endpoints indicative of endocrine activity could be added if validated. [Table C.3.5]	Female-biased phenotypic sex ratio* VTG induction in males Altered levels of estradiol and/or (keto) testosterone	?	Male-biased phenotypic sex ratio* Altered levels of estradiol and/or (keto) testosterone	?	Altered levels of thyroid hormones	Possible effects on:- VTG depression in females	Hatching success Weight Length Behaviour Gross morphology Gonado-somatic index Multiple organ histopathology Time to maturity (time to first spawn) Fecundity Fertilisation success
OECD TG 231:Amphibian	Nil	Nil	Nil	Nil	Developmental	Nil	Body weight

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
metamorphosis assay (AMA) [Table C.3.6]					stage** Hind limb length** Snout-vent length** Thyroid gland histopathology (see OECD TG 231 for interpretation of combined effects – individual changes may not be diagnostic)		
OECD TG 206: Avian reproduction test. Note: No endpoints specific to a particular endocrine disruption modality are included at	Nil	Nil	Nil	Nil	Nil	Nil	Egg production Cracked eggs Eggshell thickness

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
present but diagnostic endpoints could be added (*e.g.* vitellogenin). [Table C.3.7]							Egg viability Hatchability Body weight Gross pathology
Mammalian *in vivo* screens and tests							
OECD TG 440: Uterotrophic bioassay in rodents (UT assay) (including GD for antiestrogenicity screen) (immature female or adult after OVX) [Table C.4.1]	Uterine weight (wet and blotted) increase. Optional: keratinisation and cornification of vagina, proliferation of endometrial epithelium, changes in uterine histopathology.	Reduction of estrogen-stimulated uterine weight increase. Note: TG does not include antagonist determination. This is described in a GD (OECD 2007) Optional:	Uterine weight (wet and blotted) increase. (Aromatisable) androgens can increase uterine weight in both immature and OVX female rats.	Nil	Nil	Nil	The immature rodent assay where the HPG axis is intact, may detect other modes of action *e.g.* related to GnRH inhibition

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
		reduction of other estrogen-stimulated histopathologic changes					
OECD TG 441: Hershberger bioassay (H assay) (adult male after castration) (including GD for weanling Hershberger bioassay) [Table C.4.2]	Nil	Nil	Increase in weight of ventral prostate, seminal vesicles, LABC, cowpers glands, glans penis (+ve outcome if 2 or more tissues are increased). Note in the weanling H assay: glans penis is not included, testis weight is decreased. Optional:	Reduction of androgen-stimulated weights of ventral prostate, seminal vesicles, LABC, cowpers glands, glans penis (+ve outcome if 2 or more tissues are decreased). Note in the weanling H assay: glans penis is not included, testis weight is increased.	Optional: Changes in serum T4 and T3. Histopathologic changes in thyroid	Nil	Nil

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
			changes in serum hormones.	Optional: changes in serum hormones.			
Male pubertal assay (PP Male Assay) (US EPA OPPTS 890.1500) (no OECD TG available) [Table C.4.3]	Assay is not designed to detect this modality but the following changes may occur: Increased age at preputial separation. Decreased weight of seminal vesicles (+ coagulating glands), ventral prostate, dorsolateral prostate, LABC, epididymides. Decreased testis	Assay is not designed to detect this modality and studies using pure antagonists are lacking. However, the following changes may occur in the following endpoints: Age at preputial separation. Weight of seminal vesicles (+ coagulating glands), ventral prostate, dorsolateral	Decreased age at preputial separation. Increased weight of seminal vesicles (+ coagulating glands), ventral prostate, dorsolateral prostate, LABC, epididymides. Decreased testis weight. Histopathologic changes in testes, epididymides. Decreased serum	Increased age at preputial separation. Decreased weight of seminal vesicles (+ coagulating glands), ventral prostate, dorsolateral prostate, LABC, epididymides. Increased testis weight. Histopathologic changes in testes, epididymides. Increased serum	Increased thyroid weight Possible liver weight increase (in combination with other thyroid-related endpoints). Histopathologic changes in thyroid (follicular cell height increase & colloid area decrease) Serum T4 decreased, TSH increased.	Possible effects on: Preputial separation. Weight of seminal vesicles (+ coagulating glands), ventral prostate, dorsolateral prostate, LABC, epididymides. Histopathologic changes in testes epididymides. Serum testosterone	Changes in weight of pituitary and/or adrenals.

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
	weight. Histopathologic changes in testes epididymides. Increased serum testosterone	prostate, LABC, epididymides. Testis weight. Histopathologic changes in testes epididymides. Serum testosterone	testosterone	testosterone			
Female pubertal assay (PP Female Assay) US EPA OPPTS 890.1450) (no OECD TG available) [Table C.4.4]	Decreased age at Vaginal opening. Increased weight of uterus & decreased weight of ovaries Histopathologic changes in uterus & ovaries. Decreased age at first estrus.	Studies using pure antagonists are lacking but the following changes may occur: Increased age at Vaginal opening. Decreased weight of uterus Histopathologic changes in	Assay is not designed to detect this modality but the following changes may occur: Increased age at Vaginal opening. Decreased weight of uterus & ovaries.	Assay is not designed to detect this modality but the following changes may occur: Decreased age at Vaginal opening. Decreased weight of ovaries.	Increased thyroid weight Possible liver weight increase (in combination with other thyroid-related endpoints). Histopathologic changes in thyroid (follicular cell height increase	Possible effects on: Age at vaginal opening. Weight of uterus and ovaries. Histopathologic changes in uterus & ovaries. Estrus cyclicity.	Changes in weight of pituitary and/or adrenals.

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
	Changes in estrus cyclicity.	uterus & ovaries. Increased age at first estrus. Changes in estrus cyclicity.	Histopathologic changes in uterus & ovaries. Increased age at first estrus. Changes in estrus cyclicity.	Histopathologic changes in uterus & ovaries.	& colloid area decrease) Serum T4 decrease, TSH increased.		
OECD TG 407: Repeated dose 28-day oral toxicity study in rodents [Table C.4.5]	Histopathologic changes in ovary, uterus/cervix, vagina. Decrease in weight of epididymides, prostate + seminal vesicles with coagulating glands. Histopathologic changes in	Studies using pure antagonists are lacking. However, changes may occur in the following: Histopathologic changes in ovary, uterus/cervix, vagina. Changes in weight of	Histopathologic changes in ovary, uterus/cervix, vagina. Increase in weight of prostate + seminal vesicles with coagulating glands. Decrease in weight of testes. Histopathologic	Decrease in weight of epididymides, prostate + seminal vesicles with coagulating glands. Histopathologic changes in testes, epididymides, prostate + seminal vesicles with coagulating	Possible liver weight increase (in combination with other thyroid-related endpoints). Histopathologic changes in thyroid (follicular cell height increase & colloid area decrease)	Possible effects on: Histopathologic changes in ovary, uterus/cervix, vagina. Weight of, prostate + seminal vesicles with coagulating glands. Optional	Histopathologic changes in adrenal. Optional: Histopathologic changes in pituitary and mammary glands.

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
	testes, epididymides, prostate + seminal vesicles with coagulating glands Optional endpoints: Increase in weight of uterus (slight), decrease in weight of ovaries. Changes in vaginal smears. Histopathologic changes in mammary glands (males).	epididymides, prostate + seminal vesicles with coagulating glands. Histopathologic changes in testes, epididymides, prostate + seminal vesicles with coagulating glands Optional endpoints: Uterine/ovary weight. Changes in vaginal smears. Histopathologic changes in mammary	changes in testes, epididymides, Optional endpoints: Ovary/ weight (decrease). Changes in vaginal smears. Histopathologic changes in mammary glands.	glands Optional endpoints: ovary weight (decrease).	Optional: Serum T3 and T4 decreased, TSH increased. Increased thyroid weight.	endpoints: Uterine and ovary weight Changes in vaginal smears. Histopathologic changes in mammary gland histopathology.	

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
		glands.					
OECD TG 416: 2-generation reproduction toxicity study (including guidance on OECD TG 415: 1-generation study) [Table C.4.6]	Change in AGD in male and female pups. Changes in estrus cyclicity (P, F1). Decreased age at Vaginal opening (F1). Increased age at preputial separation (F1). Changes in weights of: (P, F1) uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating	Studies using pure antagonists are lacking. However, changes may occur in the following: AGD in male and female pups. Estrus cyclicity (P, F1). Age at Vaginal opening (F1). Age at preputial separation (F1). Weights of: (P, F1) uterus, ovaries, testes, epididymides, prostate, seminal	Studies using agonists are lacking. However, changes may occur in the following: Increased AGD in male pups, change in AGD in female pups. Estrus cyclicity (P, F1). Age at Vaginal opening (F1). Age at preputial separation (F1). Weights of: (P, F1) uterus, ovaries, testes,	Decreased AGD in male pups, change in AGD in female pups. Changes in estrus cyclicity (P, F1). Changes in age at vaginal opening (F1). Increased age at preputial separation (F1). Changes in weights of: (P, F1) uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating	Increased thyroid weight. Possible liver weight increase (in combination with other thyroid-related endpoints). Histopathologic changes in thyroid (follicular cell height increase & colloid area decrease)	Possible effects on: AGD in male and female pups. Estrus cyclicity (P, F1). Age at Vaginal opening (F1). Age at preputial separation (F1). Changes in weights of: (P, F1) uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating glands).	Changes in : Weights of adrenals Time to mating Male fertility Female fertility Gestation length Dystocia Placental weight Number of implantations, corpora lutea Number of live births and pre and post implantation loss

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
	glands). Histopathologic changes in vagina, uterus (+ cervix), ovaries, testis, epididymis, prostate, seminal vesicles and coagulating glands. Changes in sperm parameters: Sperm numbers (testicular homog resistant spermatids & cauda epididymides sperm reserve), sperm motility, sperm morphology (P,	vesicles (+ coagulating glands). Histopathologic changes in the above organs Sperm parameters: Sperm numbers (testicular homog resistant spermatids & cauda epididymides sperm reserve), sperm motility, sperm morphology (P, F1).	epididymides, prostate, seminal vesicles (+ coagulating glands). Histopathologic changes in the above organs Sperm parameters: Sperm numbers (testicular homog resistant spermatids & cauda epididymides sperm reserve), sperm motility, sperm morphology (P, F1).	glands). Histopathologic changes in the above organs Changes in sperm parameters: Sperm numbers (testicular homog resistant spermatids & cauda epididymides sperm reserve), sperm motility, sperm morphology (P, F1).		Histopathologic changes in the above organs. Changes in sperm parameters: Sperm numbers (testicular homog resistant spermatids & cauda epididymides sperm reserve), sperm motility, sperm morphology (P, F1).	Litter size Sex ratio (F1, F2). Litter/pup weight Pup survival index Abnormalities in pup development (F1, F2).

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
	F1).						
OECD TG 443: Extended One-Generation Reproductive Toxicity Study [Table C.4.7]	Change in AGD in male and female pups. Changes in estrus cyclicity (P, F1). Decreased age at Vaginal opening (F1). Increased age at preputial separation (F1). Genital abnormalities. Changes in weights of: (P, F1) uterus, ovaries, testes, epididymides, prostate, seminal	Studies using pure antagonists are lacking. However, changes may occur in the following: Change in AGD in male and female pups. Estrus cyclicity (P, F1). Age at Vaginal opening (F1). Age at preputial separation (F1). Genital abnormalities. Weights of: (P,	Studies using agonists are lacking. However, changes may occur in the following: Increased AGD in male pups, change in AGD in female pups.Age at preputial separation (F1). Genital abnormalities. Weights of: (P, F1) uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+	Decreased AGD in male pups, change in AGD in female pups.Increased age at preputial separation (F1). Genital abnormalities. Nipple retention. Changes in weights of: (P, F1) testes, epididymides, prostate, seminal vesicles (+ coagulating glands). Histopathologic changes in the above organs	Increased thyroid weight. Possible liver weight increase (in combination with other thyroid-related endpoints). Histopathologic changes in thyroid. Serum T4, decreased, TSH increased.	Possible effects on: AGD in male and female pups. Estrus cyclicity (P, F1). Age at Vaginal opening (F1). Age at preputial separation (F1). Genital abnormalities. Changes in weights of: (P, F1) uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+	Changes in weights of adrenals and pituitary. Histopathologic changes in adrenals. Changes in : Time to mating Male fertility Female fertility Dystocia Gestation length Number of implantations, corpora lutea Number of

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
	vesicles (+ coagulating glands). Histopathologic changes in vagina, uterus (+ cervix), ovaries, testis, epididymis, prostate, seminal vesicles and coagulating glands. Histopathologic changes (proliferative) in mammary glands. Changes in sperm parameters: Sperm numbers	F1) uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating glands). Histopathologic changes in the above organs Histopathologic changes in mammary glands. Changes in sperm parameters: Sperm numbers sperm motility, sperm morphology (P, F1).	coagulating glands). Histopathologic changes in the above organs and in mammary glands. Changes in sperm parameters: Sperm numbers sperm motility, sperm morphology (P, F1).	and in mammary glands. Changes in sperm parameters: Sperm numbers sperm motility, sperm morphology (P, F1).		coagulating glands). Histopathologic changes in the above organs Changes in sperm parameters: Sperm numbers sperm motility, sperm morphology (P, F1). Histopathologic changes in mammary glands.	ovarian follicles Number of live births and post implantation loss Litter size Viability index Placental weight Sex ratio (F1). Litter/pup weight Pup survival index Abnormalities in pup development (F1). Apical endpoints from the developmental neuro- and

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
	sperm motility, sperm morphology (P, F1).						immunotoxicity cohorts may be sensitive to endocrine modulation.

*Simultaneous measurement of genotypic sex ratio (in medaka, zebrafish or stickleback at present) allows a more powerful detection of any effects on phenotypic sex ratio. However, sufficient power can be achieved by using an appropriate number of animals with phenotypic sexing alone.

**Accelerated or asynchronous development is considered by many authorities to be diagnostic of thyroid active chemicals, in addition to abnormal thyroid histopathology. Retarded development may be due either to thyroid-active chemicals or to systemic toxicants.

***Primary histopathological criteria in gonads include the following: Males – increased spermatogonia; testis-ova; testicular degeneration; Leydig cell hyperplasia/hypertrophy. Females – increased oocyte atresia; perifollicular cell hyperplasia/hypertrophy; decreased yolk formation; changes in overian staging.

C. Specific Guidance for the Test Guidelines Addressed

C.1 Introduction to Specific Guidance

C.1.1 This Introduction applies to all assays covered by the GD, including those in Annex 2, although it should be noted that Annex 2 guidance remains provisional until those assays have been fully validated with EDs.

C.1.2 As indicated earlier, the information given in Section C (and to a more limited extent in Annex 2) is intended to provide guidance on the interpretation of data from individual assays, and on a possible next step for obtaining additional data, if required by a given user. It is important to understand that the guidance should be used flexibly in the light of local regulatory circumstances and available data – it is not a rigid prescription, but should be considered as a decision-support tool.

C.1.3 Discussion of each assay takes the form of textual guidance which describes the basis of the assay and any special considerations or limitations, when and why the assay is likely to be used, and what broad conclusions may be appropriate when one is in possession of positive, negative, or equivocal results. This is followed by a table (known as a 'building block') that elaborates that guidance for each of a number of data scenarios. Thus, for each type of assay result, the guidance varies depending on the type and amount of pre-existing data (both *in vitro* and *in vivo*). The intention has been to cover all the major possible scenarios, but the document cannot address all eventualities. Furthermore, it is implicit that expert advice will need to be consulted at many points in these building blocks – they are not recipes which can be followed blindly. Note that some scenarios are much less likely to occur than others – for example, it is unlikely (but still possible) that a higher tier procedure such as a fish life-cycle test will have been performed in the absence of various screening assays. A large range of possible scenarios has, therefore, been described for the sake of completeness.

C.1.4 When considering a possible 'next step' in evidence-gathering that could follow from a particular result in an *in vitro* assay, guidance is given in the next section about suitable *in vivo* testing with mammalian or wildlife species. Similar guidance is not given concerning possible mammalian tests that might be conducted following positive wildlife tests, and *vice versa*. At the present state of knowledge, such guidance is not considered to be reliable. However, it is clear that a positive result in an ED-responsive mammalian assay could be interpreted as an alert about possible related effects in wildlife, and the reverse also applies (although mammalian assays will often have been performed before any with wildlife). Due to the difficulties associated with reading-across from mammalian toxicity data to possible effects in non-mammalian wildlife, it may be considered that positive mammalian assays should generally result in some wildlife testing if the hazards experienced by the latter group are to be taken into account. On the other hand, insufficient data yet exist to be confident that negative mammalian data imply an absence of effects in wildlife.

C.1.5 It will be apparent that the underlying approach when implementing this guidance is to consider the weight of available evidence – situations in which a single assay provides conclusive evidence that a chemical is an ED may not be common although there will be exceptions. For example, feminized AGD in male offspring (observed in OECD TG 416 and possibly in OECD TG 421/422) may be considered as conclusive evidence of an endocrine disrupting effect. OECD GD 43 (GD on Mammalian Reproductive Toxicity Testing and Assessment; OECD 2008c) states "A statistically significant change in AGD that cannot be explained by the size of the animal indicates effects of the exposure and should be used for setting the NOAEL." It is vital to consider all relevant data on the test chemical, including their quantity, their type, and their quality. For example, without adequate mechanistic data from (Q)SARs or *in vitro* assays, or from the *in vivo* assay under consideration, it will often not be possible to conclude with confidence that any apical effects have been caused by an endocrine disrupting mode of action. Indeed, any linkage between mechanistic data

and apical responses will probably have to be assessed according to the weight of evidence and is unlikely to be confirmed absolutely. Another example of the use of weight of evidence concerns *in vivo* screening assays which may indicate that a chemical can interfere with the endocrine system in intact animals, but will sometimes not be able to provide data on apical effects, or supply information which could be used on its own in a full risk assessment. In such situations, more complete apical data may have to be obtained from a higher tier test, which will then be evaluated in conjunction with the screening data. Note, however, that negative data from a higher tier test should generally be given more weight than positive data from a lower tier screen, assuming the same class of vertebrates has been employed at both tiers, the quality of the data is good, the suspected mechanism or mode of action is adequately covered by apical endpoints, and a sensitive life stage has been used in the higher tier negative test.

C.1.6 The guidance in this document is considered reliable for EATS modalities, although the assays in Annex 2 have not yet been fully validated. However, the field of endocrine disruption continues to develop, so for that reason, this is a 'living document' which will be subject to amendment as new data are generated, new modalities are described, and new assays are published as Technical Guidelines.

C.1.7 Users of this GD should be aware that comparisons of no-effect doses or concentrations from different types of test may be very difficult or impossible. This is obvious if one is trying to compare an oral dose in a mammalian or avian test with an ambient concentration in an aquatic test. However, caution should also be used when making comparisons within these two major types of test if different methods have been used to calculate the no-effect dose or concentration (*e.g.* if test concentrations in one test were nominal and in the other were measured).

C.2 *In Vitro* Screens

C.2.1 ER binding Assay (US EPA OPPTS 890.1250)

C.2.1.1 Modality detected/endpoints: Binding to ER isoforms

Background to the Assay

C.2.1.2 The ER Binding Assay is an *in vitro* screening assay to detect substances that bind to ERs. The assay has been in use for many years and there are different variations of the protocol. The most commonly used protocol utilises rat uterine cytosol as a source of ER without further purifications of ER isoforms. Binding therefore occurs to a mixture of ERα and ERβ, although the primary isoform in rat uterine cytosol is ERα. Human ERα prepared as a recombinant protein is now available and will replace the use of rat cytosol when successful validation has been completed. There is no OECD TG available for this assay but the method is in common use and an OECD test guideline is likely to be developed in 2011. The ER binding assay was chosen to be one of the suite of assays comprising US EPA's "Tier 1" and has been validated in that context (USEPA, 2009a). The US EPA (OPPTS) guideline is therefore available (USEPA, 2009). In this context, the assay provides information on the ability of a compound to interact with ERs but is not intended to be used to show that the interaction is, specifically, one-site competitive binding, or to characterize precisely the strength of the binding. The assay determines the ability of a chemical to displace a radiolabeled ligand (17β-estradiol) from ER and provides a positive or negative result for the ability to bind to ER.

C.2.1.3 Chemicals that bind to ER may induce hormone-dependent transcriptional activity (agonist) or block normal hormone function by preventing the endogenous hormone from binding to the receptor (antagonist). The binding assay does not distinguish between these. The hormone-binding domain of the ER is highly conserved across vertebrate species and therefore represents a simple evaluation of estrogenic potential that is relevant to many taxa. A positive result in guideline OPPTS 890.1250 requires demonstration of a concentration response curve for the ability of the test chemical to displace radiolabelled 17β-estradiol. The concentration response curve allows the determination of potency *i.e.* IC50 (concentration at which 50% of radioligand is displaced by the test chemical) and relative binding affinity by comparing the Log (IC50) of 17β-estradiol with that of the test chemical. The OECD Validation Management Group for non-animal tests is discussing the statistical analysis of this assay which will identify a positive result and this will be included in the new OECD test guideline.

C.2.1.4 The ER binding assay may suffer from variability in response, if not performed exactly as stated in the protocol *e.g.* if the receptor concentration in the cytosol is too low or too high, or the tubes are not kept cold at all times during the experiment. Performance criteria are therefore specified in order to demonstrate that the assay is functioning correctly. Proficiency chemicals are also used on each run to demonstrate the sensitivity of the experiment (reference standard: 17β-estradiol; weak positive control: norethynodrel and negative control: octyltriethoxysilane). Compliance with the performance criteria should be checked before evaluating results from this assay to ensure that most have been met. Small deviations are unlikely to have compromised the assay but judgement should be made on a case-by-case basis.

When/Why the Assay May be Used

C.2.1.5 Although the ER binding assay may be used at any stage in the hazard assessment process, the most likely use scenario is during initial assessment of chemicals for their ability to interact with endocrine systems *in vitro*, *i.e.* EATS modalities. Assays for interaction with other modalities *e.g.* AR and steroidogenesis interference, are likely to be conducted at the same time so that all results can be considered together. TR and other assays concerning mechanisms of thyroid disruption may be conducted but the methods for these are not in common use and are not validated (see paragraphs A.3 and C.1). The ER binding assay does not include the use of a xenobiotic metabolising system but consideration should be given to the inclusion of this (Jacobs *et al*, 2008; OECD, 2008a) depending upon the circumstances *e.g.* if the metabolism of a chemical is unknown, although the methods for inclusion of xenobiotic metabolising systems are not yet validated (see paragraph B.18). Alternatively, for a chemical with known metabolites, these could also be tested in the ER binding assay. Another use scenario may be following effects obtained in higher tier tests, for example delayed or accelerated puberty onset in females, which could be indicative of an effect mediated by ER. Selection of the most appropriate tests has to be on a case-by-case basis but also considering the need to minimise animal testing. A further example could be results obtained in other apical assays, *e.g.* in OECD TG 408 (90-day toxicity test), where effects on reproductive organs could be investigated further by testing in the ER binding assay in combination with AR and steroidogenesis based assays.

Introduction to the Table of Scenarios

C.2.1.6 Table C.2.1 gives guidance on a further step to take in the event of a positive (+) or negative (-) result in the ER binding assay and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.2.1.7 The results of the ER binding assay are given in the second column. Criteria for positive, negative and equivocal results are given in the OPPTS guideline. A result is judged positive ("interactive with ER") if the chemical will displace at least 50% of radiolabeled estradiol from the receptor. The lowest point on the fitted response curve, within the range of data, will therefore be less than 50% and a log IC50 can be obtained. A positive result should be obtained in at least 2 out of 3 independent test runs. Chemicals with limited solubility may be problematic in this assay if some binding is seen at high concentrations. The maximum concentration of chemical to be used in the assay is 1mM. The guideline provides detailed guidance on classification of a chemical as "interactive", "equivocal", "not interactive" or "equivocal up to the limit of concentrations tested". It is important that quality and proficiency criteria are demonstrated for both positive and negative results.

C.2.1.8 Equivocal results for the guideline are not included in the table because these data require further interrogation about the result itself. In the event of an equivocal result, the considerations mentioned above about control quality and proficiency criteria should be taken into account and further investigations made. Equivocal results at high concentrations may result from solubility issues.

Existing Data to be Considered

C.2.1.9 Existing "Mechanism" *in vitro* data are assumed to be available from AR based assays (level 2) and the steroidogenesis assay. Assays may also be available for interference with thyroid modalities. The ER binding assay is most likely to be performed before the ER STTA (OECD TG 455) assay and so the ability of the chemical to affect ER-mediated gene expression may not be

known. In practice, it is possible that data from some or all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. The ER binding assay and ER STTA both provide data about the intrinsic ability of a chemical to interact with ER but each has their own advantages and disadvantages. The ER binding assay will not distinguish between agonists and antagonists whilst some chemicals testing positive in the ER STTA assay may have affected the reporter gene activity through non-ER related mechanisms. Consistent results in both assays give more confidence in the presence or absence of an ER-related mode of action.

C.2.1.10 Existing "Effects" data refer to *in vivo* effects "of concern" *i.e.* data from level 4 or 5 mammalian or wildlife assays. These may come from varied sources and will depend upon the type of substance (*e.g.* new chemicals, HPV chemical, pesticide). Thus, available data may range from repeated dose toxicity studies (28-day, 90-day), combined repeat dose/reproductive screening assays or fish screening assays, to chronic toxicity studies and multi-generation reproductive tests in mammalian or wildlife species. Some studies fail to identify EDs that weakly affect oestrogen or androgen receptors as was demonstrated on the basis of data generated in the validation process of OECD TG 407 assay with endocrine endpoints. In this validation only moderate EDs, such as nonylphenol and DDE, and strong EDs such as ethinylestradiol and flutamide (acting via ER and AR respectively) were detected. Thus OECD TG 407 cannot be regarded as a screening assay for endocrine activity. This means that when a relatively insensitive test is positive for both endocrine-specific and apical endpoints, this should be taken as an indication that the substance is a possible ED. Caution should be exercised, however, because endocrine endpoints may be impacted secondary to non-endocrine toxicity (*e.g.* hepatotoxicity) and *in vivo* apical endpoints can be affected by all modes of action including endocrine modalities. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.2.1.11 Data may also be available from level 3 mammalian assays (H and UT assays) but as the UT assay primarily detects (*in vivo*) the same modality as ER binding it is unlikely that it would be conducted before ER binding. An AMA may also be available but as this test primarily detects thyroid disruption in amphibians it is unlikely to provide useful data for E-modalities.

C.2.1.12 When considering the results of the ER binding assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.2.1.13 The scenarios (A to R) presented in Table C.2.1 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.2.1.14 Scenarios A to C represent positive results in the ER binding assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. A positive result in an ER binding assay is strong evidence for (anti)estrogenic activity that may or may not be supported by the *in vivo* effects data. In the case of positive *in vivo* effects data there may be sufficient evidence to conclude concern for endocrine disruption and therefore no need for further screening. Negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. A conclusion of lack of concern for endocrine

disruption in the presence of positive effects data can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.1.15 Scenarios D to F represent positive results in the ER binding assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Unless the metabolic profile of the test substance is known, the first option should be to conduct these *in vitro* assays with an added metabolising system. If the metabolic profile is known, then an *in vivo* test may be advisable. The choice of tests will depend upon the available *in vivo* effects data. As in scenarios A to C, negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. A conclusion of lack of concern for endocrine disruption in the presence of positive effects data (Scenario E) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.1.16 Scenarios G to I represent positive results in the ER binding assay in the presence of various combinations of missing or equivocal data. The next step to take in these eventualities will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step. In some cases it may be necessary to conduct *in vivo* tests and some guidance is given in the final column. As above, a conclusion of lack of concern for endocrine disruption in the presence of some positive effects data (Scenario H) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.1.17 Scenarios J to L represent negative results in the ER binding assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. The limitations of the ER binding assay should be considered first (*e.g.* lack of metabolic activation, possible involvement of other binding proteins). The positive *in vitro* mechanistic data indicates possible alternative ATS mechanisms. To confirm lack of ER-related activity in the presence of *in vivo* data, an ER STTA could be performed. Otherwise *in vivo* tests will confirm or refute EATS activity.

C.2.1.18 Scenarios M to O represent negative results in the ER binding assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend upon the weight of evidence and may not be possible in some cases. However, in the presence of negative data from robust level 4 and 5 assays further animals testing is probably not justified. The limitations of the ER binding assay should also be considered (as described for scenarios J to L). To confirm lack of ER-related activity in the presence of *in vivo* data, an ER STTA could be performed. Otherwise *in vivo* tests will confirm or refute EATS activity (Scenarios M and O).

C.2.1.19 Scenarios P to R represent negative results in the ER binding assay in the presence of various combinations of missing or equivocal data. The limitations of the ER binding assay should be considered first (as described for scenarios J to L). As with the positive result scenarios above (paragraph C.2.1.15) the next step to take for Scenarios P to R will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.2.1.20 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.2.1 is meant to provide a succinct guide and may not cover all circumstances or possibilities. In general, a decision about whether or not to conduct *in vivo* mammalian or wildlife tests will depend on the weight of evidence of new and existing data. If most available data (*e.g*. the results of the ER binding assay, results from an ER transcription activation assay, predictions from QSARs, 'read-across' from data on similar substances, and results from mammalian *in vivo* assays) suggest that the substance has the potential to cause endocrine disruption via the estrogen receptor (*i.e*. the level of suspicion about endocrine disrupting action is high – corresponding to Scenario A), then consideration should be given to the conduct of a higher level test.

C.2.1.21 For wildlife species, higher level tests with fish (*i.e* TG 234, the FLCTT or the MMGT) are recommended. Choice about which of these tests is most appropriate will be driven *inter alia* by mode of action considerations, and by whether multi-generation effects are to be expected. Such tests are unlikely to be needed if exposure of the natural environment is not expected. On the other hand, if available data only raise a low or moderate level of suspicion about endocrine disrupting action (*e.g*. the data appear to conflict with each other), then consideration should be given to the conduct of a fish screen (*i.e*. OECD TG 229 or 230).

C.2.1.22 For mammals, similar considerations apply but lower level tests (*e.g.* Level 3 or 4) should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study.

Table C.2.1. ER binding Assay (US EPA OPPTS 890.1250). Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from AR based assays and the steroidogenesis assay (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". Data from the ER STTA are assumed to be unavailable but a decision about the next step to be taken will also depend upon the availability of this assay.

Existing Results: ** "Effects (in vivo effects of concern)" assumes various information, such as data from repeat oral toxicity studies, reproduction/developmental toxicity screening tests, read across from analogues, will be available.

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	Interaction with ER combined with effects on AR/T/S and potential for adverse effects via multiple mechanisms.	Perform ER STTA or assay from upper levels *e.g.* UT assay (level 3) or female PP assay (level 4) or ext-1 or 2-gen assays or partial/full fish life cycle tests (level 4/5).	If existing data are from level 5 there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism , although some Level 4 assays (*e.g.* TG 234) may be sufficient for this

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results and *in vivo.* results but may also be metabolised to a metabolite that also has positive results *in vitro* and *in vivo*.
B	+	+	-	Interaction with ER combined with effects on AR/T/S but effects not detected in *in vivo* studies. Interaction with ER does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform binding assay or ER STTA with added metabolising system or assay from upper levels *e.g.* UT assay or fish screen (OECD TG 229/230) (level 3) or female PP assay (level 4)	If existing data are from an adequate level 5 study there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234) may be sufficient for this

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species.
C	+	+	Eq/0	Interaction with ER combined with effects on AR/T/S but no or equivocal data from *in vivo* studies Interaction with ER may not result in adverse effects.	Perform ER STTA or Perform assay from upper levels *e.g.* UT assay or fish screen (OECD TG 229/230) (level 3) or female PP assay (level 4)	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
D	+	-	+	Interaction with ER and potential for adverse effects. .	Perform ER STTA or	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					Perform assay from upper levels (*e.g*. UT assay or fish screen OECD TG 229/230) (level 3) or female PP assay (level 4)	concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g*. TG 234) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
E	+	-	-	Interaction with ER but effects not detected in *in vivo* studies. Interaction with ER does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform binding assay or ER STTA with added metabolising system or assay from upper levels *e.g.* UT assay or fish	If existing data are from an adequate level 5 assay there sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					screen (OECD TG 229/230) (level 3) or female PP assay (level 4).	comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species. Check data on chemical analogues.
F	+	-	Eq/0	Interaction with ER but no or equivocal data from *in vivo* studies.	Perform ER STTA or perform assay from upper levels *e.g.* UT assay or fish screen (OECD TG 229/230)	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					(level 3) or female PP assay (level 4).	to an inactive metabolite *in vivo*. Check data on chemical analogues.
G	+	Eq/0	+	Interaction with ER and potential for adverse effects via ER. May act via EATS mechanism and may or may not require metabolic activation.	Perform ER STTA	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Check data on chemical analogues. Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
H	+	Eq/0	**-**	Interaction with ER but effects not detected in *in vivo* studies. Interaction with ER does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	For the "0" scenario, perform ER STTA. For the "Eq" scenario perform ER STTA.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						different between wildlife species. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	Interaction with ER with unknown potential for effects in *in vivo* studies. May act via ER and may or may not require metabolic activation. Unknown potential for adverse effects.	For the "0" scenario, ER STTA with added metabolising system. For the "Eq" scenario UT assay or fish screen (OECD TG 229/230) (level 3) if existing data indicates this is needed	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. Check data on chemical analogues.
J	-	+	+	No evidence for interaction with ER. Effects on AR/T/S and potential for adverse effects via EATS mechanisms.	Perform ER binding assay or ER STTA with added metabolising system. or Perform assay from upper levels (*e.g*. UT assay or fish screen	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption. If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism,

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					OECD TG 229/230) (level 3) or female PP assay (level 4)	although some Level 4 assays (*e.g.* TG 234) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues. Further mechanistic studies may help determine MoA.
K	-	+	-	No evidence for interaction with ER. Effects on AR/T/S but effects not detected in *in vivo* studies. Metabolic differences explain *in vitro*/*in vivo* ATS differences.	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female PP assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose.

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						If existing data are from UT assay then level 4 assay will provide data on multiple modalities Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* ATS activity is not realised. Consider possible routes of exposure, implications of metabolism.
L	-	+	Eq/0	No evidence for interaction with ER. Effects on AR/T/S but effects not detected in *in vivo* studies. Unknown potential for adverse effects. Metabolic differences explain *in vitro*/*in vivo* ATS differences.	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3), or male or female PP assay (level 4)	Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* ATS activity is not realised. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.
M	-	-	+	No evidence for interaction with ER. Metabolic differences explain *in vitro*/*in vivo* differences. Effects seen in existing studies are via non-EATS or non-endocrine mechanism.	Perform ER STTA with added metabolising system or Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female PP assay (level	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption. If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism,

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					4).	although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Metabolic activation of chemical may occur *in vivo*. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
N	-	-	-	No evidence for interaction with ER. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female PP assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Check data on chemical analogues.
O	**-**	**-**	Eq/0	No evidence for interaction with ER. Metabolic differences explain *in vitro*/*in vivo* differences. Unknown potential for adverse effects via other mechanisms.	Perform ER STTA with added metabolising system or fish screen (OECD TG 229/230) (level 3) or male or female PP assay (level 4) if existing data indicates this is needed.	Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.
P	-	Eq/0	+	No evidence for interaction with ER. Metabolic differences explain *in vitro*/*in vivo* differences. Unknown potential for adverse effects via other mechanism.	Perform ER STTA with added metabolising system	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption. If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
Q	-	Eq/0	-	No evidence for interaction with ER. No evidence of adverse effects.	Perform ER STTA with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient

Scenarios	Result of ER binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
R	-	Eq/0	Eq/0	No evidence for interaction with ER. Unknown potential for adverse effects via other mechanism.	For the "0" scenario perform ER STTA with added metabolising system or Perform assay from upper levels *e.g.* UT assay or fish screen (OECD TG 229/230) (level 3) or female PP assay (level 4).	Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.

C.2.2 AR binding Assay (US EPA.OPPTS 890.1150)

C.2.2.1 Modality detected/endpoints: Binding to AR

Background to the Assay

C.2.2.2 The AR Binding Assay is an *in vitro* screening assay to detect substances that bind to AR. The assay has been in use for a number of years and there are different variations of the protocol. The most commonly used protocol utilises rat prostate cytosol as a source of AR without further purification. Human AR is now available as a recombinant protein and will replace the use of rat prostate cytosol when successful validation has been completed. The AR binding assay was chosen to be one of the suite of assays comprising US EPA's "Tier 1" and has been validated in that context (USEPA, 2007). There is no OECD test guideline for the assay but the US EPA (OPPTS) guideline is available (USEPA, 2009b). In this context, the assay provides information on the ability of a compound to interact with AR but is not intended to be used to show that the interaction is, specifically, one-site competitive binding, or to characterize precisely the strength of the binding. The assay determines the ability of a chemical to displace a radiolabeled ligand (R1881) from AR (in a rat ventral prostate tissue homogenate) and provides a positive or negative result for the ability to bind to AR.

C.2.2.3 Chemicals that bind to AR may induce hormone-dependent transcriptional activity (agonist) or block normal hormone function by preventing the endogenous hormone from binding to the receptor (antagonist). The binding assay does not distinguish between these. The AR ligand binding domain among vertebrate species is well conserved, so that substances that bind to AR derived from one species are expected to bind to the AR from other vertebrate species. The results from this assay are therefore relevant to many taxa. A positive result in guideline OPPTS 890.1150 requires demonstration of a concentration response curve for the ability of the test chemical to displace radiolabelled R1881. The concentration response curve allows the determination of potency *i.e.* IC50 (concentration at which 50% of radioligand is displaced by the test chemical) and relative binding affinity by comparing the log (IC50) of R1881with that of the test chemical.

C.2.2.4 Performance criteria are specified for the assay in order to demonstrate that the assay is functioning correctly. Proficiency chemicals are also used on each run to demonstrate the sensitivity of the experiment (reference standard: R1881 and weak positive control: dexamethasone). Compliance with the performance criteria should be checked before evaluating results from this assay. Small deviations are unlikely to have compromised the assay but judgement should be made on a case-by-case basis.

When/Why the Assay May be Used

C.2.2.5 Although the AR binding assay may be used at any stage in the hazard assessment process, the most likely use scenario is during initial assessment of chemicals for their ability to interact with endocrine systems *in vitro*, *i.e.* EATS modalities. Assays for interaction with other modalities *e.g.* ER and steroidogenesis interference, are likely to be conducted at the same time so that all results can be considered together. TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. The AR binding assay does not include the use of a xenobiotic metabolising system but consideration should be given to the inclusion of this (Jacobs *et al*, 2008; OECD, 2008a) depending upon the circumstances *e.g.* if the metabolism of a chemical is unknown, although the methods for inclusion of xenobiotic metabolising systems are not yet validated (see paragraph B.18). Alternatively, for a chemical with known metabolites, these could also be tested in the AR binding assay. Another use scenario may be following effects obtained in higher tier tests, for example delayed or accelerated puberty onset in males, which could be indicative of an effect

mediated by AR. Selection of the most appropriate tests has to be on a case-by-case basis but also considering the need to minimise animal testing. A further example could be results obtained in other apical assays, *e.g.* OECD TG 408 (90-day toxicity test), where effects on reproductive organs may be investigated further by testing in the AR binding assay in combination with ER and steroidogenesis based assays.

Introduction to the Table of Scenarios

C.2.2.6 Table C.2.2 gives guidance on a further step to take in the event of a positive (+) or negative (-) result in the AR binding assay and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.2.2.7 The results of the AR binding assay are given in the second column. Criteria for positive, negative and equivocal results are given in the OPPTS guideline. A result is judged positive if the lowest point on the fitted response curve, within the range of data, is less than 50%. This means that more than 50% of radiolabeled R1881 has been displaced from the receptor and a log IC50 can be obtained. A positive result should be obtained in at least 2 out of 3 independent test runs. Chemicals with limited solubility may be problematic in this assay if some binding is seen at high concentrations. The maximum concentration of chemical to be used in the assay is 1mM. The guideline provides detailed guidance on classification of a chemical as "binder", "equivocal", "non-binder", or "untestable" (does not reach 50% reduction in binding and is not soluble above 10^{-6} M). It is important that quality and proficiency criteria are demonstrated for both positive and negative results.

C.2.2.8 Equivocal results for the guideline are not included in the table because these data require further interrogation about the result itself. In the event of an equivocal result, the considerations mentioned above about control quality and proficiency criteria should be taken into account and further investigations made. Equivocal results at high concentrations may result from solubility issues.

Existing Data to be Considered

C.2.2.9 Existing "Mechanism" *in vitro* data are assumed to be available from ER based assays (level 2) and the steroidogenesis assay. Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from some or all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. The AR binding assay and AR transactivation assays both provide data about the intrinsic ability of a chemical to interact with AR but the binding assay will not distinguish between agonists and antagonists whilst some chemicals testing positive in the transactivation assays may have affected the reporter gene activity through non-AR related mechanisms. Consistent results in both assays give more confidence about the presence or absence of an AR-related mode of action.

C.2.2.10 Existing "Effects" data refer to *in vivo* effects "of concern" *i.e.* data from level 4 or 5 tests. These may come from varied sources and will depend upon the type of substance (*e.g.* new chemicals, HPV chemical, pesticide). Thus, available data may range from repeated dose toxicity studies (28-day, 90-day) or combined repeat dose/reproductive screening assays to chronic toxicity studies and multi-generation reproductive tests. Some studies fail to identify EDs that weakly affect oestrogen or androgen receptors as was demonstrated on the basis of data generated in the validation process of OECD TG 407 assay with endocrine endpoints. In this validation only moderate EDs, such as nonylphenol and DDE, and strong EDs such as ethinylestradiol and flutamide (acting via ER and AR respectively) were detected. Thus OECD TG 407 cannot be regarded as a screening assay for

endocrine activity. This means that when a relatively insensitive test is positive for both endocrine-specific and apical endpoints, this should be taken as an indication that the substance is a possible ED. Caution should be exercised, however, because endocrine endpoints may be impacted secondary to non-endocrine toxicity (*e.g.* hepatotoxicity) and *in vivo* apical endpoints can be affected by all modes of action including endocrine modalities. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.2.2.11 Data may also be available from level 3 tests (H and UT assays) but as the H assay primarily detects (*in vivo*) the same modality as AR binding it is unlikely that it would be conducted before AR binding. An AMA may also be available but as this test primarily detects thyroid disruption in amphibians it is unlikely to provide useful data for E-modalities.

C.2.2.12 When considering the results of the AR binding assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.2.2.13 The scenarios (A to R) presented in Table C.2.2 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.2.2.14 Scenarios A to C represent positive results in the AR binding assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. A positive result in an AR binding assay is strong evidence for (anti)androgenic activity that may or may not be supported by the *in vivo* effects data. In the case of positive *in vivo* effects data there may be sufficient evidence to conclude concern for endocrine disruption and therefore no need for further screening. Negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. A conclusion of lack of concern for endocrine disruption in the presence of positive effects data can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.2.15 Scenarios D to F represent positive results in the AR binding assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Unless the metabolic profile of the test substance is known, the first option should be to conduct these *in vitro* assays with an added metabolising system. If the metabolic profile is known, then an *in vivo* test may be advisable. The choice of tests will depend upon the available *in vivo* effects data. As in scenarios A to C, negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. A conclusion of lack of concern for endocrine disruption in the presence of positive effects data (Scenario E) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.2.16 Scenarios G to I represent positive results in the AR binding assay in the presence of various combinations of missing or equivocal data. The next step to take in these eventualities will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight

of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step. As above, a conclusion of lack of concern for endocrine disruption in the presence of some positive effects data (Scenario H) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.2.17 Scenarios J to L represent negative results in the AR binding assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. The limitations of the AR binding assay should be considered first (*e.g.* lack of metabolic activation, possible involvement of other binding proteins). The positive *in vitro* mechanistic data indicates possible alternative ETS mechanisms. To confirm lack of AR-related activity in the presence of *in vivo* data, an AR STTA could be performed. Otherwise *in vivo* tests will confirm or refute EATS activity.

C.2.2.18 Scenarios M to O represent negative results in the AR binding assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend upon the weight of evidence and may not be possible in some cases. However, in the presence of negative data from robust level 4 and 5 assays further animals testing is probably not justified. The limitations of the AR binding assay should also be considered (as described for scenarios J to L). To confirm lack of AR-related activity in the presence of *in vivo* data, an AR STTA could be performed. Otherwise *in vivo* tests will confirm or refute EATS activity (Scenarios M and O).

C.2.2.19 Scenarios P to R represent negative results in the AR binding assay in the presence of various combinations of missing or equivocal data. The limitations of the AR binding assay should be considered first (as described for scenarios J to L). As with the positive result scenarios above (paragraph C.2.2.16) the next step to take for Scenarios P to R will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.2.2. 20 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.2.2 is meant to provide a succinct guide and may not cover all circumstances or possibilities. In general, a decision about whether or not to conduct *in vivo* mammalian or wildlife tests will depend on the weight of evidence of new and existing data. If most available data (*e.g*. the results of the AR binding assay, results from an AR transcription activation assay, predictions from QSARs, 'read-across' from data on similar substances, and results from mammalian *in vivo* assays) suggest that the substance has the potential to cause endocrine disruption via the androgen receptor (*i.e*. the level of suspicion about endocrine disrupting action is high – corresponding to Scenario A), then consideration should be given to the conduct of a higher level test.

C.2.2.21 For wildlife species, higher level tests with fish (*i.e* TG 234 (FSDT), the FLCTT or the MMGT) are recommended. Choice about which of these tests is most appropriate will be driven *inter alia* by mode of action considerations, and by whether multi-generation effects are to be expected. Such tests are unlikely to be needed if exposure of the natural environment is not expected. On the other hand, if available data only raise a low or moderate level of suspicion about endocrine disrupting action (*e.g*. the data appear to conflict with each other), then consideration should be given to the conduct of a fish screen (*i.e*. OECD TG 229 or 230 or the AFSS).

C.2.2.22 For mammals, similar considerations apply but lower level tests (*e.g.* Level 3 or 4) should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study.

Table C.2.2. AR binding Assay (US EPA OPPTS 890.1150). Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER based assays and the steroidogenesis assay (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". Data from the AR STTA are assumed to be unavailable but a decision about the next step to be taken will depend upon the availability of this assay.

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes various information, such as data from repeat oral toxicity studies, reproduction/developmental toxicity screen tests, read across from analogues, will be available.

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	Interaction with AR combined with effects on ER/T/S and potential for adverse effects via multiple mechanisms.	Perform assay AR STTA or assay from upper levels *e.g.* H assay (level 3) or fish screen (AFSS) (level 3) or male PP assay (level 4) or ext-1 or 2-gen assays or partial/full fish life cycle tests	If existing data are from level 5 there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism,

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					(level 4/5).	although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results and *in vivo.* results but may also be metabolised to a metabolite that also has positive results *in vitro* and *in vivo*.
B	+	+	-	Interaction with AR combined with effects on AR/T/S but effects not detected in *in vivo* studies. Weak interaction with AR does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform binding assay or AR STTA with added metabolising system or assay from upper levels *e.g.* H assay or fish screen (OECD TG 229/230 or AFSS) (level 3)	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					or male PP assay (level 4).	of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species.
C	+	+	Eq/0	Interaction with AR combined with effects on ER/T/S but no or equivocal data from *in vivo* studies Weak interaction with AR may not result in adverse effects.	Perform assay from upper levels *e.g* H assay or fish screen (OECD TG 229/230 or AFSS) (level 3) or male PP assay	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					(level 4).	
D	+	**-**	+	Interaction with AR and potential for adverse effects.	Perform AR STTA or Perform assay from upper levels *e.g.* H assay or fish screen (OECD TG 229/230 or AFSS) (level 3) or male PP assay, (level 4)	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
E	+	**-**	**-**	Interaction with AR but effects not detected in *in vivo* studies. Weak interaction with AR does not result in adverse effects.	Perform binding assay or AR STTA with added metabolising	If existing data are from an adequate level 5 assay there sufficient information to conclude absence of concern for endocrine disruption

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				Metabolic differences explain *in vitro*/*in vivo* differences.	system or assay from upper levels *e.g.* H assay or fish screen (OECD TG 229/230 or AFSS) (level 3) or male PP assay (level 4).	(the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species.

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						Check data on chemical analogues.
F	+	-	Eq/0	Interaction with AR but no or equivocal data from *in vivo* studies. Weak interaction with AR does not result in adverse effects.	Perform AR STTA or perform assay from upper levels *e.g.* H assay or fish screen (OECD TG 229/230 or AFSS) (level 3) or male PP assay (level 4).	AR transactivation assay results will indicate whether AR binding affects transcription. Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
G	+	Eq/0	+	Interaction with AR and potential for adverse effects via AR or other ETS mechanisms. May act via EATS mechanism and may or may not require metabolic activation.	Perform AR STTA	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						or fish screen (OECD TG 229/230) will provide data on multiple modalities. Check data on chemical analogues. Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
H	+	Eq/0	**-**	Interaction with AR but effects not detected in *in vivo* studies. Weak interaction with AR does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	For the "0" scenario, perform AR STTA. For the "Eq" scenario perform AR STTA with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230)

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	Interaction with AR with unknown potential for effects in *in vivo* studies. May act via AR and may or may not require metabolic activation. Unknown potential for adverse effects.	For the "0" scenario, AR STTA with added metabolising system. For the "Eq" scenario H assay or fish screen (OECD TG 229/230 or AFSS) (level 3) if existing data indicates this is needed	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. Check data on chemical analogues.

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
J	-	+	+	No evidence for interaction with AR. Effects on ER/T/S and potential for adverse effects via EATS mechanisms.	Perform AR binding assay or AR transactivation assay with added metabolising system. or Perform assay from upper levels (*e.g*. UT assay or fish screen OECD TG 229/230 or OECD TG 231) (level 3) or male PP assay (level 4)	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption. If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g*. TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues. Further mechanistic studies may help determine MoA.
K	-	+	-	No evidence for interaction with AR. Effects on ER/T/S but effects not detected in	Perform assay from upper levels	If existing data are from an adequate level 5 assay there may be sufficient

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				in vivo studies. Metabolic differences explain *in vitro*/*in vivo* ATS differences.	*e.g.* fish screen (OECD TG 229/230) (level 3) or male or female PP assay (level 4).	information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* ETS activity is not realised. Consider possible routes of exposure, implications of metabolism.
L	**-**	+	Eq/0	No evidence for interaction with AR. Effects on ER/T/S but effects not detected in *in vivo* studies. Unknown potential for adverse effects. Metabolic differences explain *in vitro*/*in vivo*	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level	Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* ETS activity is not realised. Consider possible routes of exposure, implications of

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				ETS differences.	3), male or female PP (level 4).	metabolism. Check data on chemical analogues.
M	-	-	+	No evidence for interaction with AR. Metabolic differences or route of exposure explain *in vitro*/*in vivo* differences. Effects seen in existing studies are via non-EATS or non-endocrine mechanism.	Perform AR STTA with added metabolising system or Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female PP assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption. If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Metabolic activation of chemical may occur *in vivo*. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						Further mechanistic studies would help determine MoA.
N	-	-	-	No evidence for interaction with AR. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may perform assay from upper levels *e.g.* fish screen (AFSS) (level 3), male or female PP assay (level 4)	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Check data on chemical analogues.
O	-	-	Eq/0	No evidence for interaction with AR. Metabolic differences explain *in vitro*/*in vivo* differences. Unknown potential for adverse effects via other mechanisms.	Perform AR STTA with added metabolising system or	Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					fish screen (OECD TG 229/230) (level 3) or male or female PP assay (level 4)if existing data indicates this is needed.	
P	-	Eq/0	+	No evidence for interaction with AR. Metabolic differences explain *in vitro*/*in vivo* differences. Unknown potential for adverse effects via other mechanism.	Perform AR STTA with added metabolising system	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption. If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities.

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
Q	-	Eq/0	-	No evidence for interaction with AR. No evidence of adverse effects.	Perform AR STTA with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Check data on chemical analogues. Further mechanistic studies would

Scenarios	Result of AR binding assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						help determine MoA.
R	-	Eq/0	Eq/0	No evidence for interaction with AR. Unknown potential for adverse effects via other mechanism.	For the "0" scenario perform AR STTA with added metabolising system or Perform assay from upper levels *e.g.* H assay or fish screen (OECD TG 229/230) (level 3) or male PP assay (level 4).	Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.

C.2.3 OECD TG 455: The Stably Transfected Human ERα Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of Chemicals (ER STTA) (including guidance for the antagonism assay)

C.2.3.1 Modality detected/endpoints: Activation of reporter gene linked to ER (agonist assay). Inhibition of activation of reporter gene linked to ER (antagonist assay).

Note: antagonism is not addressed in OECD TG 455 as the antagonist assay is currently in validation.

Background to the Assay

C.2.3.2 The Stably Transfected hERα Transcriptional Activation Assay (ER STTA) is an *in vitro* screening assay to detect substances that bind to hERα and activate the transcription of estrogen responsive genes. It is an *in vitro* tool that provides mechanistic data. Several ER STTA assays in common use can be found in the literature (*e.g.* Andersen *et al*, 2001; Escande *et al*, 2006; Takeyoshi *et al*, 2002; Du *et al*, 2010). One of the first versions of this assay used was the "yeast estrogen screen" (Routledge and Sumpter, 1996); Odum *et al*, 1997; Sheahan *et al*, 2002) which is still widely used for screening of environmental samples. The assay described in OECD TG 455 utilises the hERα-HeLa-9903 cell line and a luciferase reporter gene. OECD TG 455 only addresses agonist interaction with hERα as this was the focus for validation of the assay. Published assays addressing antagonism are, however, readily available (*e.g.* Takeyoshi *et al*, 2002; Du *et al*, 2010). Other STTA assays are also being validated via OECD initiatives and these will include the antagonist assay (*e.g.* Witters *et al*, 2010). A generic performance based OECD TG, covering these assays, will replace OECD TG 455 once it has been developed and approved. This guidance therefore covers both agonism and antagonism.

C.2.3.3 OECD TG 455 provides a positive or negative result for the ability of a chemical to induce hERα-mediated transactivation of luciferase gene expression (agonist assay) compared to a vehicle control. The antagonist assay determines whether a reduction in response occurs when cells are co-exposed to chemical and a potent estrogen agonist compared to the potent estrogen agonist alone. A positive response for OECD TG 455 is the ability of a chemical to achieve an agonist response equal to 10% of that induced by the positive control 17β-estradiol *i.e.* the PC10, in at least two out of two or three runs of the assay. A measure of potency is also provided by the magnitude of the effect and the concentration at which it occurs. To be acceptable the results should also meet the performance standards given in the assay. Small deviations are unlikely to have compromised the assay but judgement should be made on a case-by-case basis.

C.2.3.4 The hERα-HeLa-9903 assay showed a high degree of sensitivity and specificity in the validation studies, when compared with the UT assay which determines the ability of a chemical to elicit an estrogenic response *in vivo*. OECD TG 455 requires strict control of assay conditions in order to maintain the accuracy and reliability of response. Demonstration of laboratory proficiency with ten proficiency chemicals is required at the outset, each experiment requires four reference chemicals (positive and negative chemicals) and each plate requires positive and vehicle controls. Criteria for the degree of response with these chemicals are given in the TG. The assay also requires a minimum of 80% cell viability Compliance with the quality control criteria and with the performance criteria should be demonstrated before evaluating results from this assay.

C.2.3.5 A limitation of the OECD TG 455, related to the reporter gene luciferase, is the potential for chemicals to increase chemiluminescence via non-ERα mechanisms thus possibly giving a false positive response. This has been reported for certain phytoestrogens such as genistein and daidzein but not for industrial chemicals (Kuiper *et al*, 1998; Escande *et al*, 2006). This may be recognized by incomplete or unusual dose response curves and can be tested by performing a specific antagonist

assay (provided as an annex to the OECD TG). Other ER STTAs that do not use luciferase as a reporter gene may not have this drawback (Escande *et al*, 2006).

C.2.3.6 The ER STTA will not detect substances that act by other mechanisms *e.g.* AR, TR and steroidogenesis interference. These chemicals will, however, be detected in AR, TR and steroidogenesis specific assays and therefore results from a suite of *in vitro* tests should be considered together. The assay will not detect substances that act by affecting the HPG as an *in vivo* intact axis is required for this.

When/Why the Assay May be Used

C.2.3.7 Although the ER STTA may be used at any stage in the hazard assessment process, the most likely use scenario is during initial assessment of chemicals for their ability to interact with endocrine systems *in vitro*, *i.e.* EATS modalities. The ER STTA is frequently conducted following a positive result in the ER binding assay. Assays for interaction with other modalities *e.g.* AR, ER and steroidogenesis, are likely to be conducted at the same time so that all results can be considered together. TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. The OECD TG 445 does not include the use of a xenobiotic metabolising system but consideration should be given to the inclusion of this (Jacobs *et al*, 2008; OECD, 2008a) depending upon the circumstances *e.g.* if the metabolism of a chemical is unknown, although the methods for inclusion of xenobiotic metabolising systems are not yet validated (see paragraph B.18). Alternatively, for a chemical with known metabolites, these could also be tested in the ER STTA.

C.2.3.8 Another use scenario may be following effects obtained in higher tier tests, for example accelerated puberty onset in females, but which are not exclusively indicative of an effect on ER. Selection of the most appropriate tests has to be on a case-by-case basis but also considering the need to minimise animal testing. A further example could be results obtained in other apical assays, *e.g.* OECD TG 408 (90-day toxicity test); effects on reproductive organs may be investigated further by testing in the ER STTA in combination with AR and steroidogenesis based assays.

Introduction to the Table of Scenarios

C.2.3.9 Table C.2.3 gives guidance on a further step to take in the event of a positive (+) or negative (-) result in the ER STTA and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. “Existing results” are subdivided into “Mechanism” and “Effects” data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.2.3.10 The results of the ER STTA are given in the second column. Criteria for positive results in OECD TG 455 are given in the test guideline. A result is judged positive if the maximum response induced by the test chemical is equal to or exceeds 10% of the response of the positive control in at least 2 of 3 test runs. It is important that quality and proficiency criteria are demonstrated for both positive and negative results. At present, there are no criteria for positive results in the antagonism assay as the OECD TG only covers agonists. For the purposes of this guidance, a positive antagonist response would be a statistically significant reduction in the agonist-stimulated response compared to the agonist-stimulated control value.

C.2.3.11 Equivocal results for the guideline are not included in the table because these data require further interrogation about the result itself. This assay is a screen and therefore a clear positive or negative result should be obtained. In the event of an equivocal result, the considerations mentioned above about control quality and proficiency criteria should be taken into account and further investigations made.

Existing Data to be Considered

C.2.3.12 Existing "Mechanism" *in vitro* data are assumed to be available from AR and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform.

C.2.3.13 Existing "Effects" data refer to *in vivo* effects "of concern" *i.e.* data from level 4 or 5 mammalian or wildlife assays. These may come from varied sources and will depend upon the type of substance (*e.g.* new chemicals, HPV chemical, pesticide). Thus, available data may range from repeated dose toxicity studies (28-day, 90-day), combined repeat dose/reproductive screening assays or fish screening assays, to chronic toxicity studies and multi-generation reproductive tests in mammalian or wildlife species. Some studies fail to identify EDs that weakly affect oestrogen or androgen receptors as was demonstrated on the basis of data generated in the validation process of OECD TG 407 assay with endocrine endpoints. In this validation only moderate EDs, such as nonylphenol and DDE, and strong EDs such as ethinylestradiol and flutamide (acting via ER and AR respectively) were detected. Thus OECD TG 407 cannot be regarded as a screening assay for endocrine activity. This means that when a relatively insensitive test is positive for both endocrine-specific and apical endpoints, this should be taken as an indication that the substance is a possible ED. Caution should be exercised, however, because endocrine endpoints may be impacted secondary to non-endocrine toxicity (*e.g.* hepatotoxicity) and *in vivo* apical endpoints can be affected by all modes of action including endocrine modalities. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.2.3.14 Data may also be available from level 3 tests (H and UT assays) but as the UT assay primarily detects (*in vivo*) the same modality as the ER STTA it is unlikely that it would be conducted prior to this. An AMA may also be available but as this test primarily detects thyroid disruption in amphibians it is unlikely to provide useful data for E-modalities.

C.2.3.15 When considering the results of the ER STTA, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.2.3.16 The scenarios (A to R) presented in Table C.2.3 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.2.3.17 Scenarios A to C represent positive results in the ER STTA assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. A positive result in an ER STTA assay is strong evidence for (anti)estrogenic activity that may or may not be supported by the *in vivo* effects data. In the case of positive *in vivo* effects data there may be sufficient evidence to conclude concern for endocrine disruption and therefore no need for further screening. Negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not

present a concern for endocrine disruption. A conclusion of lack of concern for endocrine disruption in the presence of positive effects data can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.3.18 Scenarios D to F represent positive results in the ER STTA assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Unless the metabolic profile of the test substance is known, the first option should be to conduct these *in vitro* assays with an added metabolising system. If the metabolic profile is known, then an *in vivo* test may be advisable. The choice of tests will depend upon the available *in vivo* effects data. As in scenarios A to C, negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. A conclusion of lack of concern for endocrine disruption in the presence of positive effects data (Scenario E) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.3.19 Scenarios G to I represent positive results in the ER STTA assay in the presence of various combinations of missing or equivocal data. The next step to take in these eventualities will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects.. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step. As above, a conclusion of lack of concern for endocrine disruption in the presence of some positive effects data (Scenario H) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.3.20 Scenarios J to L represent negative results in the ER STTA assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. The limitations of the ER STTA assay should be considered first (*e.g.* lack of metabolic activation, possible involvement of other factors). The positive *in vitro* mechanistic data indicates possible alternative ATS mechanisms. To confirm lack of ER-related activity in the presence of *in vivo* data, an ER STTA with added metabolising capability could be performed. Otherwise *in vivo* tests will confirm or refute EATS activity.

C.2.3.21 Scenarios M to O represent negative results in the ER STTA assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend upon the weight of evidence and may not be possible in some cases. However, in the presence of negative data from robust level 4 and 5 assays further animals testing is probably not justified. The limitations of the ER STTA assay should also be considered (as described for scenarios J to L). To confirm lack of ER-related activity in the presence of *in vivo* data, an ER STTA with added metabolising capability could be performed. Otherwise *in vivo* tests will confirm or refute EATS activity (Scenarios M and O).

C.2.3.22 Scenarios P to R represent negative results in the ER STTA assay in the presence of various combinations of missing or equivocal data. The limitations of the ER STTA binding assay should be considered first (as described for scenarios J to L). As with the positive result scenarios above (paragraph C.2.3.19) the next step to take for Scenarios P to R will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.2.3.23 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.2.3 is meant to provide a succinct guide and may not cover all circumstances or possibilities. In general, a decision about whether or not to conduct *in vivo* mammalian or wildlife tests will depend on the weight of evidence of new and existing data. If most available data (*e.g*. the results of the ER transcription activation assay, predictions from QSARs, 'read-across' from data on similar substances, and results from mammalian *in vivo* assays) suggest that the substance has the potential to cause endocrine disruption via the estrogen receptor (*i.e*. the level of suspicion about endocrine disrupting action is high – corresponding to Scenario A), then consideration should be given to the conduct of a higher level test.

C.2.3.24 For wildlife species, higher level tests with fish (*i.e* the TG 234 (FSDT), the FLCTT or the MMGT) are recommended. Choice about which of these tests is most appropriate will be driven *inter alia* by mode of action considerations, and by whether multi-generation effects are to be expected. Such tests are unlikely to be needed if exposure of the natural environment is not expected. On the other hand, if available data only raise a low or moderate level of suspicion about endocrine disrupting action (*e.g*. the data appear to conflict with each other), then consideration should be given to the conduct of a fish screen (*i.e*. OECD TG 229 or 230).

C.2.3.25 For mammals, similar considerations apply but lower level tests (*e.g.* Level 3 or 4) should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study.

Table C.2.3 OECD TG 455: The Stably Transfected Human ERα Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of Chemicals (ER STTA) (including guidance for the antagonist assay). Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and S based assays (level 2). The ER binding assay is likely to be performed prior to the ER STTA. TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes various information, such as data from repeat oral toxicity studies, reproduction/developmental toxicity screen tests, read across from analogues, will be available.

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	ER (ant)agonism combined with effects on AR/T/S and potential for adverse effects via multiple mechanisms.	Perform assay from upper levels *e.g.* UT assay (level 3) or female PP assay (level 4) or ext-1 or 2-gen assays or partial/full fish life cycle tests	If existing data are from level 5 there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					(level 4/5).	effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results and *in vivo.* results but may also be metabolised to a metabolite that also has positive results *in vitro* and *in vivo*.
B	+	+	-	ER (ant)agonism combined with effects on AR/T/S but effects not detected in *in vivo* studies. Weak ER (ant)agonism does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform ER STTA with added metabolising system or assay from upper levels *e.g.* UT assay or fish screen (OECD TG 229/230) (level 3) or	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					female PP assay (level 4).	endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species.
C	+	+	Eq/0	ER (ant)agonism combined with effects on AR/T/S but no or equivocal data from *in vivo* studies Weak ER (ant)agonism may not result in adverse effects.	Perform assay from upper levels *e.g.* UT assay or fish screen (OECD TG 229/230) (level 3) or female PP assay (level 4).	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
D	+	-	+	ER (ant)agonism and potential for adverse effects.	Perform assay from upper levels *e.g.* UT assay or fish screen (OECD TG 229/230) (level 3) or female PP assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
E	+	-	-	ER (ant)agonism but effects not detected in *in vivo* studies. Weak ER (ant)agonism does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform ER STTA with added metabolising system or assay from upper levels *e.g.* UT	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					assay or fish screen (OECD TG 229/230) (level 3) or female PP assay (level 4).	mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species. Check data on chemical analogues.
F	+	-	Eq/0	ER (ant)agonism but no or equivocal data from *in vivo* studies. Weak ER (ant)agonism does not result in adverse effects.	Perform assay from upper levels *e.g.* UT assay or fish screen	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical.

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					(OECD TG 229/230) (level 3) or female PP assay (level 4).	Check data on chemical analogues.
G	+	Eq/0	+	ER (ant)agonism and potential for adverse effects via ER (ant)agonism or other ATS mechanisms. May act via EATS mechanism and may or may not require metabolic activation.	Perform assay from upper levels *e.g.* UT assay or fish screen (OECD TG 229/230) (level 3) or female PP assay (level 4)	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Check data on chemical analogues. Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
H	+	Eq/0	**-**	ER (ant)agonism but effects not detected in *in vivo* studies. Weak ER (ant)agonism does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform ER STTA with added metabolising system	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						metabolism of chemicals can be different between wildlife species. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	ER (ant)agonism with unknown potential for effects in *in vivo* studies. May act via ER mechanism and may or may not require metabolic activation. Unknown potential for adverse effects.	Perform ER STTA with added metabolising system or UT assay or fish screen (OECD TG 229/230) (level 3), if existing data indicates this is needed	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
J	-	+	+	No evidence for ER (ant)agonism. Effects on AR/T/S and potential for adverse effects via EATS mechanisms.	Perform ER STTA with added metabolising system. or Perform assay from upper levels (*e.g*. UT assay or fish screen OECD TG 229/230 (level 3) or female PP assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption . If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g*. TG 234 (FSDT)) may be sufficient

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues. Further mechanistic studies may help determine MoA.
K	-	+	-	No evidence for ER (ant)agonism. Effects on AR/T/S but effects not detected in *in vivo* studies.	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female PP assay(level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						then level 4 assay will provide data on multiple modalities. Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* EATS activity is not realised. Consider possible routes of exposure, implications of metabolism.
L	-	+	Eq/0	No evidence for ER (ant)agonism. Effects on AR/T/S but effects not detected in *in vivo* studies. Unknown potential for adverse effects. Metabolic differences explain *in vitro*/*in vivo* ATS differences.	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female PP assay(level 4).	Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* EATS activity is not realised. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.
M	-	-	+	No evidence for ER (ant)agonism. Metabolic differences or route of exposure explain *in vitro*/*in vivo* differences. Effects seen in existing studies are via non-EATS or non-endocrine mechanism.	Perform ER STTA with added metabolising system or Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female PP assay (level 4) .	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption. If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose.

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Metabolic activation of chemical may occur *in vivo*. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
N	-	-	-	No evidence for ER (ant)agonism. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female PP assay (level 4)	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						then level 4 assay will provide data on multiple modalities. Check data on chemical analogues.
O	-	-	Eq/0	No evidence for ER (ant)agonism. Unknown potential for adverse effects via other mechanisms.	Perform ER STTA with added metabolising system or fish screen (OECD TG 229/230) (level 3) or male or female PP assay (level 4) if existing data indicates this is needed.	Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.
P	-	Eq/0	+	No evidence for ER (ant)agonism. Unknown potential for adverse effects via other mechanism.	Perform ER STTA with added metabolising system	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption. If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.*

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						TG 234 (FSDT) may be sufficient for this purpose. If existing data are from UT assay then level 4 assay will provide data on multiple modalities. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
Q	-	Eq/0	-	No evidence for ER (ant)agonism. No evidence of adverse effects.	Perform ER STTA with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assays then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. If existing data are from UT assay

Scenarios	Result of OECD TG 455 (ER STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						then level 4 assay will provide data on multiple modalities. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
R	-	Eq/0	Eq/0	No evidence for ER (ant)agonism. Unknown potential for adverse effects via other mechanism.	For the "0" scenario perform ER STTA with added metabolising system or perform UT assay or fish screen (OECD TG 229/230) (level 3), if existing data indicates this is needed.	Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.

C.2.4 OECD TG 456: H295R Steroidogenesis Assay

C.2.4.1 Modality detected/endpoints: Interference with steroidogenesis/Inhibition and induction of estradiol and testosterone synthesis.

Background to the Assay

C.2.4.2 The H295R Steroidogenesis Assay is an *in vitro* screening assay to detect substances that affect production of estradiol and testosterone. It provides a positive or negative result for the ability of a chemical to induce or inhibit the production of estradiol and testosterone. The assay utilises a human adrenocarcinoma cell line (NCI-H295R cells) that have the characteristics of undifferentiated human fetal adrenal cells. This cell line expresses all the key enzymes involved in steroidogenesis, from cholesterol to estradiol and testosterone. This expression would allow for the detection of other hormones but the assay was only validated for estradiol and testosterone. The cells represent a unique *in vitro* system because *in vivo*, expression of these enzymes is developmental stage specific with no one tissue expressing all the enzymes at once.

C.2.4.3 Chemicals may induce steroidogenesis; this can be determined by increased production of estradiol and testosterone. Alternatively, chemicals may inhibit steroidogenesis; this can be determined by decreased production of estradiol and testosterone. Results are expressed as fold change compared with the negative control. In the validation of the assay, forskolin induced estradiol and testosterone production whilst prochloraz inhibited estradiol and testosterone production. The validation of the steroidogenesis assay demonstrated that whilst not always directly predictive of a specific type response *in vivo*, the chemicals chosen in the validation studies would always be flagged as a disrupter of steroidogenesis or a reproductive toxicant (OECD, 2010e). The assay is therefore used somewhat as a "black box" where a positive result indicates that a chemical is a possible disrupter of steroidogenesis but without defining the exact mechanism of action.

C.2.4.4 An adequate response with positive control chemicals (forskolin and prochloraz), and other proficiency chemicals, is required in the OECD TG to demonstrate laboratory proficiency. The assay also requires the assessment of the cytotoxic effect of a chemical, as measurement of cell viability is an important feature of the TG. A minimum of 80% cell viability is needed for the hormone production assessment to be considered adequate. Limitations of the assay are that xenobiotic metabolising capability is unknown but likely to be limited and production of other hormones (*e.g.* gluco-and mineralocorticoids) by the cells may affect estradiol and testosterone levels. The current assay does not detect 5-alpha-reductase inhibitors (*e.g.* finasteride) that inhibit the conversion of testosterone to dihydrotestosterone. Although 5-alpha reductase is present in H295R cells, dihydrotestosterone is not a validated endpoint and therefore these chemicals will not be identified. 5-Alpha-reductase inhibitors are detected by OECD TG 441 (H assay).

C.2.4.5 The assay will not detect substances that act by affecting the HPG as an *in vivo* intact axis is required for this. The effect of AR, ER and TR ligands on this assay is also not clear, although the steroidogenesis assay is not designed to detect these substances, it is not known whether they affect steroidogenesis. These chemicals will, however, be detected in AR, ER and TR specific assays and therefore results from a suite of *in vitro* tests should be considered together.

C.2.4.6 The steroidogenesis assay requires that strict control is made of the age at which the cells are used. The capacity of the cells to produce estradiol changes with increasing number of cell passages. In addition, chemicals and cell matrices may interfere with hormone measurements. The TG includes quality control measures to ensure the accuracy and reliability of results. Compliance with the quality control criteria and with the performance criteria for the positive control substances forskolin and prochloraz and with the other proficiency chemicals should be demonstrated before evaluating results

from this assay. Small deviations are unlikely to have compromised the assay but judgement should be made on a case-by-case basis.

When/Why the Assay May be Used

C.2.4.7 Although the steroidogenesis assay may be used at any stage in the hazard assessment process, the most likely use scenario is during initial assessment of chemicals for their ability to interact with endocrine systems *in vitro*, *i.e.* EATS modalities. Assays for interaction with other modalities *e.g.* AR and ER, are likely to be conducted at the same time so that all results can be considered together. TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. Data from the aromatase assay may also be available, chemicals testing positive in this assay are likely to also give positive results in the steroidogenesis assay as aromatase is one of the key enzymes in the steroidogenesis pathway. The steroidogenesis TG does not include the use of a xenobiotic metabolising system but consideration should be given to the inclusion of this (Jacobs *et al*, 2008; OECD, 2008a) depending upon the circumstances *e.g.* if the metabolism of a chemical is unknown, although the methods for inclusion of xenobiotic metabolising systems are not yet validated (see paragraph B.18). Alternatively, for a chemical with known metabolites, these could also be tested in the steroidogenesis assay. Another use scenario may be following effects obtained in higher tier tests, for example delayed puberty onset in females, but which are not exclusively indicative of an effect on ER. Selection of the most appropriate tests has to be on a case-by-case basis but also considering the need to minimise animal testing. A further example could be results obtained in other apical assays, *e.g.* OECD TG 408 (90-day toxicity test), where effects on reproductive organs may be investigated further by testing in the steroidogenesis assay in combination with AR and ER based assays.

Introduction to the Table of Scenarios

C.2.4.8 Table C.2.4 gives guidance on a further step to take in the event of a positive (+) or negative (-) result in the steroidogenesis assay and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. “Existing results” are subdivided into “Mechanism” and “Effects” data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.2.4.9 The results of the steroidogenesis assay are given in the second column. Criteria for positive results are given in the draft test guideline. A result is judged positive if the fold difference is statistically significant from the solvent control at two adjacent concentrations in at least 2 tests, or when a single concentration data point is significantly different from the solvent control, and this can be confirmed by being significantly different in at least one more run within a +/- 1 concentration increment of the respective experiment. The latter allows for effects that may be seen close to the maximum concentration (1mM). It is important that quality and proficiency criteria are demonstrated for both positive and negative results.

C.2.4.10 Equivocal results for the guideline are not included in the table because these data require further interrogation about the result itself. This assay is a screen and therefore a clear positive or negative result should be obtained. In the event of an equivocal result, the considerations mentioned above about control quality and proficiency criteria should be taken into account and further investigations made.

Existing Data to be Considered

C.2.4.11 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR based assays and the aromatase assay (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform.

C.2.4.12 Existing "Effects" data refer to *in vivo* effects "of concern" *i.e.* data from level 4 or 5 mammalian or wildlife assays. These may come from varied sources and will depend upon the type of substance (*e.g.* new chemicals, HPV chemical, pesticide). Thus, available data may range from repeated dose toxicity studies (28-day, 90-day), combined repeat dose/reproductive screening assays or fish screening assays, to chronic toxicity studies and multi-generation reproductive tests in mammalian or wildlife species. Some studies fail to identify EDs that weakly affect oestrogen or androgen receptors as was demonstrated on the basis of data generated in the validation process of OECD TG 407 assay with endocrine endpoints. In this validation only moderate EDs, such as nonylphenol and DDE, and strong EDs such as ethinylestradiol and flutamide (acting via ER and AR respectively) were detected. The aromatase inhibitor CGS 18320B was detected by the OECD TG 407 assay, but this chemical was developed as a pharmaceutical aromatase inhibitor and therefore is a strong ED. The ability to detect chemicals that weakly interfere with steroidogenesis is not known. Thus OECD TG 407 cannot be regarded as a screening assay for endocrine activity. This means that when a relatively insensitive test is positive for both endocrine-specific and apical endpoints, this should be taken as an indication that the substance is a possible ED. Caution should be exercised, however, because endocrine endpoints may be impacted secondary to non-endocrine toxicity (*e.g.* hepatotoxicity) and *in vivo* apical endpoints can be affected by all modes of action including endocrine modalities. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.2.4.13 Data may also be available from H and UT assays (level 3) but as these assays do not generally detect steroidogenesis interference they are only useful in these cases for purposes of elimination.

C.2.4.14 When considering the results of the steroidogenesis assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.2.4.15 The scenarios (A to R) presented in Table C.2.4 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.2.4.16 Scenarios A to C represent positive results in the steroidogenesis assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. A positive result in a steroidogenesis assay is strong evidence for disruption of steroidogenesis that may or may not be supported by the *in vivo* effects data. Inhibition of steroidogenesis (but not induction) could be followed up by a confirmatory aromatase assay if this is not already available. In the case of positive *in vivo* effects data there may be sufficient evidence to conclude concern for endocrine disruption and

therefore no need for further screening. Negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. A conclusion of lack of concern for endocrine disruption in the presence of positive effects data can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.4.17 Scenarios D to F represent positive results in the steroidogenesis assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. As above, inhibition of steroidogenesis could be followed up by a confirmatory aromatase assay if this is not already available. Unless the metabolic profile of the test substance is known, the first option should be to conduct these *in vitro* assays with an added metabolising system. If the metabolic profile is known, then an *in vivo* test may be advisable. The choice of tests will depend upon the available *in vivo* effects data. As in scenarios A to C, negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. A conclusion of lack of concern for endocrine disruption in the presence of positive effects data (Scenario E) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.4.18 Scenarios G to I represent positive results in the steroidogenesis assay in the presence of various combinations of missing or equivocal data. As above, inhibition of steroidogenesis could be followed up by a confirmatory aromatase assay if this is not already available. The next step to take for missing or equivocal data will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step. As above, a conclusion of lack of concern for endocrine disruption in the presence of some positive effects data (Scenario H) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.4.19 Scenarios J to L represent negative results in the steroidogenesis assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. The limitations of the steroidogenesis assay should be considered first (*e.g.* lack of metabolic activation, possible involvement of other factors). The positive *in vitro* mechanistic data indicates possible alternative EAT mechanisms. To confirm lack of steroidogenesis activity in the presence of *in vivo* data, a steroidogenesis with added metabolising capability could be performed. Otherwise *in vivo* tests will confirm or refute EATS activity.

C.2.4.20 Scenarios M to O represent negative results in the steroidogenesis assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend upon the weight of evidence and may not be possible in some cases. However, in the presence of negative data from robust level 4 and 5 assays further animals testing is probably not justified. The limitations of the steroidogenesis assay should also be considered (as described for scenarios J to L). To confirm lack of steroidogenesis-related activity in the presence of *in vivo* data, a steroidogenesis assay with added metabolising capability could be performed. Otherwise *in vivo* tests will confirm or refute EATS activity (Scenarios M and O).

C.2.4.21 Scenarios P to R represent negative results in the steroidogenesis assay in the presence of various combinations of missing or equivocal data. The limitations of the steroidogenesis assay

should be considered first (as described for scenarios J to L). As with the positive result scenarios above (paragraph C.2.4.18) the next step to take for Scenarios P to R will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.2.4.22 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.2.4 is meant to provide a succinct guide and may not cover all circumstances or possibilities. In general, a decision about whether or not to conduct *in vivo* mammalian or wildlife tests will depend on the weight of evidence of new and existing data. If most available data (*e.g*. the results of the steroidogenesis assay, predictions from QSARs, 'read-across' from data on similar substances, and results from mammalian *in vivo* assays) suggest that the substance has the potential to cause endocrine disruption via interference with steroidogenesis (*i.e*. the level of suspicion about endocrine disrupting action is high – corresponding to Scenario A), then consideration should be given to the conduct of a higher level test.

C.2.4.23 For wildlife species, higher level tests with fish (*i.e* TG 234 (FSDT), the FLCTT or the MMGT) are recommended. Choice about which of these tests is most appropriate will be driven *inter alia* by mode of action considerations, and by whether multi-generation effects are to be expected. Such tests are unlikely to be needed if exposure of the natural environment is not expected. On the other hand, if available data only raise a low or moderate level of suspicion about endocrine disrupting action (*e.g*. the data appear to conflict with each other), then consideration should be given to the conduct of a fish screen (*i.e*. OECD TG 229 or 230).

C.2.4.24 For mammals, similar considerations apply but lower level tests (*e.g.* Level 3 or 4) should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study.

Table C.2.4. OECD TG 456: H295R Steroidogenesis Assay. Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER and AR based assays (level 2). Data on aromatase inhibition may also be available. TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes various information, such as data from repeat oral toxicity studies, reproduction/developmental toxicity screen tests, read across from analogues, will be available.

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	Inhibition/induction of steroidogenesis combined with effects on ER/AR/T/S and potential for adverse effects via multiple mechanisms.	Perform assay from upper levels *e.g.* male or female pubertal assay (level 4) or ext-1 or 2-gen assays or partial/full fish life cycle tests (level 4/5).	If existing data are from level 5 there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.*

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						TG 234 (FSDT)) may be sufficient for this purpose. Compare steroidogenesis assay results with other *in vitro* results to help discern mechanism. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results and *in vivo.* results but may also be metabolised to a metabolite that also has positive results *in vitro* and *in vivo*.
B	+	+	-	Inhibition/induction of steroidogenesis combined with effects on ER/AR/T but effects not detected in *in vivo* studies. Weak steroidogenesis inhibition/induction does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform steroidogenesis assay with added metabolising system OR assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.*

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						TG 234 (FSDT)) may be sufficient for this purpose. Compare steroidogenesis assay results with other *in vitro* results to help discern mechanism. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species.
C	+	+	Eq/0	Inhibition/induction of steroidogenesis combined with effects on ER/AR/T but no or equivocal data from *in vivo* studies. Weak steroidogenesis inhibition/induction may not result in adverse effects.	Perform assay upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	Compare steroidogenesis assay results with other *in vitro* results to help discern mechanism. Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
D	+	**-**	+	Inhibition/induction of steroidogenesis and potential for adverse effects.	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3)	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					or male or female pubertal assay (level 4).	most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
E	+	-	-	Inhibition/induction of steroidogenesis but effects not detected in *in vivo* studies. Weak steroidogenesis inhibition/induction does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform steroidogenesis assay with added metabolising system OR assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					pubertal assay (level 4).	of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species. Check data on chemical analogues.
F	+	-	Eq/0	Inhibition/induction of steroidogenesis but no or equivocal data from *in vivo* studies. Weak steroidogenesis inhibition/induction may not result in adverse effects.	Perform assay upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
G	+	Eq/0	+	Inhibition/induction of steroidogenesis and potential for adverse effects via steroidogenesis interference or other EAT mechanisms. May act via non- steroidogenesis interference	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3)	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				mechanism and may or may not require metabolic activation.	or male or female pubertal assay (level 4).	most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Check data on chemical analogues. Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
H	+	Eq/0	**-**	Inhibition/induction of steroidogenesis but effects not detected in *in vivo* studies. Weak steroidogenesis inhibition/induction does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform steroidogenesis assay with added metabolising system	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						endpoints over more extensive parts of the life cycle of the organism . Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	Steroidogenesis inhibition/induction with unknown potential for effects in *in vivo* studies. May act via non- steroidogenesis interference mechanism and may or may not require metabolic activation. Unknown potential for adverse effects.	Perform steroidogenesis assay with added metabolising system or assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4) if existing data	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					indicates this is needed	
J	-	+	+	No evidence for steroidogenesis interference. Effects on ER/AR/T and potential for adverse effects via EAT mechanisms.	Perform steroidogenesis assay with added metabolising system or Perform assay from upper levels (*e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption. If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues. Further mechanistic studies may help determine MoA.
K	-	+	-	No evidence for steroidogenesis interference. Effects on ER/AR/T but effects not detected in *in vivo* studies. Metabolic differences explain *in vitro*/*in vivo* EATS differences.	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3)	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					or male or female pubertal assay (level 4).	most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* EAT activity is not realised. Consider possible routes of exposure, implications of metabolism.
L	-	+	Eq/0	No evidence for steroidogenesis interference. Effects on ER/AR/T but effects not detected in *in vivo* studies. Unknown potential for adverse effects. Metabolic differences explain *in vitro*/*in vivo* EAT differences.	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* EAT activity is not realised. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.
M	-	-	+	No evidence for steroidogenesis interference. Metabolic differences or route of exposure explain *in vitro*/*in vivo* differences. Effects seen in existing studies are via non-	Perform steroidogenesis assay with added metabolising	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption.

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				EATS or non-endocrine mechanism.	system or Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Metabolic activation of chemical may occur *in vivo*. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
N	-	-	-	No evidence for steroidogenesis interference. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3)	If existing data are from adequate level 4 or 5 assays there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
					or male or female pubertal assay (level 4).	effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Check data on chemical analogues.
O	**-**	**-**	Eq/0	No evidence for steroidogenesis interference. Unknown potential for adverse effects via other mechanisms.	Perform steroidogenesis assay with added metabolising system or assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4) if existing data indicates this is needed.	Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.
P	-	Eq/0	+	No evidence for steroidogenesis interference. Unknown potential for adverse effects via other mechanism.	Perform steroidogenesis assay with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						purpose. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
Q	-	Eq/0	-	No evidence for steroidogenesis interference. No evidence of adverse effects.	Perform steroidogenesis assay with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
R	-	Eq/0	Eq/0	No evidence for steroidogenesis interference. Unknown potential for adverse effects via	For the "0" scenario perform	Consider possible routes of exposure, implications of

Scenarios	Result of steroid-ogenesis assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				other mechanism.	steroidogenesis assay with added metabolising system or perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal (level 4) if existing data indicates this is needed.	metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.

C.2.5 Aromatase Assay (US EPA OPPTS 890.1200)

C.2.5.1 Modality detected/endpoints: Inhibition of aromatase (CYP19) enzyme activity.

Background to the Assay

C.2.5.2 The Aromatase Assay is an *in vitro* screening assay to detect substances that inhibit aromatase – the cytochrome P450 enzyme complex (CYP 19) responsible for the conversion of androgens to estrogens during steroidogenesis. Inhibition of aromatase enzyme activity alters the levels of circulating estrogens in males and females which may lead to effects on reproductive organs and other targets such as mammary gland. Aromatase is found in many vertebrate taxa, including mammals and fish and therefore the results of this assay are applicable to both human health and wildlife populations (USEPA, 2007a).

C.2.5.3 The assay determines the conversion of radiolabeled [1-^{3}H]-androstenedione to estrone. The progress of the reaction can be followed by measuring formation of either of the reaction products: estrone or water. The most common assay in usage (and the one described in guideline OPPTS 890.1200 (USEPA, 2009c) determines the formation of tritiated water as the end product of the reaction. Aromatase enzyme may be obtained from a number of sources *e.g.* human placenta or rat ovary, but human recombinant aromatase has recently become available and this is the preferred source as it is directly relevant to humans, is easily obtained and does not require the use of laboratory animals. Guideline OPPTS 890.1200 utilises the human recombinant enzyme.

C.2.5.4 Inhibition of aromatase may also be determined in the H295R steroidogenesis assay. This assay detects substances that affect production of estradiol and testosterone but the steroidogenesis assay contains all the enzymes involved in steroidogenesis, from cholesterol to estradiol and testosterone. Aromatase is the final enzyme in this pathway. Chemicals causing aromatase inhibition will be detected in the steroidogenesis assay by causing reduced production of estradiol from the H295R cells but as the assay is not specific for aromatase it would not be possible to discern which enzyme(s) activity is altered. The H295R steroidogenesis assay, as an intact cell system, will also detect chemicals that induce aromatase enzyme activity whilst the aromatase assay itself is not capable of detecting inducers.

C.2.5.5 The aromatase assay may be subject to variability, for example due to degradation of the enzyme, and therefore performance criteria are specified in guideline OPPTS 890.1200 in order to demonstrate that the assay is functioning correctly. An adequate response with the proficiency chemicals econazole, fenarimol, nitrofen (inhibitors) and atrazine (non-inhibitor) should be demonstrated and the inhibitor 4-hydroxyandrostenedione is used as a positive control chemical in each experiment. Compliance with the performance criteria should be checked before evaluating results from this assay. A positive result in guideline OPPTS 890.1200 requires demonstration of inhibition of aromatase activity that fits a 4-parameter nonlinear regression model and such that the concentration response curve crosses 50% inhibition. The concentration response curve allows the determination of potency *i.e.* IC50 (concentration at which the activity of aromatase is reduced to 50% of control values). In some cases, variability may be due to limited solubility of a chemical. The maximum concentration of chemical to be used in the assay is 1mM.

When/Why the Assay May be Used

C.2.5.6 Although the aromatase assay may be used at any stage in the hazard assessment process, the most likely use scenario is during initial assessment of chemicals for their ability to interact with endocrine systems *in vitro*, *i.e.* EATS modalities. Assays for interaction with other modalities *e.g.* AR,

ER and the steroidogenesis assay, are likely to be conducted at the same time so that all results can be considered together. TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. If the *in vitro* assays are not conducted at the same time then positive results in the steroidogenesis assay could be followed by an aromatase assay to confirm and clarify a mode of action. The aromatase assay does not include the use of a xenobiotic metabolising system but consideration should be given to the inclusion of this (Jacobs *et al*, 2008; OECD, 2008a) depending upon the circumstances *e.g.* if the metabolism of a chemical is unknown, although the methods for inclusion of xenobiotic metabolising systems are not yet validated (see paragraph B.18). Alternatively, for a chemical with known metabolites, these could also be tested in the aromatase assay. Another use scenario may be following effects obtained in higher tier tests, for example delayed puberty onset in females, but which are not exclusively indicative of an effect on ER. Selection of the most appropriate tests has to be on a case-by-case basis but also considering the need to minimise animal testing. A further example could be results obtained in other apical assays, *e.g.* OECD TG 408 (90-day toxicity test), where effects on reproductive organs may be investigated further by testing in the aromatase and steroidogenesis assays in combination with AR and ER based assays.

Introduction to the Table of Scenarios

C.2.5.7 Table C.2.5 gives guidance on a further step to take in the event of a positive (+) or negative (-) result in the aromatase assay and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.2.5.8 The results of the aromatase assay are given in the second column. Criteria for positive, negative and equivocal results are given in guideline OPPTS 890.1200. A result is judged positive if the average concentration response curve crosses 50% of control activity ("inhibitor"). A negative result is obtained if the average lowest portion of concentration response curve is greater than 75% of control activity or data do not fit the regression model ("non-inhibitor"). "Equivocal" results lie between these limits. It is important that quality and proficiency criteria are demonstrated for both positive and negative results.

Equivocal results for the guideline are not included in the table because these data require further interrogation about the result itself. This assay is a screen and therefore a clear positive or negative result should be obtained. In the event of an equivocal result, the considerations mentioned above about control quality and proficiency criteria should be taken into account and further investigations made.

Existing Data to be Considered

C.2.5.9 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR based and steroidogenesis assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform.

C.2.5.10 Existing "Effects" data refer to *in vivo* effects "of concern" *i.e.* data from level 4 or 5 mammalian or wildlife assays. These may come from varied sources and will depend upon the type of substance (*e.g.* new chemicals, HPV chemical, pesticide). Thus, available data may range from repeated dose toxicity studies (28-day, 90-day), combined repeat dose/reproductive screening assays or fish screening assays, to chronic toxicity studies and multi-generation reproductive tests in mammalian or wildlife species. Some studies fail to identify EDs that weakly affect oestrogen or

androgen receptors as was demonstrated on the basis of data generated in the validation process of the OECD TG 407 assay with endocrine endpoints. In this validation only moderate EDs, such as nonylphenol and DDE, and strong EDs such as ethinylestradiol and flutamide (acting via ER and AR respectively) were detected. The aromatase inhibitor CGS 18320B was detected by the OECD TG 407 assay, although this chemical was developed as a pharmaceutical aromatase inhibitor and therefore is a strong ED, but the ability to detect chemicals that weakly inhibit with aromatase is not known. Thus OECD TG 407 cannot be regarded as a screening assay for endocrine activity. This means that when a relatively insensitive test is positive for both endocrine-specific and apical endpoints, this should be taken as an indication that the substance is a possible ED. Caution should be exercised, however, because endocrine endpoints may be impacted secondary to non-endocrine toxicity (*e.g.* hepatotoxicity) and *in vivo* apical endpoints can be affected by all modes of action including endocrine modalities. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.2.5.11 Data may also be available from H and UT assays (level 3) but as these assays do not generally detect aromatase interference they are only useful in these cases for purposes of elimination.

C.2.5.12 When considering the results of the aromatase assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.2.5.13 The scenarios (A to R) presented in Table C.2.5 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.2.5.14 Scenarios A to C represent positive results for aromatase inhibition in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. A positive result is strong evidence for inhibition of aromatase that may or may not be supported by the *in vivo* effects data. In the case of positive *in vivo* effects data there may be sufficient evidence to conclude concern for endocrine disruption and therefore no need for further screening. Negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. A conclusion of lack of concern for endocrine disruption in the presence of positive effects data can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.5.15 Scenarios D to F represent positive results for aromatase inhibition in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Unless the metabolic profile of the test substance is known, the first option should be to conduct these *in vitro* assays with an added metabolising system. If the metabolic profile is known, then an *in vivo* test may be advisable. The choice of tests will depend upon the available *in vivo* effects data. As in scenarios A to C, negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. A conclusion of lack of concern for

endocrine disruption in the presence of positive effects data (Scenario E) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.5.16 Scenarios G to I represent positive results for aromatase inhibition in the presence of various combinations of missing or equivocal data. The next step to take in these eventualities will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step. As above, a conclusion of lack of concern for endocrine disruption in the presence of some positive effects data (Scenario H) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.2.5.17 Scenarios J to L represent negative results for aromatase inhibition in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. The limitations of the aromatase assay should be considered first (*e.g.* lack of metabolic activation, possible involvement of factors). The positive *in vitro* mechanistic data indicates possible alternative EAT mechanisms. To confirm lack of aromatase activity in the presence of *in vivo* data, an aromatase assay with added metabolising capability could be performed. Otherwise *in vivo* tests will confirm or refute EATS activity.

C.2.5.18 Scenarios M to O represent negative results in for aromatase inhibition in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend upon the weight of evidence and may not be possible in some cases. However, in the presence of negative data from robust level 4 and 5 assays further animals testing is probably not justified. The limitations of the aromatase assay should also be considered (as described for scenarios J to L). To confirm lack of aromatase inhibition in the presence of *in vivo* data, an aromatase assay with added metabolising capability could be performed. Otherwise *in vivo* tests will confirm or refute EATS activity (Scenarios M and O).

C.2.5.19 The limitations of the aromatase assay should be considered first (as described for scenarios J to L). As with the positive result scenarios above (paragraph C.5.4.16) the next step to take for Scenarios P to R when negative results in the aromatase assay are obtained in the presence of various combinations of missing or equivocal data will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.2.5.20 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.2.5 is meant to provide a succinct guide and may not cover all circumstances or possibilities. In general, a decision about whether or not to conduct *in vivo* mammalian or wildlife tests will depend on the weight of evidence of new and existing data. If most available data (*e.g*. the results of the steroidogenesis assay, predictions from QSARs, 'read-across' from data on similar substances, and results from mammalian *in vivo* assays) suggest that the substance has the potential to cause endocrine disruption via inhibition of aromatase (*i.e*. the level of suspicion about endocrine disrupting action is high – corresponding to Scenario A), then consideration should be given to the conduct of a higher level test.

C.2.5.21 For wildlife species, higher level tests with fish (*i.e* TG 234 (FSDT), the FLCTT or the MMGT) are recommended. Choice about which of these tests is most appropriate will be driven *inter*

alia by mode of action considerations, and by whether multi-generation effects are to be expected. Such tests are unlikely to be needed if exposure of the natural environment is not expected. On the other hand, if available data only raise a low or moderate level of suspicion about endocrine disrupting action (*e.g*. the data appear to conflict with each other), then consideration should be given to the conduct of a fish screen (*i.e*. OECD TG 229 or 230).

C.2.5.22 For mammals, similar considerations apply but lower level tests (*e.g.* Level 3 or 4) should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study.

Table C.2.5. Aromatase Assay (US EPA OPPTS 890.1200). Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER and AR based assays (level 2). It is assumed that data from the steroidogenesis assay are also available. TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes various information, such as data from repeat oral toxicity studies, reproduction/developmental toxicity screen tests, read across from analogues, will be available.

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	Inhibition of aromatase combined with effects on ER/AR/T/S and potential for adverse effects via multiple mechanisms.	Perform assay from upper levels *e.g.* male or female pubertal assay (level 4) or ext-1 or 2-gen assays or partial/full fish life cycle tests	If existing data are from level 5 there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
					(level 4/5).	comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Compare aromatase assay results with other *in vitro* results to help discern mechanism. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results and *in vivo.* results but may also be metabolised to a metabolite that also has positive results *in vitro* and *in vivo*.
B	+	+	-	Inhibition of aromatase combined with effects on ER/AR/T/S but effects not detected in *in vivo* studies. Weak aromatase inhibition does not result in	Perform aromatase assay with added metabolising	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	system OR assay from upper levels *e.g.* fish screen (OECD TG 229/230) (leve 3) or male or female pubertal assay (level 4).	disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Compare aromatase assay results with other *in vitro* results to help discern mechanism. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						metabolism of chemicals can be different between wildlife species.
C	+	+	Eq/0	Inhibition of aromatase combined with effects on ER/AR/T but no or equivocal data from *in vivo* studies. Weak aromatase inhibition may not result in adverse effects.	Perform assay upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	Compare aromatase assay results with other *in vitro* results to help discern mechanism. Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
D	+	-	+	Inhibition of aromatase and potential for adverse effects.	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) male or female pubertal assay or (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism,

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
E	+	**-**	**-**	Inhibition of aromatase but effects not detected in *in vivo* studies. Weak aromatase inhibition does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform aromatase assay with added metabolising system OR assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose.

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species. Check data on chemical analogues.
F	+	**-**	Eq/0	Inhibition of aromatase but no or equivocal data from *in vivo* studies. Weak aromatase inhibition may not result in adverse effects.	Perform assay upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
G	+	Eq/0	+	Inhibition of aromatase and potential for adverse effects via aromatase inhibition or other EATS mechanisms.	Perform assay from upper levels *e.g.* fish screen	If existing data are from an adequate level 5 assay there may be sufficient information to conclude

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				May act via non- aromatase inhibition mechanism and may or may not require metabolic activation.	(OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Check data on chemical analogues. Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
H	+	Eq/0	**-**	Inhibition of aromatase but effects not detected in *in vivo* studies. Weak aromatase inhibition does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo*	Perform aromatase assay with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				differences.		provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species. Check data on chemical analogues.

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	Inhibition of aromatase with unknown potential for effects in *in vivo* studies. May act via non- aromatase inhibition mechanism and may or may not require metabolic activation. Unknown potential for adverse effects.	Perform aromatase assay with added metabolising system, or assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4) if existing data indicates this is needed.	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
J	-	+	+	No evidence for aromatase inhibition. Effects on ER/AR/T/S and potential for adverse effects via EAT mechanisms.	Perform aromatase assay with added metabolising system or Perform assay	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption. If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
					from upper levels (*e.g.* fish screen OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues. Further mechanistic studies may help determine MoA.
K	-	+	-	No evidence for aromatase inhibition. Effects on ER/AR/T/S but effects not detected in *in vivo* studies. Metabolic differences explain *in vitro*/*in vivo* EATS differences.	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4)	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* EATS activity is not realised. Consider possible routes of exposure, implications of metabolism.
L	-	+	Eq/0	No evidence for aromatase inhibition. Effects on ER/AR/T/S but effects not detected in *in vivo* studies. Unknown potential for adverse effects. Metabolic differences explain *in vitro*/*in vivo* EAT differences.	Perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4)	Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* EATS activity is not realised. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.
M	-	-	+	No evidence for aromatase inhibition. Metabolic differences or route of exposure explain *in vitro*/*in vivo* differences.	Perform aromatase assay with added	If existing data are from an adequate level 5 assay there may be sufficient information to conclude

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				Effects seen in existing studies are via non-EATS or non-endocrine mechanism.	metabolising system or perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	concern for endocrine disruption. If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Metabolic activation of chemical may occur *in vivo*. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
N	-	-	-	No evidence for aromatase inhibition. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
					perform assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay (level 4).	disruption (the ext-1 gen assay provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Check data on chemical analogues.
O	**-**	**-**	Eq/0	No evidence for aromatase inhibition. Unknown potential for adverse effects via other mechanisms.	Perform aromatase assay with added metabolising system or assay from upper levels *e.g.* fish screen (OECD TG 229/230) (level 3) or male or female pubertal assay	Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
					(level 4) if existing data indicates this is needed.	
P	-	Eq/0	+	No evidence for aromatase inhibition. Unknown potential for adverse effects via other mechanism.	Perform aromatase assay with added metabolising system	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
Q	-	Eq/0	-	No evidence for aromatase inhibition. No evidence of adverse effects.	Perform steroidogenesis assay with added metabolising system	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						provides the most information). If existing data are from level 3 or 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism, although some Level 4 assays (*e.g.* TG 234 (FSDT)) may be sufficient for this purpose. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
R	-	Eq/0	Eq/0	No evidence for aromatase inhibition. Unknown potential for adverse effects via other mechanism.	For the "0" scenario perform steroidogenesis assay with added metabolising system or perform assay from upper levels *e.g.* fish screen (OECD TG	Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.

Scenarios	Result of aromatase assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
					229/230) (level 3) or male or female pubertal (level 4) if existing data indicates this is needed.	

C.3 Wildlife Screens and Tests

C.3.1 OECD TG 229: Fish Short Term Reproduction Assay (FSTRA).

C.3.1.1 Modality detected/endpoints: Estrogens (♂VTG ↑; ♂ 2° sex characteristics ↓); Anti-estrogens (♀VTG↓); Androgens (♂ 2° sex characteristics in ♀); Anti-androgens (♂ 2° sex characteristics ↓); Aromatisable androgens (♂VTG ↑); Aromatase inhibitors (♀VTG↓); Non-specific effects on HPG axis, plus other reprotox (fecundity ↓); (Optional endpoint – gonadal histo-pathology. This may assist with diagnosis of MOA). Note that this assay may, in some cases, have low statistical power or sensitivity to detect anti-androgenic activity through effects on secondary sexual characteristics. However, if gonad histopathology has been optionally studied, changes in Leydig cells resulting from anti-androgen exposure may have been observed. Finally, diagnostic endpoints (*i.e.* indicators of hormonal activity) and the apical endpoint (*i.e.* fecundity) should be considered together to obtain maximum value from this assay.

Background to the Assay

C.3.1.2 This assay is primarily designed as a screen for the types of *in vivo* endocrine disruption activity in fish which are listed above, but it does also provide information on adverse effects on fecundity which could be used in assessing the environmental risks of an individual chemical based on a PEC/PNEC approach (although note that only 3 test concentrations are normally used, so precision of a NOEC/ECx may be relatively low). The fecundity endpoint, which although not necessarily diagnostic of endocrine action, does indicate that apical effects on reproduction are occurring, is sensitive to known EDs. However, the validation studies demonstrated high variability for fecundity (and consequently low power to detect an effect) under certain sub-optimal test conditions. If the assay gives a positive result, this may be due to a positive indicator of hormonal activity (VTG level, secondary sexual characteristic development), which may or may not be associated with decrease in fecundity. Each of these three possible combinations of positive response should be considered separately, (although the distinctions between indicators of hormonal activity and apical effects are not always clear) so they have been listed individually as points 1, 2 and 3 in the Possible Conclusions column of Table C.3.1. It should be noted, however, that due to the relatively short exposure time employed in this screen (3 weeks), effects of some chemicals on fecundity might not be as apparent as in longer-term exposures, especially for bioaccumulative chemicals. Also, as only 3 test concentrations are employed, even a reliable short-term NOEC or ECx for fecundity cannot be precisely derived.

When/Why the Assay May be Used

C.3.1.3 Although OECD TG 229 could, in principle, be used at any stage in the hazard assessment process, the most likely use scenario will be when there are relatively few data available about the possible endocrine disrupting properties of a chemical. The results from this assay are most likely to be available after deployment of a battery of *in vitro* and *in vivo* screens (*e.g.* the USEPA's Endocrine Disruptor Screening Program), or as a supplement to existing data which suggest possible endocrine disruption activity. It is also possible that no existing endocrine-relevant data are available (*i.e.* OECD TG 229 has been used as a primary screen), but in that case a positive result in the screen should ideally be followed up with relevant *in vitro* screening to investigate the suspected mode of action.

Existing Data to be Considered

C.3.1.4 Existing data available before deployment of OECD TG 229 might include *in vivo* results obtained with other vertebrates (*e.g.* a positive uterotrophic assay with rodents), or one or more of a range of *in silico* or *in vitro* results which suggest that the modalities indicated above may occur *in vivo*. Such indicators of possible *in vivo* activity might include (Q)SAR predictions of endocrine activity, 'read-across' from *in vivo* results obtained with chemically related chemicals, or positive results from an *in vitro* screen for estrogen or androgen receptor-mediated activity, or for effects on steroidogenesis (especially aromatase inhibition). OECD TG 229 may also be used as part of a battery of screening assays. Conduct of OECD TG 229 would be particularly relevant if knowledge is sought about the test chemical's effects on the mature reproductive phase of the fish lifecycle (as opposed to effects on the immature sexual development phase), because it provides some apical information on reproductive success and gonad histopathology. However, this assay is also likely to be responsive to many chemicals which act primarily on sexual development.

Scenarios: Positive and Negative Results Combined with Existing Data

C.3.1.4a The scenarios (A to R) presented in Table C.3.1 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.3.1.5 Positive results obtained with one or more of the indicators of hormonal activity (Table C.3.1, Scenarios A-I, sub-section 2) result in the conclusion that the test chemical is a possible ED *in vivo*. If both an indicator of hormonal activity and fecundity give a response (Table C.3.1, Scenarios A-I, sub-section 1), this provides strong evidence for *in vivo* endocrine activity on the hypothalamic-pituitary-gonadal (HPG) axis with potential adverse effects. If only fecundity responds (Table C.3.1, Scenarios A-I, sub-section 3), it suggests that the chemical is a reproductive or general systemic toxicant, with a reduced probability that it is an ED that acts on one or more of the endocrine modalities covered in the Conceptual Framework (although existing positive *in vitro* data, or positive *in vivo* data from other species, would have to be weighed against this conclusion).

C.3.1.6 As indicated above, although a combined effect on fecundity and an indicator of hormonal activity in OECD TG 229 suggests that the test chemical is a reproductive toxicant acting through one or more EATS pathways (assuming that the concentration giving this response is not sufficiently high to cause systemic toxicity), a result of this type would generally need to be followed up with a more comprehensive reproduction test if countries need further evidence (*e.g.* a Fish Lifecycle Toxicity Test –FLCTT - or Medaka Multi-Generation Test – MMGT) which is able to provide a more reliable and reproducible NOEC or ECx for adverse effects. An exception might be if there are no indications of endocrine activity (either from this or other screens/tests), although in such a case, a NOEC or ECx for reproductive effects would still need to be derived for a non-endocrine risk assessment (*e.g.* using data from OECD TG 210). Equally, if one or more biomarkers for hormonal activity alone respond without a corresponding response from apical endpoints, this would also need to be followed up with more comprehensive testing to show whether any adverse apical effects occur at other parts of the lifecycle, if countries need further evidence whether the chemical is an ED. In other words, in order to increase evidence in relation to ED, a positive result of whichever type in OECD TG 229 could be followed by fish partial- or full lifecycle testing at Level 5. Existing data suggesting endocrine-specific activity (*e.g.* positive *in vitro* data, or positive *in vivo* data from other species) will strengthen the case for additional testing still further if the intention is to establish a firmer link between endocrine activity and adverse effects.

C.3.1.7 The situation in which OECD TG 229 gives a negative result (Table C.3.1, Scenarios J-R) needs careful consideration of the weight of evidence based on any existing data. If these data suggest that the chemical is endocrine-active both *in vitro* and *in vivo* (Scenario J), then it is possible that OECD TG 229 is simply insufficiently sensitive, perhaps due to rapid metabolism, or because the main mode of action (MOA) acts more potently during sexual development, or because fish in general are simply insensitive to the chemical under consideration. In some of these circumstances, it might therefore be appropriate to conduct a Fish Sexual Development Test (FSDT) (TG 234), or alternatively, a FLCTT to confirm that there is no endocrine activity in fish.

C.3.1.8 If OECD TG 229 and existing *in vivo* data are all negative, but *in vitro* data reveal some endocrine activity (Scenario K), the probability is that the test chemical is not sufficiently potent to produce endocrine effects *in vivo* in fish, *e.g*. because it is rapidly metabolised to ED-inactive metabolites. In such a situation, further testing in fish is probably not necessary. However, if the chemical is known to bioaccumulate slowly, it may be that exposures in the *in vivo* tests have been insufficiently prolonged, in which case longer-term testing might be justified. Equally, if the *in vitro* or histopathology data reveal anti-androgenic or thyroid activity, consideration may be given to conducting the Androgenised Female Stickleback Screen or the Amphibian Metamorphosis Assay (OECD TG 231), respectively.

C.3.1.9 On the other hand, if OECD TG 229 and the *in vitro* tests are negative (Scenario M), but there are positive existing *in vivo* data, the nature of those existing data should be considered. Unless the existing data are from fish, the chemical is probably not a possible ED acting on fish reproduction, but it may act via MOAs not covered by the *in vitro* screens, or it may be more potent in species or life-stages that have not been tested. In this situation, the existing *in vivo* data should be used to guide decisions about whether to conduct any further testing, either for modalities such as anti-androgenicity or thyroid activity, or including lifestages represented in TG 234 (FSDT) or in the FLCTT.

C.3.1.10 Finally, a negative OECD TG 229 screen, set against a background of negative *in vitro* and *in vivo* data (Scenario N), suggests that the test chemical is not an ED acting on reproduction in fish, and no further testing for estrogenic, anti-estrogenic, androgenic or steroidogenic MOAs will generally be necessary. It remains possible that it has anti-androgenic or thyroid activity, although this scenario is unlikely if relevant *in vitro* tests for these modalities have shown negative results and if no effects have been detected by gonadal histopathology.

C.3.1.11 In each of the above scenarios, it is possible that existing data will be equivocal, or there may be no existing data (Scenarios C, F, G, H, I, L, O, P, Q and R). This will weaken the conclusions which can be drawn about a negative OECD TG 229 test, and this is reflected in Table C.3.1. However, a lack of mechanistic data on endocrine activity should ideally be rectified before any further *in vivo* testing is finally rejected. Indeed, as a general principle, it is desirable to obtain mechanistic data before any *in vivo* testing. On the other hand, if OECD TG 229 is positive, further *in vivo* testing may be needed to establish a more precise NOEC or ECx for any adverse effects, even if all other existing data are equivocal, or if there are no existing data. Again, however, it will always be desirable to obtain some mechanistic information before conducting further *in vivo* testing.

C.3.1.12 The scenario in which the results of OECD TG 229 are themselves equivocal has not been dealt with in Table C.3.1, for reasons of brevity. In this context, an equivocal result might be a non-monotonic concentration-response (*e.g.* no effect at a high concentration but effects at a lower), or a result which borders on statistical significance. Without knowing the exact circumstances, reliable advice cannot be given, but the opinions of an experienced ecotoxicologist should be sought. Clearly, however, such equivocal results do not necessarily rule out the existence of *in vivo* endocrine activity. For example, vitellogenin induction in males at a high concentration might be masked by any systemic toxicity, while fecundity depression might just fail to reach a statistically significant level because the sometimes high variability of this endpoint combined with a relatively small sample size might have reduced the power of the test to detect a difference from the controls. If these or other possible reasons for false negatives are suspected with good reason, the screen could be repeated (*e.g.*

conduct it at lower concentrations which avoid systemic toxicity), or a more appropriate version of it (*e.g.* more fish per replicate) could be designed and conducted. However, note that a repeat test in the event of systemic toxicity would not be needed providing at least one tested concentration was not subject to such toxicity.

C.3.1.13 In summary, positive results in the OECD TG 229 screen indicate that a chemical is either a reproductive toxicant, or a possible endocrine disrupter, or both. In most cases, more comprehensive *in vivo* testing would be needed if the intention is to derive a long-term NOEC/ECx and/or to confirm whether or not the chemical is an actual endocrine disrupter. In this connection, it should also be borne in mind that effects solely on fecundity might be caused by systemic toxicity rather than endocrine disruption or specific reproductive toxicity, if test concentrations were very high. Negative results in OECD TG 229 do not necessarily mean that the chemical is not a possible ED – a judgement about the endocrine disruption potential and the possible need for additional testing will have to be made in the light of existing *in vitro* and *in vivo* data.

Table C.3.1 OECD TG 229: Fish Short-Term Reproduction Assay. Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

The assay under discussion could either be positive for both apical endpoints and indicators of hormonal activity, or positive just for apical endpoints, or positive just for indicators of hormonal activity. For each scenario, each of these 3 possibilities is addressed separately in the Possible Conclusions column, taking into consideration other existing data.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR, and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". There is no evidence at present that equivalent *in vitro* assays with systems derived from fish offer advantages over their mammalian counterparts.

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter.

Note that under some scenarios, a Medaka Multi-Generation Test (MMGT) is recommended as a possible Next Step. This test is still being validated, so it is described relatively briefly in Annex 2.

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	1) Strong evidence for *in vivo* endocrine activity with potential adverse effects (reproductive toxicity) in fish	Consider performing a fish lifecycle test,	An alternative approach would be to deploy the fish sexual development test, especially if

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				2) Strong evidence for *in vivo* endocrine activity in fish 3) Evidence for *in vivo* endocrine activity in other species, and strong evidence for reproductive toxicity in fish	especially if the intention is to obtain precise data on a reproductive or developmental NOEC/ECx.	sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a Fish Lifecycle Toxicity Test (FLCTT) or multi-generation test (MMGT) may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the young via the eggs.
B	+	+	-	1) Strong-moderate evidence for *in vivo* endocrine activity with potential adverse effects (reproductive toxicity) in fish 2) Strong-moderate evidence for *in vivo* endocrine activity in fish 3) Moderate-weak evidence for *in vivo* endocrine activity, and strong evidence for reproductive toxicity in fish	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive or developmental NOEC/ECx.	An alternative approach would be to deploy the fish sexual development test, especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a fish lifecycle or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the young via the eggs. If the negative *in vivo* data are from another fish endocrine assay, consider possible reasons for the disparity (*e.g.* differences in species sensitivity) before conducting a lifecycle test.
C	+	+	Eq/0	1) Moderate evidence for *in vivo* endocrine activity with potential adverse effects in fish 2) Moderate evidence for *in vivo* endocrine activity in fish 3) Weak evidence for *in vivo* endocrine activity, and strong evidence for reproductive toxicity in fish	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive or developmental NOEC/ECx.	An alternative approach would be to deploy the fish sexual development test, especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a fish lifecycle or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the young via the

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						eggs. If no existing fish data are available, it may be worth performing a Fish Sexual Development Test before a lifecycle test in order to obtain information on whether sexual development is the most sensitive part of the lifecycle. Such information could influence the design of the lifecycle test.
D	+	-	+	1) Strong evidence for *in vivo* endocrine activity with potential adverse effects in fish 2) Strong evidence for *in vivo* endocrine activity in fish 3) Evidence for *in vivo* endocrine activity in other species, and strong evidence for reproductive toxicity in fish	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive or developmental NOEC/ECx.	The negative *in vitro* data suggest that the test chemical may be metabolically activated *in vivo* or may not act via the screened receptor. An alternative approach would be to deploy the fish sexual development test, especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a fish lifecycle or multi-generation test may be driven

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the young via the eggs.
E	+	-	-	1) Moderate-strong evidence for *in vivo* endocrine activity with potential adverse effects in fish 2) Moderate-strong evidence for *in vivo* endocrine activity in fish 3) Weak-moderate evidence for *in vivo* endocrine activity, and strong evidence for reproductive toxicity in fish	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive or developmental NOEC/ECx.	The negative *in vitro* data suggest that the test chemical may be metabolically activated *in vivo*. An alternative approach would be to deploy the fish sexual development test, especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a fish lifecycle or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the young via the eggs.

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						If the negative *in vivo* data are from another fish endocrine assay, consider possible reasons for the disparity (*e.g.* differences in species sensitivity) before conducting a lifecycle test.
F	+	-	Eq/0	1) Moderate-strong evidence for *in vivo* endocrine activity with potential adverse effects in fish 2) Moderate-strong evidence for *in vivo* endocrine activity in fish 3) Weak evidence for *in vivo* endocrine activity, and strong evidence for reproductive toxicity in fish	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive or developmental NOEC/ECx.	The negative *in vitro* data suggest that the test chemical may be metabolically activated *in vivo*. The decision about whether to conduct a fish lifecycle or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the young fish via the eggs. If no existing fish data are available, it may be worth performing a Fish Sexual Development Test before a lifecycle test in order to obtain information on whether sexual

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						development is the most sensitive part of the lifecycle. Such information could influence the design of the lifecycle test.
G	+	Eq/0	+	1) Strong evidence for *in vivo* endocrine activity with potential adverse effects in fish, but mechanism unconfirmed 2) Strong evidence for *in vivo* endocrine activity in fish, but mechanism unconfirmed 3) Moderate evidence for *in vivo* endocrine activity (mechanism unconfirmed), and strong evidence for reproductive toxicity in fish	Obtain more predictive mechanistic data and then consider performing a fish lifecycle test	An alternative approach would be to deploy the fish sexual development test, especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a fish lifecycle or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the young fish via the eggs.
H	+	Eq/0	-	1) Moderate-strong evidence for *in vivo* endocrine activity with potential adverse effects in fish, but mechanism unconfirmed 2) Moderate-strong evidence for *in vivo* endocrine activity in fish, but mechanism unconfirmed	Obtain more predictive mechanistic data and then consider performing a fish lifecycle test	An alternative approach would be to deploy the fish sexual development test, especially if sexual development is expected to give a response at lower concentrations than reproduction.

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				3) Weak-moderate evidence for *in vivo* endocrine activity (mechanism unconfirmed), and strong evidence for reproductive toxicity in fish		The decision about whether to conduct a fish lifecycle or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the young fish via the eggs. If the negative *in vivo* data are from a fish test (*e.g.* the 21 d fish assay), consider possible reasons for the disparity (*e.g.* differences in species sensitivity) before conducting a lifecycle test.
I	+	Eq/0	Eq/0	1) Moderate evidence for *in vivo* endocrine activity with potential adverse effects in fish, but mechanism unconfirmed 2) Moderate evidence for *in vivo* endocrine activity in fish, but mechanism unconfirmed 3) Weak evidence for *in vivo* endocrine activity (mechanism unconfirmed), but strong evidence for reproductive toxicity in fish	Obtain more predictive mechanistic data and then consider performing a fish lifecycle test	The decision about whether to conduct a fish lifecycle or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) may generally be sufficient if the chemical is not expected to be transferred to the young via the

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						eggs. If no existing fish data are available, it may be worth performing a Fish Sexual Development Test before a lifecycle test in order to obtain information on whether sexual development is the most sensitive part of the lifecycle. Such information could influence the design of the lifecycle test.
J	-	+	+	Based on the existing data, the chemical has endocrine activity *in vivo*. The lack of response in OECD TG 229 suggests that fish are not responsive, unless the existing data are from fish	If existing *in vivo* data are from fish, consider performing a Fish Sexual Development Test (unless reproduction is known to be the most sensitive life-stage).	
K	-	+	-	There is no evidence that the chemical is an ED *in vivo*, probably because it is very weakly acting, rapidly metabolised or simply does not reach the target site	Probably no further action, but see comments in right-hand column.	It is possible that EDs which bioaccumulate slowly may only cause effects *in vivo* after exposure times longer than 3

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						weeks. If this is suspected, and depending on which part of the lifecycle is suspected of being the most sensitive, consider performing a Fish Sexual Development Test or a Fish Lifecycle Toxicity Test. It is also possible that the chemical may be an anti-androgen *in vivo* (consider performing the Androgenised Female Stickleback Screen), or a thyroid-active chemical *in vivo* (consider performing the Amphibian Metamorphosis Assay – OECD TG 231)
L	-	+	Eq/0	The chemical may not be an ED *in vivo*, but the confidence in this conclusion is relatively low as there is only one unequivocal *in vivo* test result (a negative).	If the existing *in vivo* data are equivocal and from a fish, consider performing a fish assay (OECD TG 229 or 230) with a different species, or consider a	It is also possible that the chemical may be an anti-androgen *in vivo* (consider performing the Androgenised Female Stickleback Screen), or a thyroid-active chemical *in vivo* (consider performing the Amphibian Metamorphosis Assay – OECD TG 231)

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
					longer-term test (TG 234 (FSDT) or lifecycle).	
M	-	-	+	The chemical is probably not an ED acting on reproduction in fish. However, it has endocrine activity in another species and may act through MOAs not covered by the available *in vitro* assays, or it may be more potent in a species other than that tested, or over a longer exposure period.	If further evidence is required, consider using the existing *in vivo* data to help choose a longer-term test with an appropriate species.	It is also possible that the chemical may be an anti-androgen *in vivo* (consider performing the Androgenised Female Stickleback Screen), or a thyroid-active chemical *in vivo* (consider performing the Amphibian Metamorphosis Assay – OECD TG 231), although lack of *in vitro* binding affinity with receptors suggests this is unlikely. Use the existing *in vivo* data to guide any further testing.
N	-	-	-	The chemical is probably not an ED acting on reproduction in fish. There is a possibility that the chemical is able to affect sexual development in fish, but the probability of this is low given the apparent absence of estrogenic, androgenic or steroidogenic properties.	No further action with respect to estrogenic, anti-estrogenic, androgenic or steroidogenic MOAs.	It is still possible that the chemical may be an anti-androgen *in vivo* (consider performing the Androgenised Female Stickleback Screen), or a thyroid-active chemical *in vivo* (consider performing the Amphibian Metamorphosis Assay – OECD TG 231), although lack of *in vitro* binding affinity with receptors

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						suggests this is unlikely.
O	-	-	Eq/0	The chemical is probably not an ED acting on reproduction in fish. There is a possibility that the chemical is able to affect sexual development in fish, but the probability of this is low given the apparent absence of estrogenic, androgenic or steroidogenic properties.	Probably no further action. However, see comments in right-hand column.	If the paucity of *in vivo* data are a concern, performance of a screening test (OECD TG 229 or 230) with a different species, or a longer-term test (*i.e.* TG 234 (FSDT) or lifecycle) could be considered. It is also possible that the chemical may be an anti-androgen *in vivo* (consider performing the Androgenised Female Stickleback Screen), or a thyroid-active chemical *in vivo* (consider performing the Amphibian Metamorphosis Assay – OECD TG 231), although lack of *in vitro* binding affinity with receptors suggests this is unlikely.
P	-	Eq/0	+	The chemical may not be an ED acting on reproduction in fish, but confidence in this conclusion is low given the lack of mechanistic *in vitro* data and the availability of positive existing *in vivo* data.	Obtain more predictive mechanistic data, and then consider further testing.	If the mechanistic data confirm that the chemical has potential endocrine action, consider conducting a fish assay (OECD TG 229 or 230) with another species, or a longer term test (TG

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						234 (FSDT) or lifecycle). Use the existing *in vivo* data as a guide to test design. If the mechanistic data reveal anti-androgenic or thyroid activity, perform the Androgenised Female Stickleback Screen or the Amphibian Metamorphosis Assay (OECD TG 231), respectively.
Q	-	Eq/0	-	The chemical is probably not an ED acting on reproduction in fish, but the lack of more predictive mechanistic data are a concern, even though the existing *in vivo* data are negative.	Obtain more predictive mechanistic data, and then consider further testing.	If the mechanistic data confirm that the chemical has potential endocrine action, consider conducting a fish assay (OECD TG 229 or 230) with another species, or a longer term test (TG 234 (FSDT) or lifecycle). Use the existing *in vivo* data as a guide to test design. If the mechanistic data reveal anti-androgenic or thyroid activity, perform the Androgenised Female Stickleback Screen or the Amphibian Metamorphosis Assay (OECD TG 231), respectively.
R	-	Eq/0	Eq/0	The chemical is probably not an ED acting on reproduction in fish, but confidence in this	Obtain more predictive	If the mechanistic data confirm that the chemical has potential

Scenarios	Result of OECD TG 229 assay	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				conclusion is low given the lack of mechanistic *in vitro* and existing *in vivo* data.	mechanistic data, and then consider further testing.	endocrine action, consider conducting a fish assay (OECD TG 229 or 230) with another species, or a longer term test (TG 234 (FSDT) or lifecycle). Use the existing *in vivo* data as a guide to test design. If the mechanistic data reveal anti-androgenic or thyroid activity, perform the Androgenised Female Stickleback Screen or the Amphibian Metamorphosis Assay (OECD TG 231), respectively.

C.3.2 OECD TG 230: 21-Day Fish Assay.

C.3.2.1 Modality detected/endpoints: Estrogens (♂VTG ↑; ♂ 2° sex characteristics ↓); Anti-estrogens (♀VTG↓); Androgens (♂ 2° sex characteristics in ♀); Anti-androgens (♂ 2° sex characteristics ↓); Aromatisable androgens (♂VTG ↑); Aromatase inhibitors (♀VTG↓). Note that this assay has low statistical power to identify anti-androgenic activity.

Background to the Assay

C.3.2.2 This assay is designed as a screen for the types of *in vivo* endocrine disruption activity in fish which are listed above. The endpoints are indicators of hormonal activity and there are no apical measures of adverse effects that can be attributed to a single EATS modality (although it is possible that some substances could cause cessation of spawning). A variation of this assay specifically designed for the detection of androgens and anti-androgens, the Androgenised Female Stickleback Screen (AFSS), is described in a separate section of this document.

When/Why the Assay May be Used

C.3.2.3 Although data from OECD TG 230 could, in principle, be available at any stage in the hazard assessment process, the most likely scenario will be when there are relatively few data available about the possible endocrine disrupting properties of a chemical. The assay is most likely to be used either as part of a battery of *in vitro* and *in vivo* screens, or to follow up on existing data which suggest possible endocrine disruption activity. It is also possible that no existing endocrine-relevant data are available (*i.e.* OECD TG 230 has been used as a primary screen), but in that case a positive result in the screen should ideally be followed up with relevant *in vitro* screening in an attempt to confirm the suspected mode of action. Possible conclusions to be derived from the results of OECD TG 230, and guidance about potential additional studies to increase evidence, are summarised below in Table C.3.2.

Existing Data to be Considered

C.3.2.4 Existing data available before deployment of OECD TG 230 might include *in vivo* results obtained with other vertebrates (*e.g.* a uterotrophic assay with rodents), or one or more of a range of *in silico* or *in vitro* results which suggest that the modalities indicated above may occur *in vivo*. Such indicators of possible *in vivo* activity might include (Q)SAR predictions of endocrine activity, 'read-across' from *in vivo* results obtained with chemically related chemicals, or positive results from an *in vitro* screen for estrogen or androgen receptor-mediated activity, or for effects on steroidogenesis (especially aromatase inhibition).

Scenarios: Positive and Negative Results Combined with Existing Data

C.3.2.4a The scenarios (A to R) presented in Table C.3.2 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.3.2.5 Positive results obtained with one or more of the endpoints (Table C.3.2, Scenarios A-I) result in the conclusion that the test chemical is a possible ED *in vivo*. This would ideally need to be followed up with more comprehensive testing to show whether adverse apical effects related to endocrine impacts occur at any part of the lifecycle (and hence to discover whether the chemical is an ED acting through EATS pathways). In other words, a positive result in OECD TG 230 may trigger TG 234 (FSDT) at Level 4 or fish lifecycle testing at Level 5. Existing data suggesting endocrine activity will strengthen the case for additional testing.

C.3.2.6 The situation in which OECD TG 230 gives a negative result (Table C.3.2, Scenarios J-R) needs careful consideration of any existing data. If the weight of evidence of these data suggests that the chemical is endocrine-active both *in vitro* and *in vivo* in other species (Scenario J), then the probability is that OECD TG 230 is simply insufficiently responsive in that case, or fish in general may be unresponsive. In some of these circumstances, it might be appropriate to conduct a Fish Sexual Development Test (FSDT) (TG 234), or alternatively, a fish lifecycle test (either FLCTT or MMGT) to confirm that there is no endocrine activity in fish.

C.3.2.7 If OECD TG 230 and existing *in vivo* data are all negative, but *in vitro* data reveal some endocrine activity (Scenario K), the probability is that the test chemical is not sufficiently potent to produce endocrine effects *in vivo* in adult fish, or it may be rapidly metabolised. However, TG 230 does not include some endpoints which are included in TG 229 (fecundity and histopathology) which is able to detect certain endocrine-active substances not detected by TG 230 alone. In such a situation, further testing may or may not be necessary. A lack of effects in adult fish does not preclude the possibility that endocrine-mediated effects may manifest in fish exposed during a more sensitive life stage, *e.g.*, as embryos or larvae. If the chemical is known to bioaccumulate slowly, it may be that exposures in the *in vivo* tests are not of sufficient duration, in which case longer-term testing might be justified. If the *in vitro* data reveal anti-androgenic or thyroid activity, consideration should be given to conducting the Androgenised Female Stickleback Screen or the Amphibian Metamorphosis Assay (OECD TG 231), respectively.

C.3.2.8 On the other hand, if OECD TG 230 and the *in vitro* tests are negative, but there are positive existing *in vivo* data (Scenario M), the chemical is probably not a possible ED with the modalities listed above, but it may act via estrogen- or androgen-related MOAs not covered by the *in vitro* screens, or it may be more potent in species or life-stages that have not been tested. In this situation, the relevant existing *in vitro* and *in vivo* data should be used to guide decisions about whether to conduct any further testing, either for modalities such as anti-androgenicity or including lifestages represented in the TG 234 (FSDT) or in FLCTT/MMGT.

C.3.2.9 Finally, a negative OECD TG 230 screen, set against a background of negative *in vitro* and *in vivo* data (Scenario N) that includes relevant *in vivo* data for fish, suggests that the test chemical is not a possible ED in fish or other vertebrates, and no further testing for estrogenic, anti-estrogenic, androgenic or steroidogenic MOAs will generally be necessary. It remains possible that it has anti-androgenic or thyroid activity, although negative *in vitro* tests for these modalities would suggest that this scenario is unlikely.

C.3.2.10 In each of the above scenarios, it is possible that existing data will be equivocal, or there may be no existing data (Scenarios C, F, G, H, I, L, O, P, Q and R). This will weaken the conclusions which can be drawn about a negative OECD TG 230 test, and this is reflected in Table C.3.2. However, a lack of mechanistic data on endocrine activity should usually be rectified before any further *in vivo* testing is finally rejected. Indeed, as a general principle, it is desirable to obtain mechanistic data before any *in vivo* testing. On the other hand, if OECD TG 230 is positive, further *in vivo* testing is generally indicated, particularly when existing data are equivocal, or if there are no existing data.

C.3.2.11 The scenario in which the results of OECD TG 230 are themselves equivocal has not been dealt with in Table C.3.2, for reasons of brevity. In this context, an equivocal result might be an

inconsistent concentration-response (*e.g.* no effect at a high concentration but effects at a lower), or a result which borders on statistical significance. Without knowing the exact circumstances, reliable advice cannot be given, but the opinions of an experienced ecotoxicologist should be sought. Clearly, however, such equivocal results do not necessarily rule out the existence of *in vivo* endocrine activity. For example, vitellogenin induction in males at a high concentration might be masked by any systemic toxicity, while VTG depression in females might just fail to reach a statistically significant level because VTG levels were relatively low to begin with. If these or other possible reasons for false negatives are suspected with good reason, the screen could be repeated if none of the test concentrations have given reliable data (*e.g.* conduct it at lower concentrations which avoid systemic toxicity), or a more appropriate version of it (*e.g.* ensure females have high VTG levels at the start of the test) could be conducted. However, note that a repeat test in the event of systemic toxicity would not be needed providing at least one tested concentration was not subject to such effects.

C.3.2.12 In summary, positive results in the OECD TG 230 screen indicate that a chemical is a possible endocrine disrupter. More predictive *in vivo* testing would then be necessary to produce a long-term NOEC/ECx and/or to confirm whether or not the chemical is an actual endocrine disrupter with adverse effects *in vivo*. Negative results in OECD TG 230 do not necessarily mean that the chemical is not a possible ED – a judgement about its endocrine disruption potential and the possible need for additional testing will have to be made based on a weight of evidence evaluation of existing *in vitro* and *in vivo* data.

Table C.3.2 OECD TG 230: 21-Day Fish Assay. Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available. Note that there are no apical endpoints in this assay considered to be diagnostic of an EATS modality.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR, and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". There is no evidence at present that equivalent *in vitro* assays with systems derived from fish offer advantages over their mammalian counterparts.

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter.

Note that under some scenarios, a Medaka Multi-Generation Test (MMGT) is recommended as a possible Next Step. This test is still being validated, so it is described relatively briefly in Annex 2.

Scenarios	Result of OECD TG 230 assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	Strong evidence for *in vivo* endocrine activity in fish and other organisms	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive or developmental	An alternative approach would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.

Scenarios	Result of OECD TG 230 assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					NOEC/ECx.	
B	+	+	-	Strong evidence for *in vivo* endocrine activity in fish, despite lack of *in vivo* effects in existing tests.	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive or developmental NOEC/ECx.	An alternative approach would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs. If the negative *in vivo* data are from a fish test (*e.g.* OECD TG 229), consider possible reasons for the disparity (*e.g.* differences in species sensitivity) before conducting a lifecycle test.
C	+	+	Eq/0**	Strong evidence for *in vivo* endocrine activity in fish despite equivocal or absent *in vivo* data in other species	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive or developmental NOEC/ECx.	The decision about whether to conduct a fish one generation or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs. If no existing fish data are available, it may be worth performing TG 234 (FSDT) before a possible lifecycle test in order to obtain information on whether sexual development is a sensitive part of the lifecycle. Such information could influence the design of the lifecycle test.
D	+	-	+	Strong evidence for *in vivo* endocrine activity in fish and other species, but confidence	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive	The negative *in vitro* data suggest that the test chemical may be metabolically activated *in vivo,* or it may operate via mechanisms not covered by the *in vitro* screens. An alternative approach to a lifecycle test would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a FLCTT or MMGT may

Scenarios	Result of OECD TG 230 assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				about MOA is reduced by negative mechanistic data.	or developmental NOEC/ECx.	be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
E	+	-	-	Moderate-strong evidence for *in vivo* endocrine activity in fish, but confidence is reduced by negative *in vitro* data and negative *in vivo* activity in other species.	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive or developmental NOEC/ECx.	The negative *in vitro* data suggest that the test chemical may be metabolically activated *in vivo*, or it may operate via mechanisms not covered by the *in vitro* screens An alternative approach to a lifecycle test would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a fish one generation or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs. If the negative *in vivo* data are from a fish test (*e.g.* OECD TG 229), consider possible reasons for the disparity (*e.g.* differences in species sensitivity) before conducting a lifecycle test.
F	+	-	Eq/0	Moderate – strong evidence for *in vivo* endocrine activity in fish, but confidence is reduced by negative *in*	Consider performing a fish lifecycle test, especially if the intention is to obtain precise data on a reproductive or developmental	The negative *in vitro* data suggest that the test chemical may be metabolically activated *in vivo*, or it may operate via mechanisms not covered by the *in vitro* screens. The decision about whether to conduct a fish one generation or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.

Scenarios	Result of OECD TG 230 assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				vitro data and equivocal or absent *in vivo* activity in other species.	NOEC/ECx.	If no existing fish data are available, it may be worth performing TG 234 (FSDT) before a possible lifecycle test in order to obtain information on whether sexual development is a sensitive part of the lifecycle. Such information could influence the design of the lifecycle test.
G	+	Eq/0	+	Strong evidence for *in vivo* endocrine activity in fish, but mechanism unconfirmed.	Obtain mechanistic data and then consider performing a fish lifecycle test.	An alternative approach to a lifecycle test would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
H	+	Eq/0	-	Strong-moderate evidence for *in vivo* endocrine activity in fish, but mechanism unconfirmed.	Obtain mechanistic data and then consider performing a fish lifecycle test.	An alternative approach to a lifecycle test would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct a FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs. If the negative *in vivo* data are from a fish test (*e.g.* OECD TG 229), consider possible reasons for the disparity (*e.g.* differences in species sensitivity) before possibly conducting a lifecycle test.
I	+	Eq/0	Eq/0	Moderate evidence for *in vivo* endocrine activity in fish,	Obtain mechanistic data and then consider performing a fish	The decision about whether to conduct a FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be

Scenarios	Result of OECD TG 230 assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				but mechanism unconfirmed.	lifecycle test.	transferred to the fry via the eggs. If no existing fish data are available, it may be worth performing TG 234 (FSDT) before a possible lifecycle test in order to obtain information on whether sexual development is a sensitive part of the lifecycle. Such information could influence the design of the lifecycle test.
J	-	+	+	Based on the existing data, the chemical has endocrine activity *in vivo*. The lack of response in OECD TG 230 suggests that fish are not responsive, unless the existing data are from fish.	Consider performing TG 234 (FSDT)	It is possible that the failure to give a positive result in OECD TG 230 was caused by the relatively short exposure time (3 weeks). If this is suspected (*e.g.* the chemical only bioaccumulates slowly), or if the existing *in vivo* data are from a fish, TG 234 (FSDT) or potentially a lifecycle test would be able to study the effects of longer exposure and confirm whether there is a hazard to fish. Choice of test should be guided by the existing *in vivo* data.
K	-	+	-	There is no evidence that the chemical is a possible ED *in vivo*, probably because it is very weakly	Probably no further action, but see comments in right-hand column.	It is possible that EDs which bioaccumulate slowly may only cause effects *in vivo* after exposure times longer than 3 weeks. If this is suspected, and depending on which part of the lifecycle is suspected of being the most sensitive, consider performing TG 234 (FSDT), or a fish lifecycle test. It is also possible that the chemical may be an anti-androgen *in vivo* (consider performing the Androgenised Female Stickleback Screen - AFSS), or a thyroid-active chemical *in vivo* (consider

Scenarios	Result of OECD TG 230 assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				acting or rapidly metabolised.		performing the Amphibian Metamorphosis Assay – AMA – OECD TG 231).
L	-	+	Eq/0	The chemical may not be an ED *in vivo*, but the confidence in this conclusion is relatively low as there is only one unequivocal *in vivo* test result (a negative).	Probably no further action, but see comments in right-hand column.	If the equivocal existing data are from a fish assay, consider performing a fish assay (OECD TG 229 or 230) with a different species, or a longer-term test (TG 234 (FSDT) or lifecycle) if the chemical is a slow bioaccumulator. It is also possible that the chemical may be an anti-androgen *in vivo* (consider performing AFSS), or a thyroid-active chemical *in vivo* (consider performing an AMA– OECD TG 231).
M	-	-	+	The chemical is apparently not a possible ED in fish but it does have activity in another species.	Use the existing *in vivo* data to help decide whether a longer-term test with an appropriate fish species is indicated.	It is also possible that the chemical may be an anti-androgen *in vivo* (consider performing an AFSS), or a thyroid-active chemical *in vivo* (consider performing an AMA – OECD TG 231), although lack of *in vitro* binding affinity with the androgen receptor suggests the former is unlikely. Use the existing *in vivo* data to guide any further testing.
N	-	-	-	The chemical is probably not a possible ED *in vivo*.	No further action with respect to estrogenic, anti-estrogenic, androgenic or steroidogenic	It is still possible that the chemical may be an anti-androgen *in vivo* (consider performing AFSS), or a thyroid-active chemical *in vivo* (consider performing an AMA – OECD TG 231), although lack of *in vitro* binding affinity with the androgen receptor suggests the former is unlikely.

Scenarios	Result of OECD TG 230 assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					MOAs.	
O	-	-	Eq/0	The chemical is probably not a possible ED in fish.	Probably no further action. However, see comments in right-hand column.	If the paucity of *in vivo* data is a concern, performance of a screening test (OECD TG 229 or 230) with a different species, or a longer-term test (*i.e.* TG 234 (FSDT) or lifecycle) could be considered. It is also possible that the chemical may be an anti-androgen *in vivo* (consider performing an AFSS), or a thyroid-active chemical *in vivo* (consider performing an AMA – OECD TG 231), although lack of *in vitro* binding affinity with the androgen receptor suggests the former is unlikely.
P	-	Eq/0	+	The chemical is probably not a possible ED in fish, but confidence in this conclusion is low given the lack of mechanistic *in vitro* data and the availability of positive existing *in vivo* data.	Obtain mechanistic data, and then consider whether further testing is desirable.	If the mechanistic data confirm that the chemical has potential endocrine action, consider conducting a fish assay (OECD TG 229 or 230) with another species, or a longer term test (TG 234 (FSDT) or lifecycle). Use the existing *in vivo* data as a guide to test choice. If the mechanistic data reveal anti-androgenic or thyroid activity, consider performing an AFSS or an AMA (OECD TG 231), respectively.
Q	-	Eq/0	-	The chemical is probably not a possible ED in fish, but the	Obtain mechanistic data, and then consider whether further testing is	If the mechanistic data confirm that the chemical has potential endocrine action, consider conducting a fish assay (OECD TG 229 or 230) with another species, or a longer term test (TG 234 (FSDT) or lifecycle). Use the existing *in vivo* data as a guide to

Scenarios	Result of OECD TG 230 assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				lack of mechanistic *in vitro* data are a concern, even though the existing *in vivo* data are negative.	desirable.	test choice. If the mechanistic data reveal anti-androgenic or thyroid activity, consider performing an AFSS or an AMA (OECD TG 231), respectively.
R	-	Eq/0	Eq/0	The chemical is probably not a possible ED in fish, but confidence in this conclusion is low given the lack of mechanistic *in vitro* and existing *in vivo* data.	Obtain mechanistic data, and then consider whether further testing is desirable.	If the mechanistic data confirm that the chemical has potential endocrine action, consider conducting a fish assay (OECD TG 229 or 230) with another species, or a longer term test (TG 234 (FSDT) or lifecycle). Use the existing *in vivo* data as a guide to test choice. If the mechanistic data reveal anti-androgenic or thyroid activity, consider performing an AFSS or an AMA (OECD TG 231), respectively.

C.3.3 Variant of OECD TG 230: Androgenised Female Stickleback Screen (AFSS) (GD 140).

C.3.3.1 Modality detected/endpoints: Androgens (♀ spiggin ↑); Anti-androgens (androgenised ♀ spiggin ↓)

Background to the Assay

C.3.3.2 This assay is designed primarily as a screen for chemicals with *in vivo* anti-androgenic activity in fish but it is also able to detect androgens. It has completed validation and has been published as an OECD GD (GD 140). The endpoints are indicators of hormonal activity and there are no apical measures of adverse effects diagnostic of a specific EATS modality. This assay is a variant of the 21-Day Fish assay (OECD TG 230) with a more limited range of endpoints, but it has more power to identify anti-androgens than OECD TG 229 or OECD TG 230.

When/Why the Assay May be Used

C.3.3.3 Although the AFSS could, in principle, be used at any stage in the hazard assessment process, the most likely use scenario will be when there are relatively few data available about the possible endocrine disrupting properties of a chemical. The assay is most likely to be used either as part of a battery of *in vitro* and *in vivo* screens, or to follow up on existing data which suggest possible endocrine disruption activity at the androgen receptor. It would not be necessary for aquatic exposure to have been predicted (because a positive in the AFSS could potentially be extrapolated to terrestrial vertebrates), but such a prediction would provide additional justification for running the screen. It is also possible that no existing endocrine-relevant data are available (*i.e.* the AFSS has been used as a primary screen), but in that case a positive result in the screen should ideally be followed up with relevant *in vitro* screening in an attempt to confirm the suspected (anti)androgenic mode of action.

Existing Data to be Considered

C.3.3.4 Existing data available before deployment of the AFSS might include *in vivo* results obtained with other vertebrates (*e.g.* a positive Hershberger assay – OECD TG 441 - with rodents), or one or more of a range of *in silico* or *in vitro* results which suggest that the modalities indicated above may occur *in vivo*. Such indicators of possible *in vivo* activity might include (Q)SAR predictions of endocrine activity, 'read-across' from *in vivo* results obtained with chemically related chemicals, or positive results from an *in vitro* screen for androgen receptor-mediated activity.

Scenarios: Positive and Negative Results Combined with Existing Data

C.3.3.4a The scenarios (A to R) presented in Table C.3.3 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.3.3.5 Positive results obtained with one of the endpoints (Table C.3.3, Scenarios A-I) result in the conclusion that the test chemical is a possible androgen or anti-androgen *in vivo*. If a regulatory authority required more evidence, positive results in the AFSS should be followed up with more comprehensive testing to show whether adverse apical effects occur at any part of the lifecycle (and

hence to provide evidence supporting a conclusion that the chemical is an actual ED). In other words, to increase confidence, a positive result in the AFSS would trigger fish lifecycle testing at Level 5 (FLCTT or MMGT), or possibly a Fish Sexual Development Test (FSDT) (TG 234) at Level 4 if it is suspected that the most responsive part of the lifecycle is sexual development. Existing data suggesting (anti)androgenic activity will strengthen the case for additional testing still further.

C.3.3.6 The situation in which the AFSS gives a negative result (Table C.3.3, Scenarios J-R) needs careful consideration of any existing data. If these data suggest that the chemical is (anti)androgenic both *in vitro* and *in vivo* (Scenario J), then the probability is that the AFSS is simply insufficiently sensitive. It might in these circumstances be appropriate to conduct TG 234 (FSDT), or alternatively, a fish lifecycle test (FLCTT or MMGT) to confirm that there is no endocrine activity in fish.

C.3.3.7 If the AFSS and existing *in vivo* data are all negative, but *in vitro* data reveal some (anti)androgenic activity (Scenario K), the probability is that the test chemical is not sufficiently potent to produce endocrine effects *in vivo* in fish or other organisms, or it may be rapidly metabolised or simply does not reach the receptor. In such a situation, further testing is probably not necessary. However, if the chemical is known to bioaccumulate slowly, it may be that exposures in the *in vivo* tests have been insufficiently prolonged, in which case longer-term testing might be justified. Equally, if existing data suggest thyroid activity, consideration should be given to conducting the Amphibian Metamorphosis Assay (OECD TG 231).

C.3.3.8 On the other hand, if the AFSS and the *in vitro* tests are negative (Scenario M), but there are positive existing *in vivo* data, the chemical is probably not an ED with (anti)androgenic activity, but it may act via MOAs not covered by the *in vitro* screens, or it may be more potent in species or life-stages that have not been tested. In this situation, the existing *in vivo* data should be used to guide decisions about whether to conduct any further testing, either for modalities such as thyroid activity, or including lifestages represented in TG 234 (FSDT) or in the FLCTT.

C.3.3.9 Finally, a negative AFSS, set against a background of negative *in vitro* and *in vivo* data (Scenario N), suggests that the test chemical is not (anti)androgenic in fish, and no further testing for this modality will generally be necessary. It remains possible that it has thyroid activity, although if any existing tests for this modality are negative, it would suggest that this scenario is unlikely.

C.3.3.10 In each of the above scenarios, it is possible that existing data will be equivocal (Scenarios C, F, G, H, I, L, O, P, Q and R), or there may be no existing data. This will weaken the conclusions which can be drawn about a negative AFSS, and this is reflected in Table C.3.3 below. However, a lack of mechanistic data on (anti)androgenic activity should ideally be rectified before any further *in vivo* testing is considered. On the other hand, if the AFSS is positive, further *in vivo* testing to obtain more evidence is generally desirable even if all existing data are equivocal, or if there are no existing data. Again, however, it will always be helpful to obtain some mechanistic information before conducting further *in vivo* testing.

C.3.3.11 The scenario in which the results of the AFSS are themselves equivocal has not been dealt with in Table C.3.3, for reasons of brevity. In this context, an equivocal result might be an inconsistent concentration-response (*e.g.* no effect at a high concentration but effects at a lower), or a result which borders on statistical significance. Without knowing the exact circumstances, reliable advice cannot be given, but the opinions of an experienced ecotoxicologist should be sought. Clearly, however, such equivocal results do not necessarily rule out the existence of *in vivo* endocrine activity. For example, spiggin induction in females at a high concentration might be masked by any systemic toxicity (although it would not be sensible to run the assay at such high concentrations), while spiggin depression in androgenised females might just fail to reach a statistically significant level because spiggin levels were relatively low to begin with. If these or other possible reasons for false negatives are suspected with good reason, the screen could be repeated (*e.g.* conduct it at lower concentrations which avoid systemic toxicity, assuming systemic toxicity in the original test occurred at all concentrations), or a more appropriate version of it (*e.g.* ensure androgenised females have high spiggin levels at the start of the test) could be conducted.

C.3.3.12 In summary, positive results in the AFSS indicate that a chemical is a possible (anti)androgen. If a regulatory authority required further evidence, more comprehensive *in vivo* testing would then be necessary to produce a long-term NOEC/ECx for adverse effects and/or to confirm whether or not the chemical is an actual (anti)androgen. Negative results in the AFSS do not necessarily mean that the chemical is not a possible (anti)androgen – a judgement about this will have to be made in the light of existing *in vitro* and *in vivo* data.

Table C.3.3 Androgenised Female Stickleback Screen (AFSS) (GD 140). Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR, and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". There is no evidence at present that equivalent *in vitro* assays with systems derived from fish offer advantages over their mammalian counterparts.

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter.

Note that this assay has been successfully validated, but it has not yet been published as an OECD TG.

Note that under some scenarios, a Medaka Multi-Generation Test (MMGT) is recommended as a possible Next Step. This test is still being validated, so it is described relatively briefly in Annex 2.

Scenario	Result of AFSS	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	Strong evidence for *in vivo* (anti)androgenic activity in fish and other organisms.	Consider performing fish lifecycle test.	An alternative approach would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction.

Scenario	Result of AFSS	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						The decision about whether to conduct FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
B	+	+	-	Strong evidence for *in vivo* (anti)androgenic activity in fish.	Consider performing fish lifecycle test.	An alternative approach would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs. If the negative *in vivo* data are from a fish test (*e.g.* OECD TG 229 or 230), consider possible reasons for the disparity (*e.g.* differences in species sensitivity) before possibly conducting a lifecycle test (FLCTT or MMGT) or TG 234 (FSDT).

Scenario	Result of AFSS	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
C	+	+	Eq/0**	Strong evidence for *in vivo* (anti)androgenic activity in fish.	Consider performing fish lifecycle test or TG 234 (FSDT).	The decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs. If no existing fish data are available, it may be worth performing TG 234 (FSDT) before a lifecycle test in order to obtain information on whether sexual development is a sensitive part of the lifecycle. Such information could influence the design of a lifecycle test (FLCTT or MMGT).
D	+	-	+	Strong evidence for *in vivo* (anti)androgenic activity in fish and other organisms, but negative *in vitro* data suggest MOA may not be via interaction with the androgen receptor, or that the test chemical may be metabolically activated *in vivo*.	Consider performing fish lifecycle test.	An alternative approach would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.

Scenario	Result of AFSS	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
E	+	-	-	Strong evidence for *in vivo* (anti)androgenic activity in fish, but negative existing data raise doubts about the MOA, or suggest that the test chemical may be metabolically activated *in vivo*.	Consider performing fish lifecycle test.	An alternative approach would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs. If the negative *in vivo* data are from a fish test (*e.g.* OECD TG 229 or 230), consider possible reasons for the disparity (*e.g.* differences in species sensitivity) before possibly conducting a lifecycle test (FLCTT or MMGT) or TG 234 (FSDT).
F	+	-	Eq/0	Strong evidence for *in vivo* (anti)androgenic activity in fish, but negative or equivocal existing data raise doubts about the MOA, or suggest that the test chemical may be	Consider performing fish lifecycle test or TG 234 (FSDT).	The decision about whether to conduct a fish one generation or multi-generation test may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.

Scenario	Result of AFSS	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				metabolically activated *in vivo*.		If no existing fish data are available, it may be worth performing TG 234 (FSDT) before a lifecycle test (FLCTT or MMGT) in order to obtain information on whether sexual development is a sensitive part of the lifecycle. Such information could influence the design of a lifecycle test.
G	+	Eq/0	+	Strong evidence for *in vivo* (anti)androgenic activity in fish and other organisms, but mechanism unconfirmed.	Obtain more predictive mechanistic data and then consider performing fish lifecycle test.	An alternative approach would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
H	+	Eq/0	-	Strong evidence for *in vivo* (anti)androgenic activity in fish, but mechanism unconfirmed.	Obtain more predictive mechanistic data and then consider performing fish lifecycle test.	An alternative approach would be to deploy TG 234 (FSDT), especially if sexual development is expected to give a response at lower concentrations than reproduction. The decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0

Scenario	Result of AFSS	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs. If the negative *in vivo* data are from a fish test (*e.g.* OECD TG 229 or 230), consider possible reasons for the disparity (*e.g.* differences in species sensitivity) before possibly conducting a lifecycle test (FLCTT or MMGT) or TG 234 (FSDT).
I	+	Eq/0	Eq/0	Strong evidence for *in vivo* (anti)androgenic activity in fish, but mechanism unconfirmed.	Obtain more predictive mechanistic data and then consider performing fish lifecycle test or TG 234 (FSDT).	The decision about whether to conduct FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs. If no existing fish data are available, it may be worth performing TG 234 (FSDT) before a lifecycle test in order to obtain information on whether sexual development is a sensitive part of the lifecycle. Such information could influence the design of a lifecycle test (FLCTT or MMGT).
J	-	+	+	No evidence for (anti)androgenic activity	Consider performing TG 234 (FSDT).	It is possible that the failure to give a positive result in the AFSS was caused by

Scenario	Result of AFSS	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				in vivo in fish. However, the chemical is an (anti)androgen in other species and this mechanism has been confirmed *in vitro*.		the relatively short exposure time (3 weeks). If this is suspected, it is worth considering whether to perform a fish lifecycle test (FLCTT or MMGT) or TG 234 (FSDT). Test design should be guided by the existing *in vivo* data.
K	-	+	-	There is no evidence that the chemical is an (anti)androgen *in vivo*, probably because it is very weakly acting or rapidly metabolised.	Probably no further action, but see comments in right-hand column.	It is possible that EDs which bioaccumulate slowly may only cause effects *in vivo* after exposure times longer than 3 weeks. If this is suspected, and depending on which part of the lifecycle is suspected of being the most sensitive, consider performing TG 234 (FSDT) or a fish lifecycle test. It is also possible that the chemical may be a thyroid-active chemical *in vivo* (consider performing the Amphibian Metamorphosis Assay – OECD TG 231), an (anti)estrogen, or an aromatase inhibitor (consider performing OECD TG 229 or 230).
L	-	+	Eq/0	The chemical may not be an (anti)androgen *in vivo*, but the confidence in this conclusion is relatively low as there is only one unequivocal *in vivo* test result (a negative).	Consider performing a fish assay (OECD TG 229 or 230) with a different species, or consider a longer-term test (TG 234 (FSDT) or lifecycle).	It is also possible that the chemical may be a thyroid-active chemical *in vivo* (consider performing the Amphibian Metamorphosis Assay – OECD TG 231), an (anti)estrogen, or an aromatase inhibitor (consider performing OECD TG 229 or 230).

Scenario	Result of AFSS	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
M	-	-	+	The chemical is probably not an (anti)androgen in fish. However, it may act through MOAs not covered by the available *in vitro* assays, or it may be more potent in a species other than that tested, or over a longer exposure period.	Use the existing *in vivo* data to help choose a possible longer-term test with an appropriate species.	It is also possible that the chemical may be a thyroid-active chemical *in vivo* (consider performing the Amphibian Metamorphosis Assay – OECD TG 231), an (anti)estrogen, or an aromatase inhibitor (consider performing OECD TG 229 or 230), although lack of *in vitro* binding affinity with the estrogen or androgen receptors suggests the 2 former possibilities are unlikely. Use the existing *in vivo* data to guide any further testing.
N	-	-	-	The chemical is probably not an (anti)androgen in fish or other organisms.	No further action with respect to (anti)androgenic MOAs.	It is still possible that the chemical may be a thyroid-active chemical *in vivo* (consider performing the Amphibian Metamorphosis Assay – OECD TG 231), an (anti)estrogen, or an aromatase inhibitor (consider performing OECD TG 229 or 230), although lack of *in vitro* binding affinity with the estrogen or, androgen receptors suggests the 2 former possibilities are unlikely.
O	-	-	Eq/0	The chemical is probably not an (anti)androgen in fish or other organisms.	Probably no further action. However, see comments in right-hand column.	If the paucity of *in vivo* data are a concern, performance of a screening test (OECD TG 229 or 230) with a different species, or a longer-term test (*i.e.* TG 234 (FSDT) or lifecycle) could be considered.

Scenario	Result of AFSS	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						It is also possible that the chemical may be a thyroid-active chemical *in vivo* (consider performing the Amphibian Metamorphosis Assay – OECD TG 231), an (anti)estrogen, or an aromatase inhibitor (consider performing OECD TG 229 or 230), although lack of *in vitro* binding affinity with the estrogen or, androgen receptors suggests the 2 former possibilities are unlikely.
P	-	Eq/0	+	The chemical is probably not an (anti)androgen in fish, but confidence in this conclusion is low given the lack of more predictive *in vitro* data and the availability of positive existing *in vivo* data.	Obtain more predictive mechanistic data, and then consider possible further testing.	If the mechanistic data confirm that the chemical has potential (anti)androgenic action, consider conducting a fish assay (OECD TG 229 or 230) with another species, or a longer term test (TG 234 (FSDT) or lifecycle). Use the existing *in vivo* data as a guide to test design. If the mechanistic data reveal (anti)estrogenic/aromatase inhibition activity, perform a fish assay (OECD TG 229 or 230). If any existing data suggest thyroid activity, consider an Amphibian Metamorphosis Assay (OECD TG 231).
Q	-	Eq/0	-	The chemical is probably not an (anti)androgen in fish or other organisms, but the lack of more	Obtain more predictive mechanistic data, and then consider possible further testing.	If the mechanistic data confirm that the chemical has potential (anti)androgenic action, consider conducting a fish assay (OECD TG 229 or 230) with another

Scenario	Result of AFSS	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				predictive mechanistic data are a concern.		species, or a longer term test (TG 234 (FSDT) or lifecycle). Use the existing *in vivo* data as a guide to test design. If the mechanistic data reveal (anti)estrogenic/aromatase inhibition activity, perform a fish assay (OECD TG 229 or 230). If any existing data suggest thyroid activity, consider an Amphibian Metamorphosis Assay (OECD TG 231).
R	-	Eq/0	Eq/0	The chemical is probably not an (anti)androgen in fish, but confidence in this conclusion is low given the lack of more predictive *in vitro* and existing *in vivo* data.	Obtain more predictive mechanistic data, and then consider possible further testing.	If the mechanistic data confirm that the chemical has potential (anti)androgenic action, consider conducting a fish assay (OECD TG 229 or 230) with another species, or a longer term test (TG 234 (FSDT) or lifecycle). Use the existing *in vivo* data as a guide to test design. If the mechanistic data reveal (anti)estrogenic/aromatase inhibition activity, perform a fish assay (OECD TG 229 or 230). If any existing data suggest thyroid activity, consider an Amphibian Metamorphosis Assay (OECD TG 231).

C.3.4 Fish Sexual Development Test (FSDT) (OECD TG 234).

C.3.4.1 Modality detected/endpoints: Estrogens (♀ and ♂VTG ↑; phenotypic sex ratio ♀↑); Anti-estrogens (♀VTG↓; phenotypic sex ratio ♂↑; sexually undifferentiated fish ↑); Androgens (phenotypic sex ratio ♂↑; ♀ VTG↓); Anti-androgens (intersex fish ↑; ♀VTG ↑; phenotypic sex ratio ♀↑); Aromatase inhibitors (♀ VTG↓; phenotypic sex ratio ♂↑); (Optional endpoints – gonadal histopathology; genetic sex in medaka and stickleback). TG 234 (FSDT) has now been fully validated for Japanese medaka, zebrafish and stickleback.

Background to the Assay

C.3.4.2 This partial lifecycle assay could potentially be used as a screen for the types of *in vivo* endocrine disruption activity in fish which are listed above (although it is considerably more expensive and time-consuming than the OECD TG 229 or 230 screens), but should generally be used as a test which can also provide apical information of use in environmental risk assessments. It includes an endpoint (altered sex ratio), which is probably indicative of endocrine action, but more importantly indicates that adverse apical effects on sexual development are occurring. Major effects on phenotypic sex ratio would be expected to damage the ability of a fish population to reproduce itself although small effects may be tolerated, but it is not possible to define the precise change in sex ratio beyond which adverse effects will occur unless specific information about a particular population is available. It should be noted that if the assay gives a positive result, this may be due to a positive indicator of hormonal activity (*e.g.* VTG), a positive for biased sex ratio, or a positive for both types of endpoint. Each of these three possible combinations of positive response should be considered separately (although the distinctions between indicators of hormonal activity and apical effects are not as clear in TG 234 (FSDT) as in other tests because it is acknowledged that sex ratio is both an apical endpoint (relevant for populations) as well as a biomarker endpoint (indicative ofmode of action)), so they have been listed individually as points 1, 2 and 3 in the Possible Conclusions column of Table C.3.4.

C.3.4.3 If only 3 test concentrations are employed, a reliable NOEC or ECx for biased sex ratio may not be obtainable, so it is desirable to use at least 5 test concentrations if it is intended to employ the data in a risk assessment. However, if the test is used for hazard or risk assessment, the stickleback should not be used because the validation data available so far show that in this species alterations of phenotypic sex ratio by test substances are uncommon. It should be noted that simultaneous measurement of both phenotypic and genotypic sex ratio (currently only possible in medaka and stickleback) will tend to provide a more robust result and so will require fewer replicates to give adequate statistical power. However, power analyses indicate that adequate power can be achieved with zebrafish as long as sufficient replication and fish per replicate are used (OECD, 2011b).

When/Why the Assay May be Used

C.3.4.4 Although TG 234 (FSDT) could, in principle, be used at any stage in the hazard assessment process, the most likely use scenario will be when there are already some *in vitro* or *in vivo* screening data available about the possible endocrine disrupting properties of a chemical. It is unlikely that no other existing endocrine-relevant data will be available (*i.e.* if TG 234 has been used as a primary screen), but in that case a positive result in TG 234 should ideally be followed up with relevant *in vitro* screening to confirm the suspected mode of action before any other *in vivo* testing is considered.

Existing Data to be Considered

C.3.4.5 Existing data available before deployment of TG 234 (FSDT) might include *in vivo* results obtained with other vertebrates (*e.g.* a positive uterotrophic assay with rodents; or positive result in the fish assays OECD TG 229 or 230), or one or more of a range of *in silico* or *in vitro* results which suggest that the modalities indicated above may occur *in vivo*. Such indicators of possible *in vivo* activity might include (Q)SAR predictions of endocrine activity, 'read-across' from *in vivo* results obtained with chemically related chemicals, or positive results from an *in vitro* screen for estrogen or

androgen receptor-mediated activity, or for effects on steroidogenesis (especially aromatase inhibition). Conduct of TG 234 (FSDT) would be particularly relevant if the test chemical is suspected to act primarily on the sexual development phase of the fish lifecycle (as opposed to the reproductive phase), because it provides apical information on phenotypic sex ratio which is fixed during the fry or juvenile stages of the species used in this test.

Scenarios: Positive and Negative Results Combined with Existing Data

C.3.4.5a The scenarios (A to R) presented in Table C.3.4 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.3.4.6 Positive results obtained with one or more of TG 234 (FSDT) indicators of hormonal activity but not with apical endpoints (Table C.3.4, Scenarios A-I, sub-section 2) result in the conclusion that the test chemical is a possible ED *in vivo*. If both an indicator of hormonal activity and sex ratio[3] give a correlated response (Table C.3.4, Scenarios A-I, sub-section 1), this provides evidence that the chemical is almost certainly an actual ED (*i.e.* it causes adverse effects through an endocrine mechanism). If only sex ratio responds (Table C.3.4, Scenarios A-I, sub-section 3), it indicates that the chemical is probably an ED, but before drawing that conclusion, existing *in vitro* and *in vivo* data should be considered and a weight-of-evidence assessment carried out.

C.3.4.7 As indicated above, an effect on sex ratio in TG 234 (FSDT) shows that the test chemical causes an adverse apical effect, is a developmental toxicant, and is probably also an ED (assuming that the concentration giving this response is not sufficiently high to cause systemic toxicity). If these results are combined with positive indicators of hormonal activity and/or positive *in vitro* screening assay data, some regulatory authorities may consider that this is sufficient to show the chemical is an ED, and/or that the information could be used in a risk assessment (providing sufficient concentrations have been tested to give an acceptably precise NOEC or ECx). Other authorities might nevertheless require further data to demonstrate that adverse effects at lower concentrations do not occur during the reproductive phase of the lifecycle, and in these circumstances, conduct of a fish lifecycle test (FLCTT or MMGT) would be appropriate. In principle, an extended version of OECD TG 229 (*i.e*. a Fish Reproduction Partial Lifecycle Assay) might also address this issue, but a suitable protocol for this has not been validated. Additional testing of this type might also be required if an indicator or indicators of hormonal activity in TG 234 (FSDT), but not sex ratio, respond positively. Existing data suggesting endocrine activity would strengthen the case for any additional testing still further.

C.3.4.8 A situation in which TG 234 (FSDT) gives a negative result needs careful consideration of any existing data. If these data suggest that the chemical is endocrine-active both *in vitro* and *in vivo* (Table C.3.4, Scenario J), then the probability is that TG 234 (FSDT) is simply insufficiently sensitive, perhaps because the main MOA acts during the reproductive phase of the lifecycle. It might then be appropriate to conduct a fish lifecycle test (FLCTT or MMGT) to confirm that there is no adverse endocrine activity in fish.

C.3.4.9 If TG 234 (FSDT) and existing *in vivo* data are all negative, but *in vitro* data reveal some endocrine activity (Scenario K), the probability is that the test chemical is not sufficiently potent to produce endocrine effects *in vivo* in fish, or it may be rapidly metabolised. In such a situation, further

[3] Note that sex ratio can be considered as an indicator or biomarker of endocrine activity in its own right, as well as an apical measurement of adverse effects, although some types of non-EDC may hypothetically be able to affect this endpoint in some species. None of these non-EDCs have yet been found.

testing is probably not necessary. However, if there is good reason to believe that the reproductive part of the lifecycle may be more responsive than sexual development, consider conducting OECD TG 229 or a lifecycle test.

C.3.4.10 Furthermore, if TG 234 (FSDT) and the *in vitro* tests are negative, but there are positive existing *in vivo* data (Scenario M), the chemical is probably not an ED acting on fish sexual development, but it may act via MOAs not covered by the *in vitro* screens, or it may be more potent in species or life-stages that have not been tested. In this situation, the existing *in vivo* data should be used to guide decisions about whether to conduct any further testing, either for modalities such as thyroid activity (*e.g.* OECD TG 231), or including other lifestages represented in OECD TG 229 or the FLCTT.

C.3.4.11 Finally, a negative TG 234 (FSDT), set against a background of negative *in vitro* and *in vivo* data (Scenario N), suggests that the test chemical is probably not an ED acting on sexual development in fish, and no further testing for estrogenic, anti-estrogenic, androgenic or steroidogenic MOAs should generally be considered unless there is reason to believe that reproduction may be more responsive than development. It remains possible that the chemical has thyroid activity, but this is unlikely if OECD TG 231 is one of the negative *in vivo* assays.

C.3.4.12 In each of the above scenarios, it is possible that existing data will be equivocal, or there may be no existing data (Scenarios C, F, G, H, I, L, O, P, Q and R). This will weaken the conclusions which can be drawn about a negative TG 234 (FSDT). However, a lack of mechanistic data on endocrine activity should ideally be rectified before any further *in vivo* testing is finally rejected. On the other hand, if TG 234 (FSDT) is positive, further *in vivo* testing may be needed even if all existing data are equivocal, or if there are no existing data. Again, however, it will always be desirable to obtain some mechanistic information before conducting further *in vivo* testing.

C.3.4.13 The scenario in which the results of TG 234 (FSDT) are themselves equivocal has not been dealt with in Table C.3.4, for reasons of brevity. In this context, an equivocal result might be a non-monotonic concentration-response (*e.g.* no effect at a high concentration but effects at a lower), or a result which borders on statistical significance. Without knowing the exact circumstances, reliable advice cannot be given, but the opinions of an experienced ecotoxicologist should be sought. Clearly, however, such equivocal results do not necessarily rule out the existence of *in vivo* endocrine activity. For example, an effect on sex ratio might just fail to reach a statistically significant level due to a random imbalance in the control sex ratio. If these or other possible reasons for false negatives are suspected with good reason, the test could be repeated (*e.g.* conduct it at lower concentrations which avoid systemic toxicity), or a more appropriate version of it (*e.g.* more fish per replicate) could be conducted.

C.3.4.14 In summary, an adverse apical response (*i.e.* biased sex ratio) in TG 234 (FSDT) indicates that a chemical is a probable ED. A combination of biased sex ratio and a positive endocrine-responsive mechanistic endpoint (*e.g.* vitellogenin) is even stronger evidence that the chemical is an actual ED. If sufficient test concentrations have been tested, this will allow a precise NOEC or ECx to be calculated. In such cases, some regulatory authorities may consider that no more data are required, while others may wish to investigate whether the reproductive stage of the lifecycle is even more sensitive than the developmental part. On the other hand, negative results in TG 234 (FSDT) do not necessarily mean that the chemical is not an ED – a judgement about this will have to be made in the light of existing *in vitro* and *in vivo* data.

Table C.3.4 Fish Sexual Development Test (FSDT) (OECD TG 234). Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

The assay under discussion could either be positive for both apical and indicators of hormonal activity endpoints, or positive just for apical endpoints, or positive just for indicators of hormonal activity. However, note that sex ratio could in most cases be considered as both an indicator of hormonal activity and an apical endpoint, and as yet, no chemicals have been found which are able to alter sex ratios by way of mechanisms other than endocrine disruption. For each scenario, each of these 3 possibilities is addressed separately in the Possible Conclusions column.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR, and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". There is no evidence at present that equivalent *in vitro* assays with systems derived from fish offer advantages over their mammalian counterparts.

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter.

Note that under some scenarios, a Medaka Multi-Generation Test (MMGT) is recommended as a possible Next Step. This test is still being validated, so it is described relatively briefly in Annex 2.

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	1) Strong evidence for adverse effects in fish and other organisms by an endocrine mechanism. 2) Strong evidence for endocrine effects, but uncertainty about whether they are adverse in fish. 3) Strong evidence for adverse effects in fish and other organisms, but mechanism may hypothetically not be via direct interaction with ER, or by aromatase inhibition, even though it is noted that currently there is no	Some regulatory authorities may consider that further evidence is not required, especially if adverse effects have been demonstrated. However, if more evidence is needed about adverse effects in fish, performance of a fish lifecycle test should be considered.	If TG 234 (FSDT) was only performed with 3 test concentrations, this may not be sufficiently precise to establish a reliable NOEC/ECx. Also, note that some EDs may be more toxic to reproduction than to sexual development, in which case TG 234 (FSDT) would be less responsive than a lifecycle test. A decision about whether to conduct FLCTT or MMLC may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**	**1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**		
				evidence for sex ratio change in fish caused by other mechanisms than those mentioned here at otherwise non-toxic concentrations of chemicals.		
B	+	+	-	1) Strong evidence for adverse effects in fish by an endocrine mechanism. 2) Strong evidence for endocrine effects in fish, but uncertainty about whether they are adverse. 3) Strong evidence for adverse effects in fish, but mechanism may	Some regulatory authorities may consider that further evidence is not required, especially if adverse effects have been demonstrated. However, if more evidence is needed about adverse effects in fish, performance of a fish lifecycle test	If TG 234 (FSDT) was only performed with 3 test concentrations, this may not be sufficiently precise to establish a reliable NOEC/ECx. Also, note that some EDs may be more toxic to reproduction than to sexual development, in which case TG 234 (FSDT) would be less responsive than a lifecycle test. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				hypothetically not be via direct interaction with ER, or by aromatase inhibition, even though it is noted that currently there is no evidence for sex ratio change in fish caused by other mechanisms than those mentioned here at otherwise non-toxic concentrations of chemicals.	should be considered.	chemical is not expected to be transferred to the fry via the eggs.
C	+	+	Eq/0**	1) Strong evidence for adverse effects in fish by an endocrine mechanism. 2) Strong evidence for endocrine effects in fish, but uncertainty	Some regulatory authorities may consider that further evidence is not required, especially if adverse effects have been demonstrated.	If TG 234 (FSDT) was only performed with 3 test concentrations, this may not be sufficiently precise to establish a reliable NOEC/ECx. Also, note that some EDs may be more toxic to reproduction than to sexual development, in which case TG 234 (FSDT) would be less responsive than a lifecycle test.

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				about whether they are adverse. 3) Strong evidence for adverse effects in fish, but mechanism may hypothetically not be via direct interaction with ER, or by aromatase inhibition, even though it is noted that currently there is no evidence for sex ratio change in fish caused by other mechanisms than those mentioned here at otherwise non-toxic concentrations of chemicals.	However, if more evidence is needed about adverse effects in fish, performance of a fish lifecycle test should be considered. This would be particularly helpful given the equivocal *in vivo* effects, or lack of *in vivo* tests, in other taxa.	A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
D	+	-	+	1) Strong evidence for	Some regulatory	If TG 234 (FSDT) was only performed with 3

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				adverse effects in fish and other organisms, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in fish, but they do not appear to be adverse. 3) Strong evidence for adverse effects in fish and other organisms, but mechanism may hypothetically not be via direct interaction with ER, or by aromatase inhibition, even though it is noted that currently there is no evidence for sex ratio	authorities may consider that further evidence is not required, especially if adverse effects have been demonstrated. However, if more evidence is needed about adverse effects in fish, performance of a fish lifecycle test should be considered.	test concentrations, this may not be sufficiently precise to establish a reliable NOEC/ECx. Also, note that some EDs may be more toxic to reproduction than to sexual development, in which case TG 234 (FSDT) would be less responsive than a lifecycle test. If *in vitro* data are negative or equivocal, it might be unsafe to conclude that an effect on sex ratio was definitely caused by endocrine disruption, although this is the most probable explanation, especially if endocrine disruption has been shown in other species. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				change in fish caused by other mechanisms than those mentioned here at otherwise non-toxic concentrations of chemicals.		
E	+	-	-	1) Strong evidence for adverse effects in fish, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in fish, but they do not appear to be adverse. 3) Strong evidence for adverse effects in fish, but mechanism may hypothetically not be	Some regulatory authorities may consider that further evidence is not required, especially if adverse effects have been demonstrated. However, if more evidence is needed about adverse effects in fish, performance of a fish lifecycle test should be considered.	If TG 234 (FSDT) was only performed with 3 test concentrations, this may not be sufficiently precise to establish a reliable NOEC/ECx. Also, note that some EDs may be more toxic to reproduction than to sexual development, in which case TG 234 (FSDT) would be less responsive than a lifecycle test. If *in vitro* data are negative or equivocal, it might be unsafe to conclude that an effect on sex ratio was definitely caused by endocrine disruption, although this is the most probable explanation, especially if endocrine disruption has been shown in other species.

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				via direct interaction with ER, or by aromatase inhibition, even though it is noted that currently there is no evidence for sex ratio change in fish caused by other mechanisms than those mentioned here at otherwise non-toxic concentrations of chemicals.		A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
F	+	-	Eq/0	1) Strong evidence for adverse effects in fish, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in fish, but they do not appear to be	Some regulatory authorities may consider that further evidence is not required, especially if adverse effects have been demonstrated. However, if more	If TG 234 (FSDT) was only performed with 3 test concentrations, this may not be sufficiently precise to establish a reliable NOEC/ECx. Also, note that some EDs may be more toxic to reproduction than to sexual development, in which case TG 234 (FSDT) would be less responsive than a lifecycle test. If *in vitro* data are negative or equivocal, it

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				adverse. 3) Strong evidence for adverse effects in fish, but mechanism may hypothetically not be via direct interaction with ER, or by aromatase inhibition, even though it is noted that currently there is no evidence for sex ratio change in fish caused by other mechanisms than those mentioned here at otherwise non-toxic concentrations of chemicals.	evidence is needed about adverse effects in fish, performance of a fish lifecycle test should be considered. This would be particularly helpful given the equivocal *in vivo* effects, or lack of *in vivo* tests, in other taxa.	might be unsafe to conclude that an effect on sex ratio was definitely caused by endocrine disruption, although this seems the most probable explanation, especially if endocrine disruption has been shown in other species. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
G	+	Eq/0	+	1) Strong evidence for adverse effects in more	Some regulatory authorities may	If TG 234 (FSDT) was only performed with 3 test concentrations, this may not be

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				than one organism, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects, but they do not appear to be adverse in fish. 3) Strong evidence for adverse effects in fish and other organisms, but mechanism may hypothetically not be via direct interaction with ER, or by aromatase inhibition, even though it is noted that currently there is no evidence for sex ratio change in fish caused	consider that further evidence is not required, especially if adverse effects have been demonstrated. However, if more evidence is needed about adverse effects in fish, performance of a fish lifecycle test should be considered. Given uncertainty about the mechanism of action, any further *in vivo* testing should be preceded by *in vitro* mechanistic studies.	sufficiently precise to establish a reliable NOEC/ECx. Also, note that some EDs may be more toxic to reproduction than to sexual development, in which case TG 234 (FSDT) would be less responsive than a lifecycle test. If *in vitro* data are negative or equivocal, it might be unsafe to conclude that an effect on sex ratio was definitely caused by endocrine disruption, although this seems the most probable explanation, especially if endocrine disruption has been shown in other species. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				by other mechanisms than those mentioned here at otherwise non-toxic concentrations of chemicals..		
H	+	Eq/0	-	1) Strong evidence for adverse effects in fish, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in fish, but they do not appear to be adverse. 3) Strong evidence for adverse effects in fish, but mechanism may hypothetically not be via direct interaction	Some regulatory authorities may consider that further evidence is not required, especially if adverse effects have been demonstrated. However, if more evidence is needed about adverse effects in fish, performance of a fish lifecycle test should be considered. Given uncertainty	If TG 234 (FSDT) was only performed with 3 test concentrations, this may not be sufficiently precise to establish a reliable NOEC/ECx. Also, note that some EDs may be more toxic to reproduction than to sexual development, in which case TG 234 (FSDT) would be less responsive than a lifecycle test. If *in vitro* data are negative or equivocal, it might be unsafe to conclude that an effect on sex ratio was definitely caused by endocrine disruption, although this seems the most probable explanation, especially if endocrine disruption has been shown in other species. A decision about whether to conduct an

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				with ER, or by aromatase inhibition, even though it is noted that currently there is no evidence for sex ratio change in fish caused by other mechanisms than those mentioned here at otherwise non-toxic concentrations of chemicals..	about the mechanism of action, any further *in vivo* testing should be preceded by *in vitro* mechanistic studies.	FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
I	+	Eq/0	Eq/0	1) Strong evidence for adverse effects in fish, possibly by an unknown endocrine mechanism. 2) Moderate-Strong evidence for endocrine effects in fish, but they do not appear to be adverse.	Some regulatory authorities may consider that further evidence is not required, especially if adverse effects have been demonstrated. However, if more evidence is needed	If TG 234 (FSDT) was only performed with 3 test concentrations, this may not be sufficiently precise to establish a reliable NOEC/ECx. Also, note that some EDs may be more toxic to reproduction than to sexual development, in which case TG 234 (FSDT) would be less responsive than a lifecycle test. If *in vitro* data are negative or equivocal, it

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				3) Strong evidence for adverse effects in fish, but mechanism may hypothetically not be via direct interaction with ER, or by aromatase inhibition, even though it is noted that currently there is no evidence for sex ratio change in fish caused by other mechanisms than those mentioned here at otherwise non-toxic concentrations of chemicals..	about adverse effects in fish, performance of a fish lifecycle test should be considered. Given uncertainty about the mechanism of action, any further *in vivo* testing should be preceded by *in vitro* mechanistic studies.	might be unsafe to conclude that an effect on sex ratio was definitely caused by endocrine disruption, although this seems the most probable explanation, especially if endocrine disruption has been shown in other species. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
J	-	+	+	The chemical is an ED *in vivo* in other species but does not appear to	Some regulatory authorities may consider that further	As OECD TG 229 only uses 3 test concentrations and exposes fish for just 3 weeks, an extended version which runs more

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				act on sexual development in fish. If any other fish tests are also negative, fish may not be responsive at all to the test chemical.	evidence is not required. However, if it is suspected that the reproductive part of the lifecycle may be responsive, consider conducting OECD TG 229 or a fish lifecycle test.	concentrations for longer would provide more comprehensive data about interference with reproduction than OECD TG 229 unmodified. However, an agreed protocol for such an extended test is not available, so an option would be to run a fish lifecycle test. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
K	-	+	-	Despite the *in vitro* mechanistic data for potential endocrine activity, there is no evidence for endocrine disruption *in vivo*. This may be because the	Some regulatory authorities may consider that sufficient data are available to show that the chemical is not an ED *in vivo*. However, if it is	As OECD TG 229 only uses 3 test concentrations and exposes fish for just 3 weeks, an extended version which runs more concentrations for longer would provide more comprehensive data about interference with reproduction than OECD TG 229 unmodified. However, an agreed protocol for such an

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				chemical is d*egra*ded to an inactive metabolite, or because it only interacts very weakly with endocrine receptors. However, it is also possible that the chemical only acts on the reproductive part of the fish lifecycle which is not exposed inTG 234 (FSDT).	suspected that the reproductive part of the lifecycle may be responsive, consider conducting either OECD TG 229 or a fish lifecycle test.	extended test is not available, so an option would be to run a fish lifecycle test. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
L	-	+	Eq/0	The chemical may not be an ED *in vivo*, but the confidence in this conclusion is relatively low as there is only one unequivocal *in vivo* test result (the negative TG	Some regulatory authorities may consider that sufficient data are available to show that the chemical is not an ED *in vivo*. However, such a	As OECD TG 229 only uses 3 test concentrations and exposes fish for just 3 weeks, an extended version which runs more concentrations for longer would provide more comprehensive data about interference with reproduction than OECD TG 229 unmodified. However, an agreed protocol for such an

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				234). However, it is also possible that the chemical only acts on the reproductive part of the fish lifecycle which is not exposed in TG 234 (FSDT).	conclusion is not well-supported. If it is suspected that the reproductive part of the lifecycle may be responsive, consider conducting either OECD TG 229 or a fish lifecycle test.	extended test is not available, so an option would be to run a fish lifecycle test. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
M	-	-	+	The chemical is probably not an ED acting on sexual development in fish, but it does have endocrine activity in other species. However, it may act through MOAs not covered by the available *in vitro* assays, or it may be more potent in a	Some regulatory authorities may consider that sufficient evidence is available. However, if it is suspected that the reproductive part of the lifecycle may be responsive, consider conducting either OECD TG 229 or a	As OECD TG 229 only uses 3 test concentrations and exposes fish for just 3 weeks, an extended version which runs more concentrations for longer would provide more comprehensive data about interference with reproduction than OECD TG 229 unmodified. However, an agreed protocol for such an extended test is not available, so an option would be to run a fish lifecycle test. A decision about whether to conduct an FLCTT or MMGT may be driven primarily

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				fish species other than that tested. It is also possible that the chemical only acts on the reproductive part of the fish lifecycle which is not exposed inTG 234 (FSDT), although such action is presumably not via one of the mechanisms mentioned above.	fish lifecycle test, possibly using a different species to that employed inTG 234 (FSDT).	by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
N	-	-	-	The chemical is probably not an ED acting on sexual development in fish, or *in vivo* in other species. It is possible that the chemical is able to	Some regulatory authorities may consider that sufficient data are available to show that the chemical is not an ED *in vivo*. However, if it is	As OECD TG 229 only uses 3 test concentrations and exposes fish for just 3 weeks, an extended version which runs more concentrations for longer would provide more comprehensive data about interference with reproduction than OECD TG 229 unmodified. However, an agreed protocol for such an

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				interfere with the reproductive part of the fish lifecycle but the probability of this is low given the apparent absence of estrogenic, androgenic or steroidogenic properties *in vitro* or *in vivo*.	suspected that the reproductive part of the lifecycle may be responsive, consider conducting either OECD TG 229 or a fish lifecycle test.	extended test is not available, so an option would be to run a fish lifecycle test. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
O	-	-	Eq/0	The chemical is probably not an ED acting on sexual development in fish. It is possible that the chemical is able to interfere with the reproductive part of the fish lifecycle, but the probability of this is low given the apparent	Some regulatory authorities may consider that sufficient data are available to show that the chemical is not an ED *in vivo*. However, if it is suspected that the reproductive part of the lifecycle may be responsive, consider	As OECD TG 229 only uses 3 test concentrations and exposes fish for just 3 weeks, an extended version which runs more concentrations for longer would provide more comprehensive data about interference with reproduction than OECD TG 229 unmodified. However, an agreed protocol for such an extended test is not available, so an option would be to run a fish lifecycle test. A decision about whether to conduct an FLCTT or MMGT may be driven primarily

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				absence of estrogenic, androgenic or steroidogenic properties.	conducting either OECD TG 229 or a fish lifecycle test.	by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to the fry via the eggs.
P	-	Eq/0	+	The chemical is probably not an ED acting on sexual development in fish, but confidence in this conclusion is low given the lack of comprehensive *in vitro* data and the availability of positive existing *in vivo* data. However, it is possible that the chemical only acts on the reproductive part of	Some regulatory authorities may consider that sufficient evidence is available. However, if it is suspected that the reproductive part of the lifecycle may be responsive, consider conducting either OECD TG 229 or a fish lifecycle test. However, it would be desirable to obtain	As OECD TG 229 only uses 3 test concentrations and exposes fish for just 3 weeks, an extended version which runs more concentrations for longer would provide more comprehensive data about interference with reproduction than OECD TG 229 unmodified. However, an agreed protocol for such an extended test is not available, so an option would be to run a fish lifecycle test. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				the fish lifecycle which is not exposed inTG 234 (FSDT).	comprehensive mechanistic data before possibly proceeding to further *in vivo* testing.	chemical is not expected to be transferred to the fry via the eggs.
Q	-	Eq/0	-	The chemical is probably not an ED acting on sexual development in fish, or *in vivo* on other species, but the lack of more predictive mechanistic data are a concern, even though the existing *in vivo* data are negative. It is nevertheless possible that the chemical is able to interfere with the reproductive part of the	Some regulatory authorities may consider that sufficient data are available to show that the chemical is not an ED *in vivo*. However, if it is suspected that the reproductive part of the lifecycle may be responsive, consider conducting either OECD TG 229 or a fish lifecycle test. It would be desirable to	As OECD TG 229 only uses 3 test concentrations and exposes fish for just 3 weeks, an extended version which runs more concentrations for longer would provide more comprehensive data about interference with reproduction than OECD TG 229 unmodified. However, an agreed protocol for such an extended test is not available, so an option would be to run a fish lifecycle test. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				fish lifecycle.	obtain comprehensive mechanistic data before any further *in vivo* testing.	the fry via the eggs.
R	-	Eq/0	Eq/0	The chemical may not be an ED acting on sexual development in fish, but confidence in this conclusion is low given the lack of comprehensive *in vitro* and existing *in vivo* data. It is nevertheless possible that the chemical is able to interfere with the reproductive part of the fish lifecycle.	Some regulatory authorities may consider that sufficient data are available to show that the chemical is not an ED *in vivo*. However, if it is suspected that the reproductive part of the lifecycle may be responsive, consider conducting either OECD TG 229 or a fish lifecycle test. However, it would be	As OECD TG 229 only uses 3 test concentrations and exposes fish for just 3 weeks, an extended version which runs more concentrations for longer would provide more comprehensive data about interference with reproduction than OECD TG 229 unmodified. However, an agreed protocol for such an extended test is not available, so an option would be to run a fish lifecycle test. A decision about whether to conduct an FLCTT or MMGT may be driven primarily by the bioaccumulative properties of the chemical – a one generation test (*e.g.* F0 eggs to F1 fry) will generally be sufficient if the chemical is not expected to be transferred to

Scenario	Result of TG 234 (FSDT)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					desirable to obtain comprehensive mechanistic data before any further *in vivo* testing.	the fry via the eggs.

C.3.5 Fish Lifecycle Toxicity Test (FLCTT) (USEPA OPPTS 850.1500).

C.3.5.1 Modality detected/endpoints: The basic FLCTT as described by Benoit (1981), USEPA (1996) and others does not contain endpoints which solely respond to endocrine disrupters. However, many of the endpoints in this apical test are nevertheless affected by EATS EDs. Of particular interest in the context of estrogens, androgens and steroidogenesis disrupters are time to sexual maturity, sex ratio of adults, fecundity and fertility, but other endpoints may also be responsive to some EDs (*e.g.* growth may respond to some thyroid disrupters). It should be noted that no cases are known in which altered sex ratio was caused by a substance other than an ED.

Background to the Assay

C.3.5.2 This assay is designed primarily as an apical test for chemicals with suspected reproductive or long-term toxicity. It has not been adopted for publication as a OECD TG, but has been widely used for several decades by regulatory agencies for assessing possible chronic effects in fish. The endpoints are all apical measures of development, growth or reproduction. Exposure of the test organisms (fathead minnow *Pimephales promelas*, in the case of Benoit 1981, but other species can be successfully used with minor changes in the protocol, including sheepshead minnow *Cyprinodon variegatus*, zebrafish *Danio rerio*, and medaka *Oryzias latipes*) usually continues from the freshly fertilised eggs of the F0 generation to the fry or young fish of the F1 generation (4-8 weeks post-hatch in the case of fathead minnow – Benoit, 1981).

C.3.5.3 It should be noted that it would be relatively straightforward to include ED-specific endpoints in this test. Depending on the species and test objectives, these could include *inter alia* sex hormones, thyroid hormones, vitellogenin, spiggin, secondary sex characteristics, gonadal histopathology, and genetic sex. It would be desirable to include such ED-specific endpoints before using the FLCTT to investigate a possible ED. Although this section only considers the basic FLCTT without endocrine-specific endpoints, the section on the Medaka Multi-Generation Test (MMGT) includes many of these indicators of hormonal activity, and their use in evaluation of test results could be realised in a modified FLCTT that included such additional endpoints.

When/Why the Assay May be Used

C.3.5.4 Although the FLCTT could, in principle, be used at any stage in the hazard assessment process, the most likely use scenario will be when there are already some data available to suggest possible endocrine disruption properties. In other words, the FLCTT will generally be used to investigate whether such potential properties result in adverse apical effects on development, growth or reproduction over an entire lifecycle. It is unlikely (and undesirable) that the FLCTT will be the first ED-responsive test procedure to be applied to a chemical, but if it is, it would be useful to include at least some of the ED-specific endpoints described above. Even if the test chemical is already suspected of being an ED, it might still be desirable to include ED-specific endpoints in the FLCTT in order to establish a closer cause-effect relationship between endocrine changes and apical effects. If this has been done, it would be helpful to consider the guidance in the MMGT section.

C.3.5.5 The choice of whether to use an FLCTT or MMGT is not primarily driven by the existence of ED-specific endpoints in the latter, because as stated above, such endpoints could easily be included in the former. Insufficient research on the comparative responsiveness to EDs of these two tests has yet been conducted, but the scanty available data suggest that for most chemicals, these tests give NOEC/ECx values of similar magnitude *[ref to medaka comparative data to be inserted when available]*. However, some theoretical considerations suggest that strongly bioaccumulative EDs may be more potent in the MMGT than the FLCTT, primarily because the longer time available for

bioaccumulation in the MMGT allows the maternal transfer of residues to both F1 and F2 offspring as well as longer overall exposure via the ambient water. If a possible ED has a high bioconcentration factor (BCF), an MMGT may therefore be a more appropriate choice than an FLCTT. It is also presumed that the MMGT stands a better chance than the FLCTT of detecting the possible epigenetic effects of EDs which may only be expressed in adult F1 or subsequent offspring, although evidence for this type of effect in fish is currently lacking (Crews, D. and McLachlan, J.A. (2006); Brown, K.H. *et al*. (2009)).

Existing Data to be Considered

C.3.5.6 Existing data available before deployment of the FLCTT for endocrine disruption hazard assessment are likely to include information on possible MOAs from (Q)SARs and/or *in vitro* screens. These will probably be accompanied by *in vivo* fish assay data from OECD TG 229 or OECD TG 230, and may also include data from TG 234 (FSDT). It would not be advisable or ethically desirable to conduct an unmodified FLCTT without mechanistic or *in vivo* screening data because it would then not be possible to link any apical effects with endocrine disruption. Furthermore, data from OECD TG 229 and/or TG 234 (FSDT) could be of use in focusing attention in the FLCTT on particularly vulnerable parts of the lifecycle. Given the high ethical and financial cost of the FLCTT, it is important to make full use of existing endocrine-related data, both before the test is begun and during data evaluation.

Scenarios: Positive and Negative Results Combined with Existing Data

C.3.5.6a The scenarios (A to R) presented in Table C.3.5 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.3.5.7 Positive results obtained with one of the FLCTT endpoints result in the conclusion that the test chemical is able to cause adverse effects *in vivo* (Table C.3.5, Scenarios A-I), but not necessarily that it is an ED. Note that if doubt exists about the test performance (e.g. highly unusual results in controls), a comparison with historical control data with respect to overall test performance might be helpful. However, the nature of these effects and any existing data will require careful consideration. If *in vitro* and/or *in vivo* data already exist which reveal possible endocrine disrupting properties (Scenarios A, B and D), a positive endpoint in the FLCTT could lead to a conclusion that the test chemical is an actual ED. Such a conclusion will be strengthened considerably if the endocrine modality previously identified is plausibly linked to the responding endpoint. For example, if the chemical has estrogenic properties and there is observed to be reduced fecundity of the F0 adults in the FLCTT, this gives added confidence in this conclusion. On the other hand, it may be harder to argue a plausible link between estrogenic properties on the one hand, and an endpoint such as growth or survival on the other, although it is known that some estrogens are able to cause changes in growth rates (Knacker *et al*., 2010). In this example, an effect solely on growth or survival, while potentially of concern from the viewpoint of environmental risk assessment, would not on its own lead to a conclusion that the chemical is an ED in fish.

C.3.5.8 If a plausible link of a responding FLCTT endpoint with previously-identified endocrine activity can be made, regulatory authorities may conclude that sufficient evidence is available to categorise the chemical as an ED (*i.e.* interference with the endocrine system has caused adverse effects *in vivo*), and no further information might then be required. Of course, if the intention is to

conduct an environmental risk assessment, it may also be necessary to consider whether or not effects observed are relevant at the population level (*e.g.* reproduction; growth; development). On the other hand, if data from prior endocrine screens and tests are negative (Scenario E), a positive response in the FLCTT would not in general support the hypothesis that the chemical is an ED in fish (although it could be argued that a change in sex ratio is likely to have been caused by an ED). It could, of course, still be subjected to an environmental risk assessment.

C.3.5.9 The scenarios in which the FLCTT gives a negative result (Table C.3.5, Scenarios J-R) lead to a tentative conclusion that the test chemical is not an ED in fish, and this conclusion is strengthened considerably if prior screens have failed to reveal endocrine activity (Scenario N). In the latter circumstances, regulatory authorities would be justified in concluding that no further action is needed. On the other hand, if one or more of those screens was positive (Scenarios J, K, L, M and P), the BCF of the chemical should be checked. If the BCF indicates that the chemical is strongly bioaccumulative, it would be worth considering the conduct of an MMGT, although as indicated above, there is little evidence at present that EDs with a high BCF are consistently more potent in such a test. If a chemical which screened positive is not bioaccumulative, the probable reasons for lack of effects in the FLCTT might be metabolism to an inactive chemical, or failure to reach the active site, and no further action would be indicated.

C.3.5.10 In each of the above scenarios, it is possible that existing data will be equivocal (Table C.3.5, Scenarios C, F, G, H, I, L, O, P, Q and R), or there may be no existing data. This will weaken the conclusions which can be drawn about a positive FLCTT, and this is reflected in Table C.3.5. However, as indicated above, it would be undesirable to proceed with an FLCTT if prior data on endocrine activity are equivocal or absent, and if there are no other effect- or exposure-related reasons for considering such a comprehensive test. On the other hand, if the FLCTT is positive, it would be essential to obtain some reliable mechanistic data before reaching a conclusion about whether or not the chemical is an ED in fish.

C.3.5.11 The scenario in which the results of the FLCTT are themselves equivocal has not been dealt with in Table C.3.5, for reasons of brevity. In this context, an equivocal result might be an inconsistent concentration-response (*e.g.* no effect at a high concentration but effects at a lower), or a result which borders on statistical significance. Without knowing the exact circumstances, reliable advice cannot be given, but the opinions of an experienced ecotoxicologist should be sought. However, if a comprehensive set of prior screens are all negative, it is doubtful whether further action is needed, because the chemical is unlikely to be an ED. If an endocrine screen is positive, some types of equivocal FLCTT results would have to be taken more seriously. For example, a non-monotonic concentration-response would not necessarily rule out the test chemical as an ED in fish. An example of this would be a chemical like ethinylestradiol which causes adverse effects on fish reproduction at low doses, but reduced reproductive success at very high doses, thus potentially giving a U-shaped response curve. Ideally, concentrations causing systemic toxicity of this type should not be tested in an FLCTT, but such toxicity may have been missed in earlier screens.

C.3.5.12 In summary, positive results in the FLCTT indicate that a chemical is a probable ED if they can be plausibly linked to an endocrine MOA established on the basis of prior mechanistic screening or concurrent observation of mechanistic effects or their biochemical/ physiological manifestations. If such screening data are unavailable or negative, it should not be concluded that a positive FLCTT is the result of endocrine disruption (although it is likely that biased sex ratio will be the result of ED). On the other hand, a negative FLCTT combined with a sufficiently comprehensive set of negative screening data could lead to a firm conclusion that a chemical is not an ED in fish. A negative FLCTT set against a background of a positive screen might, however, raise concerns *e.g.* if the chemical is strongly bioaccumulative or known to be involved in epigenesis. In this case an MMGT should be considered.

Table C.3.5 Fish Lifecycle Toxicity Test (FLCTT) (USEPA OPPTS 850.1500). Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR, and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". There is no evidence at present that equivalent *in vitro* assays with systems derived from fish offer advantages over their mammalian counterparts.

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter.

Note that although this assay has been used for many years to assess the chronic effects of chemicals, no attempt has been made to validate it for use with possible EDs, and it has not been published as an OECD TG.

Note that under some scenarios, a Medaka Multi-Generation Test (MMGT) is recommended as a Next Step. This test is still being validated, so it is described relatively briefly in Annex 2.

Scenario	Result of FLCTT	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	The test chemical is almost certainly an ED if the modality identified in	Further evidence is probably not required.	If the affected endpoint in the FLCTT cannot be plausibly linked to the known modality, the test chemical is unlikely to be

Scenario	Result of FLCTT	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				existing screens/tests can be plausibly linked to the affected endpoint.		an ED. The FLCTT is unlikely to detect epigenetic effects. If these are suspected, an MMGT may reveal them.
B	+	+	-	The test chemical is almost certainly an ED in fish if the modality identified in existing screens/tests can be plausibly linked to the affected endpoint.	Further evidence is probably not required.	If the affected endpoint in the FLCTT cannot be plausibly linked to the known modality, the test chemical is unlikely to be an ED. The FLCTT is unlikely to detect epigenetic effects. If these are suspected, an MMGT may reveal them.
C	+	+	Eq/0**	The test chemical is almost certainly an ED in fish if the modality identified in existing screens can be plausibly linked to the affected endpoint.	Further evidence is probably not required.	If the affected endpoint in the FLCTT cannot be plausibly linked to the known modality, the test chemical is unlikely to be an ED. The FLCTT is unlikely to detect epigenetic effects. If these are suspected, an MMGT may reveal them.
D	+	-	+	The test chemical may be an ED, but the negative mechanistic data reduce the confidence in this conclusion. However, if the endocrine disruption effects in existing *in vivo*	Further evidence is probably not required.	If the affected endpoint in the FLCTT cannot be plausibly linked to the endocrine effects in existing *in vivo* tests, the test chemical is unlikely to be an ED. The FLCTT is unlikely to detect epigenetic effects. If these are suspected, an MMGT may reveal them.

Scenario	Result of FLCTT	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				tests can be plausibly linked to the FLCTT responses, this increases the probability that the chemical is an ED.		
E	+	-	-	The test chemical is unlikely to be an ED[4].	Further evidence is probably not required.	It is possible that the effects observed in the FLCTT have been caused by an unknown endocrine mechanism. This would not, however, prevent the chemical being subjected to risk assessment. The FLCTT is unlikely to detect epigenetic effects. If these are suspected, an MMGT may reveal them.
F	+	-	Eq/0	The test chemical is unlikely to be an ED, but the relevance of any equivocal existing *in vivo* data to the FLCTT results should be examined.	Further evidence is probably not required.	It is possible that the effects observed in the FLCTT have been caused by an unknown endocrine mechanism – equivocal existing *in vivo* data may throw some light on this. The absence of data on a possible endocrine mechanism would, however, not prevent the chemical being subjected to risk assessment. The FLCTT is unlikely to detect epigenetic effects. If these are suspected, an MMGT may reveal them.

[4] However, note that if biased sex ratio is observed, it is likely to have been caused by an EDC.

Scenario	Result of FLCTT	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
G	+	Eq/0	+	The test chemical may be an ED, but the equivocal or absent mechanistic data reduce the confidence in this conclusion. However, if the endocrine disruption effects in existing *in vivo* tests can be plausibly linked to the FLCTT responses, this increases the probability that the chemical is an ED.	If reliable mechanistic data are not available, it would be desirable to obtain some.	The test chemical is almost certainly an ED if a modality identified in the newly commissioned mechanistic screens, or in the existing *in vivo* data, can be plausibly linked to the affected endpoint. The FLCTT is unlikely to detect epigenetic effects. If these are suspected, an MMGT may reveal them.
H	+	Eq/0	-	The test chemical may be an ED, but the equivocal or absent mechanistic data reduce the confidence in this conclusion.	If reliable mechanistic data are not available, it would be desirable to obtain some.	The test chemical is almost certainly an ED if a modality identified in the newly commissioned mechanistic screens can be plausibly linked to the affected endpoint. The FLCTT is unlikely to detect epigenetic effects. If these are suspected, an MMGT may reveal them.
I	+	Eq/0	Eq/0	The test chemical may be an ED, but the equivocal or absent mechanistic and *in vivo* data reduce the confidence in this conclusion.	If reliable mechanistic data are not available, it would be desirable to obtain some.	The test chemical is almost certainly an ED if a modality identified in the newly commissioned mechanistic screens can be plausibly linked to the affected endpoint. The FLCTT is unlikely to detect epigenetic effects. If these are suspected, an MMGT

Scenario	Result of FLCTT	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						may reveal them.
J	-	+	+	The chemical is probably not an ED in fish, unless this conclusion is contradicted by existing *in vivo* data.	If the chemical is strongly bioaccumulative, or if epigenetic effects are suspected, consider conducting an MMGT.	If any effects in an MMGT can be plausibly linked with mechanistic data, the test chemical is probably an ED.
K	-	+	-	The chemical is probably not an ED in fish.	If the chemical is strongly bioaccumulative, or if epigenetic effects are suspected, consider conducting an MMGT.	If any effects in an MMGT can be plausibly linked with mechanistic data, the test chemical is probably an ED.
L	-	+	Eq/0	The chemical is probably not an ED in fish.	If the chemical is strongly bioaccumulative, or if epigenetic effects are suspected, consider conducting an MMGT.	If any effects in an MMGT can be plausibly linked with mechanistic data, the test chemical is probably an ED.
M	-	-	+	The chemical is probably not an ED in fish.	If the chemical is strongly bioaccumulative, or if epigenetic effects are suspected, consider conducting an MMGT.	If any effects in an MMGT can be plausibly linked with *in vivo* data which indicate ED properties, the test chemical is probably an ED, but likely not by a mechanism covered by the existing *in vitro* screens.

Scenario	Result of FLCTT	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
N	-	-	-	The chemical is probably not an ED.	Further evidence is probably not required.	-
O	-	-	Eq/0	The chemical is probably not an ED in fish.	Further evidence is probably not required.	-
P	-	Eq/0	+	The chemical is probably not an ED in fish.	If reliable mechanistic data are not available, it would be desirable to obtain some.	If the newly commissioned mechanistic data are positive and the chemical is strongly bioaccumulative, or if epigenetic effects are suspected, consider conducting an MMGT.
Q	-	Eq/0	-	The chemical is probably not an ED, but confidence in this conclusion is reduced by the lack of clear mechanistic data.	Further evidence is probably not required, but confidence in the conclusion would be increased by the provision of reliable negative mechanistic data.	If any newly commissioned mechanistic data are positive and the chemical is strongly bioaccumulative, or if epigenetic effects are suspected, consider conducting an MMGT.
R	-	Eq/0	Eq/0	The chemical may not be an ED, but confidence in this conclusion is reduced by the lack of clear mechanistic and existing *in vivo* data.	Further evidence is probably not required, but confidence in the conclusion would be increased by the provision of reliable negative mechanistic	If any newly commissioned mechanistic data are positive and the chemical is strongly bioaccumulative, or if epigenetic effects are suspected, consider conducting an MMGT.

Scenario	Result of FLCTT	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					data.	

C.3.6 OECD TG 231: Amphibian Metamorphosis Assay (AMA)

C.3.6.1 Modality detected/endpoints: Thyroid activity (advanced development; asynchronous development; delayed development in absence of non-specific systemic toxicity; thyroid histopathology), but note that this covers several different modes of action, including thyroid agonists and antagonists, as well as substances interfering with thyroid hormone synthesis and transport. According to OECD TG 231, there is disagreement about the implications of the different endpoints in this larval development screen. Some experts accept that changes in one of the thyroid-relevant apical endpoints (advanced development; asynchronous development; delayed development in absence of non-specific systemic toxicity) may on their own indicate thyroid activity, while others will only reach this conclusion if one of the apical endpoints is accompanied by significant thyroid histopathology such as moderate or severe follicular hypertrophy and/or hyperplasia (OECD, 2007c). Note that the AMA is subject to indirect thyroid effects such as those that result from cytochrome P450 induction (*e.g.* phenobarbital, the model compound for the latter effect, tests positive in the AMA). Therefore, interpretation of the AMA may be complicated.

Background to the Assay

C.3.6.2 This assay is designed as a screen for thyroid activity in amphibians, and not to provide information on endocrine activity for use in assessing the environmental risks of an individual chemical based on a PEC/PNEC approach. It is important to note that there are several types of thyroid disruption, not all of which involve interactions with the thyroid receptor, and they have differential effects on the various endpoints in this screen. OECD TG 231 does not, however, allow unequivocal diagnosis of which type of thyroid disruption is occurring. It includes a specific endpoint (thyroid gland histopathology) for some types of thyroid activity, but also includes apical measurements (hind limb length, snout-vent length, developmental stage and wet weight), which are used to determine other thyroid-responsive endpoints: advanced development, asynchronous development or delayed development. The first two of these are considered by some authorities to be diagnostic of thyroid activity, while the latter is only diagnostic if non-specific systemic toxicity is absent. It should also be noted that a recent review (Pickford, 2010) concluded that for thyroid agonists, the response of amphibian thyroid histopathology is not as predictable or as sensitive as developmental stage or hind limb development. However, it is probable that a diagnosis of thyroid activity on the basis of the apical endpoints will be more robust if accompanied by thyroid histopathology, and *vice versa*.

C.3.6.3 Consequently, if the assay gives a positive result, this may be due to a combination of a positive indicator of hormonal activity (thyroid histopathology) and a positive apical endpoint (advanced development, asynchronous development, or delayed development), or a positive indicator of hormonal activity alone (possibly accompanied by a negative apical endpoint), or for an apical endpoint alone (possibly accompanied by a negative indicator of hormonal activity). Each of these possible combinations of positive response should be considered separately (although the distinctions between indicators of hormonal activity and apical effects are not always clear), so they have been listed individually as points 1, 2 and 3 in the Possible Conclusions column of Table C.3.6. It should be noted, however, that due to the relatively short exposure time employed in this screen (3 weeks), one cannot be sure if the effects of some chemicals on apical endpoints would result in adverse effects on development, growth or reproduction in the longer term. This is primarily relevant for risk assessments and not if a regulatory authority is solely concerned with hazard assessment. Also, as only 3 test concentrations are usually employed, even a reliable short-term NOEC/ECx or ECx for the apical endpoints cannot be precisely derived.

When/Why the Assay May be Used

C.3.6.4 Although OECD TG 231 could, in principle, be used at any stage in the hazard assessment process, the most likely use scenario will be when there are relatively few data available about the possible thyroid disrupting properties of a chemical. The results from this assay are most likely to be available after deployment of a battery of *in vitro* and *in vivo* screens (*e.g.* the USEPA's Endocrine Disruptor Screening Program), or as a supplement to existing data which suggest possible ED activity. A number of mammalian (rat) assays are sensitive to thyroid disruption, particularly thyroid antagonists, including the pubertal assay (male or female), the enhanced repeat dose assay (OECD TG 407), and the intact male screening assay. Note that these assays utilize different routes of exposure than OECD TG 231 and therefore, depending on the properties of the chemical, have differing potentials for the test substance to be metabolized. It should also be noted that only the AMA appears to be sensitive to thyroid agonists. It has been argued by Pickford (2010) that only one thyroid-disrupting chemical (methoxychlor) shows activity in the AMA but not in any rodent screens, but the range of chemicals tested in the former is less than in the latter.

C.3.6.5 It is possible that no endocrine-relevant data are available before the AMA is deployed (*i.e.* if OECD TG 231 has been used as a primary screen), but in that case a positive result in the screen could be followed up with relevant *in vitro* screening to investigate the suspected mode of action. However, it should be noted that *in vitro* screens essentially only exist for thyroid agonists and antagonists (*e.g.* GH_3 rat pituitary somatotroph cell proliferation; solid state thyroid receptor binding assays; transfected reporter gene assays in yeast or mammalian cell lines), while thyroid disruption can occur at other points in the endocrine system for which *in vitro* screens do not exist, or are still at the research stage (*e.g.* FRTL-5 rat cell lines sensitive to iodide uptake inhibitors) (see Para A.18). Furthermore, none of these screens have yet been validated and standardised at the international level.

Existing Data to be Considered

C.3.6.6 Existing data available before deployment of OECD TG 231 might include *in vivo* results obtained with other vertebrates (*e.g.* a positive *in vivo* assay with rats – see above), or one or more of a range of *in silico* or *in vitro* results which suggest that thyroid disruption may occur *in vivo* (but note the limitations of this approach, as indicated above). Such indicators of possible thyroid activity might include (Q)SAR predictions of thyroid activity, 'read-across' from *in vivo* results obtained with chemically related chemicals, or positive results from an *in vitro* screen for thyroid agonist/antagonist activity.

Scenarios: Positive and Negative Results Combined with Existing Data

C.3.6.6a The scenarios (A to R) presented in Table C.3.6 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.3.6.7 Positive results obtained with the thyroid histopathology endpoint (Table C.3.6, Scenarios A-I, sub-section 2) result in the conclusion that the test chemical is probably a possible ED *in vivo*. If both thyroid histopathology and an apical endpoint give a response (Table C.3.6, Scenarios A-I, sub-section 1), this may provide even stronger evidence that one is dealing with a possible ED, especially if its action is not receptor-mediated. If only an apical endpoint responds (Table C.3.6, Scenarios A-I, sub-section 3), it suggests that the chemical is a possible thyroid disrupter, but with somewhat reduced

confidence in some cases compared to sub-section 2 (although existing positive *in vitro* data, or positive *in vivo* data from other species, would have to be weighed against this conclusion). Note, however, that apical endpoints alone are probably sufficiently responsive to thyroid receptor agonists (*i.e*. in these cases thyroid histopathology is unlikely to make the assay more robust) (Daniel Pickford, pers. comm.. 2010).

C.3.6.8 As indicated above, although a positive response of OECD TG 231 indicates that the chemical is a possible thyroid disrupter, a result of this type would generally need to be followed up with a more comprehensive growth, development and/or reproduction test if countries need further evidence (*i.e.* a Larval Amphibian Growth and development Assay - LAGDA) which is able to provide a precise NOEC/ECx for adverse effects. In other words, in order to increase evidence, a positive result of whichever type in OECD TG 231 could be followed by a LAGDA at Level 5. Existing data suggesting endocrine-specific activity (*e.g.* positive *in vitro* data, or positive *in vivo* data from other species) will strengthen the case for additional testing still further. Note, however, that the LAGDA has not yet been validated or standardised (this work is on-going), and it is not a true lifecycle test which includes all aspects of reproduction. For that reason, it is worth considering whether a positive result in OECD TG 231 could be more usefully followed up under some circumstances by an FLCTT or MMGT with thyroid-specific endpoints such as thyroid hormone induction or depression, although at present the responsiveness of apical endpoints in these tests (*e.g.* growth) to thyroid-active substances is not well understood.

C.3.6.9 The situation in which OECD TG 231 gives a negative result (Table C.3.6, Scenarios J-R) needs careful consideration of any existing data. If these data suggest that the chemical is endocrine-active both *in vitro* and *in vivo* (Scenario J), then it is possible that OECD TG 231 is simply insufficiently sensitive, although most known thyroid disrupters have been shown to give a response in the AMA. Depending on the robustness of the existing data, it might therefore be appropriate to conduct a LAGDA.

C.3.6.10 If OECD TG 231 and existing *in vivo* data are all negative, but *in vitro* data reveal some endocrine activity (Scenario K), the probability is that the test chemical is not sufficiently potent to produce thyroid effects *in vivo* in amphibians or other organisms, or it may be rapidly metabolised. In such a situation, further testing is probably not necessary. However, if the chemical is known to bioaccumulate slowly, it may be that exposures in the *in vivo* tests have been insufficiently prolonged, in which case longer-term testing with the LAGDA might be justified.

C.3.6.11 On the other hand, if OECD TG 231 and the *in vitro* tests are negative (Scenario M), but there are positive existing *in vivo* data, the nature of those existing data should be considered. Unless the existing data are from another amphibian, the chemical is probably not a possible ED acting on amphibian growth or development, but it may act via MOAs not covered by the *in vitro* screens, or it may be more potent in species or life-stages that have not been tested. In this situation, the existing *in vivo* data should be used to guide decisions about whether to conduct any further testing.

C.3.6.12 Finally, a negative OECD TG 231 screen, set against a background of negative *in vitro* and *in vivo* data (Scenario N), suggests that the test chemical is not a possible thyroid-active ED, and further action is unnecessary.

C.3.6.13 In each of the above scenarios, it is possible that existing data will be equivocal, or there may be no existing data. This will weaken the conclusions which can be drawn about a negative OECD TG 231 test, and this is reflected in Table C.3.6. However, a lack of mechanistic data on thyroid activity should ideally be rectified before any further *in vivo* testing is finally conducted, although as indicated above, many thyroid modalities are not detectable in *in vitro* screens. On the other hand, if OECD TG 231 is positive, further *in vivo* testing would generally be needed to quantify any adverse effects and/or to establish a NOEC or ECx for such effects, even if all existing data are equivocal, or if there are no existing data. Again, however, it may be useful to obtain some mechanistic information before conducting further *in vivo* testing.

C.3.6.14 The scenario in which the results of OECD TG 231 are themselves equivocal has not been dealt with in Table C.3.6, for reasons of brevity. In this context, an equivocal result might be an inconsistent concentration-response (*e.g.* no effect at a high concentration but effects at a lower), or a result which borders on statistical significance. Without knowing the exact circumstances, reliable advice cannot be given, but the opinions of an experienced ecotoxicologist should be sought. Clearly, however, such equivocal results do not necessarily rule out the existence of *in vivo* endocrine activity. For example, thyroid histopathology at a high concentration might be masked by any systemic toxicity, while growth measurements might just fail to reach a statistically significant level due to unexpectedly high variability. If these or other possible reasons for false negatives are suspected with good reason, the screen could be repeated (*e.g.* conduct it at lower concentrations which avoid systemic toxicity), or a more appropriate version of it (*e.g.* more larvae per replicate) could be designed and conducted. However, note that a repeat screen in the event of systemic toxicity would not be needed providing at least one tested concentration was not subject to such effects.

C.3.6.15 In summary, certain positive results in the OECD TG 231 screen may indicate that a chemical is a possible endocrine disrupter via one of several types of thyroid activity. This suggests that more comprehensive *in vivo* testing would be needed if the intention is to derive a long-term NOEC/ECx and/or to confirm whether or not the chemical is an actual endocrine disrupter due to the occurrence of adverse effects. Negative results in OECD TG 231 do not necessarily mean that the chemical is not a possible ED – a judgement about the endocrine disruption potential and the possible need for additional testing will have to be made based on a weight of evidence evaluation of existing *in vitro* and *in vivo* data.

Table C.3.6 OECD TG 231: Amphibian Metamorphosis Assay (AMA). Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from available from ER and AR based assays and the steroidogenesis assay (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an thyroid disrupter.

The assay under discussion could either be positive for both apical endpoints and indicators of hormonal activity, or positive just for an apical endpoints or the indicator of hormonal activity. For each scenario, each of these 2 possibilities is addressed separately in the Possible Conclusions column.

Note that, under some scenarios, a Next Step could involve conduct of a Larval Amphibian Growth and Development Assay (LAGDA). This is currently being validated and is therefore only described briefly in Annex 2.

Scenarios	Result of OECD TG 231 assay (AMA)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	1) Strong evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in	Consider performing a Larval Amphibian Growth	Based on the limited scope of current *in vitro* screens, the positive *in vitro* data suggest that

Scenarios	Result of OECD TG 231 assay (AMA)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				amphibians, plus thyroid effects in other species 2) Strong evidence for *in vivo* thyroid activity in amphibians, plus thyroid effects in other species 3) Some evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians, plus thyroid effects in other species.	and Development Assay (LAGDA).	the test chemical is a thyroid (ant)agonist.
B	+	+	-	1) Strong evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians 2) Strong evidence for *in vivo* thyroid activity in amphibians 3) Some evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians	Consider performing a Larval Amphibian Growth and Development Assay (LAGDA).	Based on the limited scope of current *in vitro* screens, the positive *in vitro* data suggest that the test chemical is a thyroid (ant)agonist. Cases where chemicals are active in the AMA but not in thyroid-responsive rodent assays are rare. In this scenario, it is therefore particularly important to discover if adverse effects appear in a longer-term amphibian test.
C	+	+	Eq/0	1) Strong evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians	Consider performing a Larval Amphibian Growth	Based on the limited scope of current *in vitro* screens, the positive *in vitro* data suggest that the test chemical is a thyroid

Scenarios	Result of OECD TG 231 assay (AMA)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				2) Strong evidence for *in vivo* thyroid activity in amphibians 3) Some evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians	and Development Assay (LAGDA).	(ant)agonist.
D	+	-	+	1) Strong evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians, plus thyroid effects in other species 2) Strong evidence for *in vivo* thyroid activity in amphibians, plus thyroid effects in other species 3) Some evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians, plus thyroid effects in other species.	Consider performing a Larval Amphibian Growth and Development Assay (LAGDA).	The lack of *in vitro* thyroid activity is not evidence against any thyroid activity, due to the limited nature of current *in vitro* thyroid screens.
E	+	-	-	1) Strong evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians 2) Strong evidence for *in vivo* thyroid activity in amphibians	Consider performing a Larval Amphibian Growth and Development Assay (LAGDA).	The lack of *in vitro* thyroid activity is not evidence against any thyroid activity, due to the limited nature of current *in vitro* thyroid screens.

Scenarios	Result of OECD TG 231 assay (AMA)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				3) Some evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians.		Cases where chemicals are active in the AMA but not in thyroid-responsive rodent assays are rare. In this scenario, it is therefore particularly important to discover if adverse effects appear in a longer-term amphibian test.
F	+	-	Eq/0	1) Strong evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians 2) Strong evidence for *in vivo* thyroid activity in amphibians 3) Some evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians	Consider performing a Larval Amphibian Growth and Development Assay (LAGDA). Given the absence or equivocal nature of existing *in vivo* data, it might also be sensible to conduct a thyroid-responsive mammalian assay (*e.g.* rat pubertal)	The lack of *in vitro* thyroid activity is not evidence against any thyroid activity, due to the limited nature of current *in vitro* thyroid screens.
G	+	Eq/0	+	1) Strong evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians, plus thyroid effects in other	Consider performing a Larval Amphibian Growth and Development	If a new *in vitro* mechanistic assay is conducted, note that a negative does not mean that the test material has no thyroid

Scenarios	Result of OECD TG 231 assay (AMA)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				species 2) Strong evidence for *in vivo* thyroid activity in amphibians, plus thyroid effects in other species 3) Some evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians, plus thyroid effects in other species.	Assay (LAGDA). Given the absence or equivocal nature of the *in vitro* mechanistic data, it might also be helpful to conduct an *in vitro* screen for thyroid (ant)agonistic activity.	activity.
H	+	Eq/0	-	1) Strong evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians 2) Strong evidence for *in vivo* thyroid activity in amphibians 3) Some evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians	Consider performing a Larval Amphibian Growth and Development Assay (LAGDA). Given the absence or equivocal nature of the *in vitro* mechanistic data, it might also be helpful to conduct an *in vitro*	Cases where chemicals are active in the AMA but not in thyroid-responsive rodent assays are rare. In this scenario, it is therefore particularly important to discover if adverse effects appear in a longer-term amphibian test.

Scenarios	Result of OECD TG 231 assay (AMA)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					screen for thyroid (ant)agonistic activity.	
I	+	Eq/0	Eq/0	1) Strong evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians 2) Strong evidence for *in vivo* thyroid activity in amphibians 3) Some evidence for *in vivo* thyroid activity with potential adverse effects (developmental/growth toxicity) in amphibians	Consider performing a Larval Amphibian Growth and Development Assay (LAGDA). Given the absence or equivocal nature of the *in vitro* mechanistic data, it might also be helpful to conduct an *in vitro* screen for thyroid (ant)agonistic activity. Given the absence or equivocal nature of existing *in vivo* data, it might also be sensible to conduct a	If a new *in vitro* mechanistic assay is conducted, note that a negative does not mean that the test material has no thyroid activity.

Scenarios	Result of OECD TG 231 assay (AMA)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					thyroid-responsive mammalian assay (*e.g.* rat pubertal)	
J	-	+	+	The test chemical is probably a thyroid (ant)agonist without activity in amphibians, although it is possible that *Xenopus laevis* responds atypically in this case.	Some regulatory authorities may conclude that no further evidence is required, but it might be desirable to conduct a LAGDA with a species other than *X. laevis* if the existing data are sufficiently persuasive.	Based on the limited scope of current *in vitro* screens, the positive *in vitro* data suggest that the test chemical is a thyroid (ant)agonist.
K	-	+	-	The test chemical is probably a thyroid (ant)agonist without activity in amphibians or other taxa, although it is possible that *Xenopus laevis* responds atypically in this case.	If there is no activity in amphibian or mammals, further evidence is probably not needed.	Based on the limited scope of current *in vitro* screens, the positive *in vitro* data suggest that the test chemical is a thyroid (ant)agonist.
L	-	+	Eq/0	The test chemical is probably a thyroid (ant)agonist without activity in amphibians, although it is possible that *Xenopus laevis* responds atypically in this case.	Some regulatory authorities may conclude that no further evidence is required, but if	Based on the limited scope of current *in vitro* screens, the positive *in vitro* data suggest that the test chemical is a thyroid (ant)agonist.

Scenarios	Result of OECD TG 231 assay (AMA)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					mammalian data are absent, it might be desirable to conduct a thyroid-responsive rodent screen (*e.g.* rat pubertal).	
M	-	-	+	The test chemical is probably without thyroid activity in amphibians, although it is possible that *Xenopus laevis* responds atypically in this case.	Some regulatory authorities may conclude that no further evidence is required, but the positive existing *in vivo* data suggest that it might be helpful to perform a LAGDA with a species other than *X. laevis*.	The lack of *in vitro* thyroid activity is not evidence against any thyroid activity, due to the limited nature of current *in vitro* thyroid screens.
N	-	-	-	The test chemical is probably without thyroid activity in amphibians or other taxa.	No further action is necessary.	-
O	-	-	Eq/0	The test chemical is probably without thyroid activity in amphibians.	Some regulatory authorities may conclude that no further evidence is required, but if mammalian data are absent, it might be	The lack of *in vitro* thyroid activity is not evidence against any thyroid activity, due to the limited nature of current *in vitro* thyroid screens.

Scenarios	Result of OECD TG 231 assay (AMA)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					desirable to conduct a thyroid-responsive rodent screen (*e.g.* rat pubertal).	
P	-	Eq/0	+	The test chemical is probably without thyroid activity in amphibians, although it is possible that *Xenopus laevis* responds atypically in this case.	Some regulatory authorities may conclude that no further evidence is required, but the positive existing *in vivo* data suggest that it might be helpful to perform a LAGDA with a species other than *X. laevis*. Also, if clear *in vitro* mechanistic data are missing, it might be desirable to obtain some.	If a new *in vitro* mechanistic assay is conducted, note that a negative does not mean that the test material has no thyroid activity.
Q	-	Eq/0	-	The test chemical is probably without thyroid activity in amphibians or other taxa.	No further action is necessary.	-
R	-	Eq/0	Eq/0	The test chemical is probably without thyroid activity in amphibians.	Some regulatory authorities may	-

Scenarios	Result of OECD TG 231 assay (AMA)	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					conclude that no further evidence is required, but if mammalian data are absent, it might be desirable to conduct a thyroid-responsive rodent screen (*e.g.* rat pubertal).	

C.3.7 OECD TG 206: Avian Reproduction Test

C.3.7.1 Modality detected/endpoints: OECD TG 206 does not contain endpoints which solely respond to endocrine disrupters, and it has not been specifically validated with EDs. However, some of the endpoints in this apical test are nevertheless potentially affected by EATS EDs. Of particular interest in the context of estrogens, androgens and steroidogenesis disrupters are egg production, embryo viability, and hatchability, but other endpoints may also be responsive to some EDs (*e.g.* growth may respond to some thyroid disrupters; % cracked eggs and egg shell thickness may respond to chemicals interfering with the control of shell deposition).

Background to the Assay

C.3.7.2 This assay is designed primarily as an apical test for chemicals with suspected reproductive toxicity, but it is not a lifecycle test as it only runs from the stage of pre-laying adults to 14 day old offspring. Furthermore, only the adults are exposed to the test chemical (via the food), and any effects on sexual development would not be detectable. The endpoints are all apical measures of development, growth or reproduction. Key endpoints which might be affected by EDs include egg production, viability, and hatchability. Possible test organisms include mallard duck (*Anas platyrhynchos*), bobwhite quail (*Colinus virginiatus*) and Japanese quail (*Coturnix japonica*).

C.3.7.3 It should be noted that it would be relatively straightforward to include ED-specific endpoints in this test. Depending on the species and test objectives, these could include *inter alia* sex ratio (phenotypic and/or genotypic), sex hormones, thyroid hormones, reproductive/thyroid organ weights, gonad histopathology and gross pathology, time to first egg laying, and sexual behaviour. These types of endpoint are all included in the Avian Two-Generation Test (ATGT). However, note that the ATGT does not cover all relevant behaviours and is performed in a precocial species which reacts very differently to embryonic exposure to a test material compared with an altricial species.

When/Why the Assay May be Used

C.3.7.4 Although OECD TG 206 could, in principle, be used at any stage in the hazard assessment process, the most likely use scenario will be when there are already some data available to suggest possible endocrine disruption properties. In other words, OECD TG 206 will generally be used to investigate whether such properties result in adverse apical effects on development, growth or reproduction over the reproductive part of the avian lifecycle. OECD TG 206 could not be used as a primary screen for EDs. Another potential limitation of OECD TG 206 is that the effects of test chemicals may not become fully apparent during the test because the offspring are not directly dosed, and only receive bioaccumulated material which may be passed from their mothers via the egg.

Existing Data to be Considered

C.3.7.5 Existing data available before deployment of OECD TG 206 for ED hazard assessment are likely to include information on possible MOAs from (Q)SARs and/or *in vitro* screens. It would not be advisable to conduct an unmodified OECD TG 206 without mechanistic screening data because it would then not be possible to link any apical effects with endocrine disruption. Given the high ethical and financial cost of OECD TG 206, it is important to make full use of existing endocrine-related data, both before the test is begun and during data evaluation.

Scenarios: Positive and Negative Results Combined with Existing Data

C.3.7.5a The scenarios (A to R) presented in Table C.3.7 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

C.3.7.6 Positive results obtained with one of the OECD TG 206 endpoints which are outside the range of historical controls may result in the conclusion that the test chemical is able to cause adverse effects *in vivo* (Table C.3.7, Scenarios A-I), but not necessarily that it is an ED. Note that if doubt exists about the test performance (e.g. highly unusual results in controls), a comparison with historical control data with respect to overall test performance might be helpful. However, the nature of these effects and any existing data will require careful consideration. If *in vitro* and/or *in vivo* data already exist which reveal possible endocrine disrupting properties (Scenarios A, B and D), a positive endpoint in OECD TG 206 could lead to a tentative conclusion that the test chemical is an actual ED.

C.3.7.7 If a plausible link of a responding OECD TG 206 endpoint with previously-identified endocrine activity can be made, regulatory authorities may conclude that sufficient evidence is available to categorise the chemical as an ED (*i.e.* interference with the endocrine system has caused adverse effects *in vivo*), and no further information might then be required. However, if a more robust link between adverse effects and an endocrine modality is required (bearing in mind that none of the existing data are likely to have been generated in avian systems), or if possible effects during the sexual development part of the lifecycle are suspected, or if the chemical is suspected to cause epigenetic effects, it would be desirable to run an ATGT. Furthermore, if data on hazard are required for an environmental risk assessment, an ATGT may also be needed unless the precision of the data from OECD TG 206 (which only uses 3 test concentrations) are considered adequate for such an assessment. On the other hand, if data from prior endocrine screens and tests are negative (Scenario E), a positive response in OECD TG 206 would not support the hypothesis that the chemical is an ED in birds. It could, of course, still be subjected to an environmental risk assessment, but only if sufficient concentrations have been tested to allow derivation of an adequately precise LOEC/NOEC.

C.3.7.8 The scenarios in which OECD TG 206 gives a negative result (Table C.3.7, Scenarios J-R) lead to a tentative conclusion that the test chemical is not an ED in birds, and this conclusion is strengthened considerably if prior screens have failed to reveal endocrine activity (Scenario N). In the latter circumstances, regulatory authorities may be justified in concluding that no further action is needed. However, if it is thought possible that the sexual development part of the lifecycle is sensitive, then conduct of an ATGT should be considered. Also, if one or more of those screens was positive (Scenarios J, K, L, M and P), the BCF of the chemical should be checked. If the BCF indicates that the chemical is strongly bioaccumulative, it would also be worth considering the conduct of an ATGT. If a chemical which screened positive is not bioaccumulative, the probable reasons for lack of effects in OECD TG 206 might be metabolism to an inactive chemical, or failure to reach the active site, and no further action would be indicated.

C.3.7.9 In each of the above scenarios, it is possible that existing data will be equivocal (Table C.3.7, Scenarios C, F, G, H, I, L, O, P, Q and R), or there may be no existing data. This will weaken the conclusions which can be drawn about a positive OECD TG 206, and this is reflected in Table C.3.7. However, as indicated above, it would be undesirable to proceed with OECD TG 206 if prior data on endocrine activity are equivocal or absent. On the other hand, if OECD TG 206 is positive, it would be essential to obtain some reliable mechanistic data before reaching a conclusion about whether or not the chemical is an ED in birds.

C.3.7.10 The scenario in which the results of OECD TG 206 are themselves equivocal has not been dealt with in Table C.3.7, for reasons of brevity. In this context, an equivocal result might be an inconsistent concentration-response (*e.g.* no effect at a high concentration but effects at a lower), or a result which borders on statistical significance. Without knowing the exact circumstances, reliable advice cannot be given, but the opinions of an experienced ecotoxicologist should be sought. However, if prior screens are negative, it is doubtful if further action is needed, because the chemical is unlikely to be an ED. If an endocrine screen is positive, some types of equivocal OECD TG 206 results would have to be taken more seriously. For example, an inconsistent concentration-response would not necessarily rule out the test chemical as an ED in birds. An example of this would be a chemical which causes adverse effects on reproduction at low doses, but reduced reproductive success and ultimately mortality at very high doses, thus potentially giving a U-shaped response curve. Ideally, concentrations causing systemic toxicity of this type should not be tested in OECD TG 206, but such toxicity may have been missed in earlier screens.

C.3.7.11 In summary, positive results in OECD TG 206 indicate that a chemical may be an ED if they can be plausibly linked to an endocrine MOA established on the basis of prior screening. However, more conclusive data in this regard would be obtainable from an ATGT. If screening data are unavailable or negative, it should not be concluded that a positive OECD TG 206 is the result of endocrine disruption. On the other hand, a negative OECD TG 206 combined with negative screening data should lead to a conclusion that a chemical is probably not an ED in birds. A negative OECD TG 206 set against a background of a positive screen might, however, raise concerns if the chemical is strongly bioaccumulative, known to be involved in epigenesis, or suspected of having effects on sexual development, when an ATGT should be considered.

Table C.3.7 OECD TG 206: Avian Reproduction Test. Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR, and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". There is no evidence at present that equivalent *in vitro* assays with systems derived from fish offer advantages over their mammalian counterparts.

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter.

Note that although this assay has been used for many years to assess the sub-acute effects of chemicals, and no formal attempt has been made to validate it for use with possible EDs, the USEPA has shown that reproduction is a part of the avian life-cycle which can be responsive to EDs (http://www.epa.gov/endo/pubs/edmvac/final_avian_drp04_20_05.pdf).

Scenario	Result of OECD TG 206	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	The test chemical is probably an ED if the modality identified in existing screens/tests can be plausibly linked to the affected endpoint.	Further evidence is probably not required.	If the affected endpoint in OECD TG 206 cannot be plausibly linked to the known modality, the test chemical is unlikely to be an ED in birds. OECD TG 206 cannot detect effects on sexual development and is unlikely to detect

Scenario	Result of OECD TG 206	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an ATGT may reveal them.
B	+	+	-	The test chemical is probably an ED in birds if the modality identified in existing screens/tests can be plausibly linked to the affected endpoint.	Further evidence is probably not required.	If the affected endpoint in OECD TG 206 cannot be plausibly linked to the known modality, the test chemical is unlikely to be an ED in birds. OECD TG 206 cannot detect effects on sexual development and is unlikely to detect epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an ATGT may reveal them.
C	+	+	Eq/0**	The test chemical is probably an ED in birds if the modality identified in existing screens/tests can be plausibly linked to the affected endpoint.	Further evidence is probably not required.	If the affected endpoint in OECD TG 206 cannot be plausibly linked to the known modality, the test chemical is unlikely to be an ED in birds. OECD TG 206 cannot detect effects on sexual development and is unlikely to detect epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an ATGT may reveal them.
D	+	-	+	The test chemical may be an ED, but the negative mechanistic data reduce the confidence in this	Further evidence is probably not required.	If the affected endpoint in OECD TG 206 cannot be plausibly linked to the known modality, the test chemical is unlikely to be an ED in birds.

Scenario	Result of OECD TG 206	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				conclusion. However, if the endocrine disruption effects in existing *in vivo* tests can be plausibly linked to the OECD TG 206 responses, this increases the probability that the chemical is an ED in birds.		OECD TG 206 cannot detect effects on sexual development and is unlikely to detect epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an ATGT may reveal them.
E	+	-	-	The test chemical is unlikely to be an ED.	Further evidence is probably not required.	It is possible that the effects observed in OECD TG 206 have been caused by an unknown endocrine mechanism. This would not, however, prevent the chemical being subjected to risk assessment. OECD TG 206 cannot detect effects on sexual development and is unlikely to detect epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an ATGT may reveal them.
F	+	-	Eq/0	The test chemical is unlikely to be an ED, but the relevance of any equivocal existing *in vivo* data to the OECD TG 206 results should be examined.	Further evidence is probably not required.	It is possible that the effects observed in OECD TG 206 have been caused by an unknown endocrine mechanism – equivocal existing *in vivo* data may throw some light on this. The absence of data on a possible endocrine mechanism would, however, not prevent the chemical being subjected to risk

Scenario	Result of OECD TG 206	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						assessment. OECD TG 206 cannot detect effects on sexual development and is unlikely to detect epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an ATGT may reveal them.
G	+	Eq/0	+	The test chemical may be an ED, but the equivocal or absent mechanistic data reduce the confidence in this conclusion. However, if the endocrine disruption effects in existing *in vivo* tests can be plausibly linked to the OECD TG 206 responses, this increases the probability that the chemical is an ED.	If reliable mechanistic data are not available, it would be desirable to obtain some.	The test chemical is almost certainly an ED in birds if a modality identified in the newly commissioned mechanistic screens (see Next Step column), or in the existing *in vivo* data, can be plausibly linked to the affected endpoint. OECD TG 206 cannot detect effects on sexual development and is unlikely to detect epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an ATGT may reveal them.
H	+	Eq/0	-	The test chemical may be an ED, but the equivocal or absent mechanistic data reduce the confidence in this conclusion.	If reliable mechanistic data are not available, it would be desirable to obtain some.	The test chemical is almost certainly an ED in birds if a modality identified in the newly commissioned mechanistic screens (see Next Step column) can be plausibly linked to the affected endpoint. OECD TG 206 cannot detect effects on

Scenario	Result of OECD TG 206	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						sexual development and is unlikely to detect epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an ATGT may reveal them.
I	+	Eq/0	Eq/0	The test chemical may be an ED, but the equivocal or absent mechanistic and *in vivo* data reduce the confidence in this conclusion. Final conclusions about whether a chemical is a possible ED cannot be drawn from the results of this test alone	If reliable mechanistic data are not available, it would be desirable to obtain some.	The test chemical is almost certainly an ED in birds if a modality identified in the newly commissioned mechanistic screens (see Next Step column) can be plausibly linked to the affected endpoint. OECD TG 206 cannot detect effects on sexual development and is unlikely to detect epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an ATGT may reveal them.
J	-	+	+	The chemical is probably not an ED in birds that acts through the mechanisms tested in the available *in vitro* and *in vivo* studies .	If the chemical is strongly bioaccumulative, is suspected to affect sexual development or cause epigenetic effects, consider conducting an ATGT.	If any effects in an ATGT can be plausibly linked with mechanistic data, the test chemical is probably an ED in birds.
K	-	+	-	The chemical is probably not an ED in birds that acts through the	If the chemical is strongly bioaccumulative, is	If any effects in an ATGT can be plausibly linked with mechanistic data, the test chemical is probably an ED in birds.

Scenario	Result of OECD TG 206	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				mechanisms tested in the available *in vitro* and *in vivo* studies .	suspected to affect sexual development or cause epigenetic effects, consider conducting an ATGT.	
L	-	+	Eq/0	The chemical is probably not an ED in birds that acts through the mechanisms tested in the available *in vitro* and *in vivo* studies .	If the chemical is strongly bioaccumulative, is suspected to affect sexual development or cause epigenetic effects, consider conducting an ATGT.	If any effects in an ATGT can be plausibly linked with mechanistic data, the test chemical is probably an ED in birds.
M	-	-	+	The chemical is probably not an ED in birds that acts through the mechanisms tested in the available *in vitro* and *in vivo* studies .	If the chemical is strongly bioaccumulative, is suspected to affect sexual development or cause epigenetic effects, consider conducting an ATGT.	If any effects in an ATGT can be plausibly linked with *in vivo* data which indicate endocrine disruption properties, the test chemical is probably an ED in birds, but likely not by a mechanism covered by the existing *in vitro* screens.
N	-	-	-	The chemical is probably not an ED in birds that acts through the mechanisms tested in the	Further evidence is probably not required.	OECD TG 206 cannot detect effects on sexual development and is unlikely to detect epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an

Scenario	Result of OECD TG 206	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				available *in vitro* and *in vivo* studies .		ATGT may reveal them.
O	-	-	Eq/0	The chemical is probably not an ED in birds that acts through the mechanisms tested in the available *in vitro* and *in vivo* studies .	Further evidence is probably not required.	OECD TG 206 cannot detect effects on sexual development and is unlikely to detect epigenetic effects or effects from long-term bioaccumulation. If these are suspected, an ATGT may reveal them.
P	-	Eq/0	+	The chemical is probably not an ED in birds that acts through the mechanisms tested in the available *in vitro* and *in vivo* studies .	If reliable mechanistic data are not available, it would be desirable to obtain some.	If the newly commissioned mechanistic data are positive and the chemical is strongly bioaccumulative, or if epigenetic or developmental effects are suspected, consider conducting an ATGT.
Q	-	Eq/0	-	The chemical is probably not an ED in birds, but confidence in this conclusion is reduced by the lack of clear mechanistic data.	Further evidence is probably not required, but confidence in the conclusion would be increased by the provision of reliable negative mechanistic data.	If any newly commissioned mechanistic data are positive and the chemical is strongly bioaccumulative, or if epigenetic or developmental effects are suspected, consider conducting an ATGT.
R	-	Eq/0	Eq/0	The chemical may not be an ED in birds, but confidence in this conclusion is reduced by	Further evidence is probably not required, but confidence in the conclusion would be	If any newly commissioned mechanistic data are positive and the chemical is strongly bioaccumulative, or if epigenetic or developmental effects are suspected,

Scenario	Result of OECD TG 206	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				the lack of clear mechanistic and existing *in vivo* data.	increased by the provision of reliable negative mechanistic data.	consider conducting an ATGT.

C4 Mammalian Screens and Tests

C.4.1 OECD TG 440: Uterotrophic Bioassay in Rodents (UT assay) (including OECD GD on the use of the assay to screen for anti-estrogenicity)

C.4.1.1 Modality detected/endpoints: Estrogens (uterine wet weight and dry weight ↑); Anti-estrogens (stimulated uterine weight ↓); (Optional others *e.g.* histopathologic changes in uterus/vagina).

Background to the Assay

C.4.1.2 This assay is a short-term *in vivo* screening assay in female rodents for chemicals that interact with the ER. It is based on the increase in uterine weight (or uterotrophic response) that is elicited by ER agonists in animal models where endogenous estrogen levels are minimal. There are two variants of the assay, one uses immature animals and the other uses OVX animals. The immature rodent assay may detect modalities acting via mechanisms other than ER, as the animals have an intact HPG axis, but the ability to detect these is limited.

C.4.1.3 Non-aromatisable (non-steroidal) androgens and aromatisable androgens that may be metabolised to estrogens, have also been shown to increase uterine weight. In immature animals aromatisable androgens like testosterone elicit histopathologic changes very similar to that of estradiol suggesting that the observed changes are mediated through estrogen. For all other conditions the observed histopathologic changes are different and are considered to be mediated via the AR. In practical terms, this issue is of minor importance. Potentially aromatisable androgens can easily be identified based on their structural features, and non-steroidal androgenic chemicals are currently considered to be rare in the chemical universe. In addition, progesterone and synthetic progestins may also give a positive response (Jones and Edgren, 1973).

C.4.1.4 The Test Guideline is specific for estrogen agonists only. The validation of the assay was not considered adequate for anti-estrogens as there were insufficient pure anti-estrogens available. The test for anti-estrogens however is frequently used and is available as a GD (OECD, 2007a).

When/Why the Assay May be Used

C.4.1.5 Although OECD TG 440 can be used at any stage in the assessment process, the most likely use scenario will be following a positive result in an ER transactivation assay (ER STTA) and/or ER binding assay, in order to determine whether the positive result *in vitro* is translated into a positive result *in vivo*. It may also be used as a screen in the absence of positive *in vitro* data, when a chemical that is negative in the *in vitro* ER-interaction screens is suspected of producing estrogenic metabolites *in vivo*. In this case, the first option would be to use an additional metabolising system in the *in vitro* tests but the uterotrophic assay as an *in vivo* test will include all metabolising systems. Another possible scenario is following observation of effects in higher tier tests, for example acceleration of puberty onset in females, but which are not exclusively indicative of an effect on ER. In the EU,

chemicals included in REACH, Plant Protection Products and Biocides legislation are likely to be tested in OECD TG 416 (Two Generation Reproductive Toxicity Study) and the UT assay may then be used as a follow up. The UT assay is also likely to be carried out as part of the US EPA EDSP Tier 1 screening battery. Selection of the most appropriate tests has to be on a case-by-case basis but also considering the need to minimise animal testing.

C.4.1.6 It should be noted that the UT assay was designed to be sensitive and will detect weak and strong ER modulators. In the validation of the UT assay ethinylestradiol and oestradiol were defined as "strong" estrogens whilst nonylphenol and genistein were defined as "weak" estrogens (OECD, 2006a). Weakly acting chemicals may not always be detected as EDs when tested in higher level tests because the endocrine system in intact/adult animals has a greater ability to compensate than in the UT assay where the HPG axis is disrupted/immature. Furthermore, in case of repeat dose studies, dose levels may need adjustment to lower doses in order to cope with general toxicity.

C.4.1.7 The route of exposure is also an important consideration for the UT assay. OECD TG 440 states that chemicals may be administered by oral or subcutaneous (sc) routes but suggests that the route most relevant for human exposure should be used. The route will have consequences for absorption, distribution, metabolism and excretion and is an important consideration when interpreting results. Methoxychlor, for example, gave negative results when administered by sc injection but positive results when given orally (due to metabolism to estrogenic metabolites) (Laws *et al*, 2000).

Introduction to the Table of Scenarios

C.4.1.8 Table C.4.1 gives guidance on a further step to take in the event of a positive (+) or negative (-) result in the UT assay and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.4.1.9 The results of OECD TG 440/UT assay are given in the second column. Criteria for positive results in OECD TG 440 are given in the test guideline itself *i.e.* a statistically significant increase in uterine weight compared to the solvent control. A positive result in the assay for anti-estrogenicity would be a statistically significant decrease in uterine weight compared to the estrogen-stimulated control group. Negative results are no (statistically significant) changes in wet and dry uterine weight. It is important that quality criteria for control uterine weights are demonstrated. It is also of note that a uterotrophic response may not always be entirely of estrogenic origin *e.g.* testosterone may give a positive result, chemicals interacting with other endocrine axes may give a positive result in the immature rodent assay, diets high in phytoestrogens or energy sources may also give a positive result. Further guidance is provided in the test guideline. Optional endpoints may include histopathologic changes in uterus/vagina or vaginal cornification in the OVX rat assay. These endpoints should supplement the uterotrophic response. Changes in these endpoints in the absence of uterotrophic response should be considered equivocal.

C.4.1.10 Equivocal results for the guideline are not included in Table C.4.1 because these data require further interrogation about the result itself. In the event of an equivocal result, the considerations mentioned above about uterine weights in control animals, non ER-related changes, possible effects of phytoestrogens or high energy diets should be taken into account and further investigations made.

Existing Data to be Considered

C.4.1.11 Existing "Mechanism" *in vitro* data are assumed to be available from ER (ER binding and ER STTA), AR (AR binding and AR STTA) and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. TR-based assays are less relevant for the UT assay. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs *et al.* (2008) and an OECD Detailed Review Paper (OECD, 2008a). These methods, however, have not yet been validated.

C.4.1.12 Existing "Effects" data refer to *in vivo* effects that may come from varied sources and will depend upon the type of chemical (*e.g.* new chemicals, HPV chemical, pesticide). Thus, available data may range from repeated dose toxicity studies (28-day, 90-day) or combined repeat dose/reproductive screening assays to chronic toxicity studies and multi-generation reproductive tests. Some studies fail to identify EDs that weakly affect oestrogen or androgen receptors as was demonstrated on the basis of data generated in the validation process of OECD TG 407 assay with endocrine endpoints. In this validation only moderate EDs, such as nonylphenol and DDE, and strong EDs such as ethinylestradiol and flutamide (acting via ER and AR respectively) were detected. Thus OECD TG 407 cannot be regarded as a screening assay for endocrine activity. This means that when a relatively insensitive test is positive for both endocrine-specific and apical endpoints, this should be taken as an indication that the substance is a possible ED. Caution should be exercised, however, because endocrine endpoints may be impacted secondary to non-endocrine toxicity (*e.g.* hepatotoxicity) and *in vivo* apical endpoints can be affected by all modes of action including endocrine modalities. The ability of a given assay to detect endocrine disruption will also vary depending upon the version of the test guideline used. Older test guidelines may contain fewer endocrine sensitive endpoints than more recent ones. If data are available from single or multi-generation studies that are adequately conducted with updated guidelines that include endpoints sensitive to EDs, then there should be no reason to conduct a UT assay as the higher tier test will provide stronger evidence for hazard and risk assessment. Multi-generation studies conducted prior to the introduction of these endpoints will still provide valuable information on reproductive and endocrine organ toxicity, reproduction and development, but may not be sufficiently sensitive to EDs, in which case the UT assay would provide further valuable information. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.4.1.13 When considering the results of the UT assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.4.1.14 The scenarios (A to R) presented in Table C.4.1 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation

study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study. Further considerations, specific to each scenario are given in the Table.

C.4.1.15 Scenarios A to C represent positive results in the UT assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. A positive result in ER-based assays in combination with a positive UT assay is strong evidence for (anti)estrogenic activity that may or may not be supported by the *in vivo* effects data. Effects on endocrine endpoints in OECD TGs 407, 408, 453 or 421/422 may provide sufficient evidence to conclude concern for endocrine disruption and therefore no need for further screening. Negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. Selection of the dose level and the strain of animal should also be considered. A conclusion of lack of concern for endocrine disruption in the presence of positive effects data can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation. The possibility of other mechanisms should also not be overlooked *e.g.* positive AR-based assays may indicate an aromatisable androgen and a positive steroidogenesis assay could indicate a chemical that alters endogenous estrogen levels, both situations may give a positive result in the immature rat UT assay. Other (non-EATS) mechanisms may also be considered *e.g.* involving other receptors or endocrine axes.

C.4.1.16 Scenarios D to F represent positive results in the UT assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results in the *in vitro* assays should be viewed with caution in case a metabolite is responsible for the positive UT assay. Unless, the metabolic profile of the test substance is known then the first option should be to conduct these assays with an added metabolising system. If the metabolic profile is known, then a higher level *in vivo* test may be advisable. The choice of tests will depend upon the available *in vivo* effects data. As in scenarios A to C, negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. Selection of the dose level and the strain of animal should also be considered. A conclusion of lack of concern for endocrine disruption in the presence of positive effects data (Scenario E) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.4.1.17 Scenarios G to I represent positive results in the UT assay in the presence of various combinations of missing or equivocal data. The next step to take in these eventualities will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step. As above, a conclusion of lack of concern for endocrine disruption in the presence of some positive effects data (Scenario H) can only be made given adequate level 5 assays and sufficient mode of action data to provide a clear interpretation.

C.4.1.18 Scenarios J to L represent negative results in the UT assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. The *in vitro* mechanistic data given in the table could be any of the EATS tests *e.g.* the AR binding or steroidogenesis assay. A weak aromatase inhibitor for example could give Scenario J from a positive result in the steroidogenesis assay and a positive result in the female PP assay. All three scenarios could also arise from a chemical that binds to ER but is metabolised to a non-estrogenic metabolite leading to negative results in the UT assay and this possibility should be investigated first when considering the next step.

Endocrine active potency may also explain differences between *in vitro* and *in vivo* results, *e.g.* a weak chemical may give a positive result *in vitro* but may be negative *in vivo*. Positive *in vivo* effects data may involve other EATS, non-EATS mechanisms (*e.g.* involving other receptors or endocrine axes), more sensitive endpoints, greater statistical power or life stages that are more sensitive to the substance than immature/OVX animals in the UT assay.

C.4.1.19 Scenarios M to O represent negative results in the UT assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend upon the weight of evidence and may not be possible in some cases. However, in the presence of negative data from robust level 4 and 5 assays further animals testing is probably not justified. Where there are positive *in vivo* effects data, there could still be an estrogen-related mechanism. These effects may be related to length of exposure, route of exposure or exposure at different life stages. Other EATS or non-EATS mechanisms may also be involved.

C.4.1.20 Scenarios P to R represent negative results in the UT assay in the presence of various combinations of missing or equivocal data. As with the positive result scenarios above (paragraph C.4.1.17) the next step to take in these eventualities will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.1.21 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.4.1 is meant to provide a succinct guide and may not cover all circumstances or possibilities.

Table C.4.1 OECD TG 440: Uterotrophic Bioassay in Rodents (UT assay) (including OECD GD on the use of the assay to screen for anti-estrogenicity). Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter. These may be repeated dose toxicity tests (*e.g.* OECD TG 407,OECD TG 408 28 and 90-day studies), reproductive tests (*e.g.* reproduction screening assays or 2-generation studies) or read across from chemical analogues.

Scenarios	Result of OECD TG 440 (UT) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	Strong evidence for E/anti-E activity with (potential for) adverse effects via ER mechanism.	Perform assay from level 4 *e.g.* female pubertal assay or level 5 *e.g.* ext-1 or 2-gen assay.	If existing data are from level 4 or 5 (or less sensitive assays) then may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Consider route of exposures for UT assay and existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results and *in vivo.*

Scenarios	Result of OECD TG 440 (UT) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						results but may also be metabolised to a metabolite that also has positive results *in vitro* and *in vivo*. A positive result could have arisen from other (EATS or non-EATS) mechanisms
B	+	+	-	Strong evidence for E/anti-E activity via ER but effects not detected in other *in vivo* studies in intact animals.	Perform assay from level 4 *e.g.* female pubertal assay or level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. Consider route of exposures for UT assay and existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues A positive result could have arisen from other (EATS or non-EATS) mechanisms.
C	+	+	Eq/0	Strong evidence for E/anti-E activity via ER, but no or equivocal data from other *in vivo* studies.	Perform assay from level 4 *e.g.* female pubertal assay or level 5 *e.g.*ext-1 or 2-gen assay.	Check data on chemical analogues. Consider route of exposures for UT assay and existing effects data and possible implications of ADME characteristics of the chemical. Depending on route/kinetic and

Scenarios	Result of OECD TG 440 (UT) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						existing data considerations, may perform assay from upper levels. (levels 4 or 5). A positive result could have arisen from other (EATS or non-EATS) mechanisms
D	+	-	+	Strong evidence for E/anti-E activity. Acts via ER mechanism, but requires metabolic activation. Acts via non-ER mechanism and may or may not require metabolic activation.	Perform ER transactivation assay or binding assay with added metabolising system	If existing data are from level 4 or 5 (or less sensitive assays) then may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
E	+	-	-	Weak evidence for E/anti-E activity. Acts via non-ER mechanism. Chemical requires metabolic activation and metabolite has weak activity. Weak E/anti-E activity via ER does not result in adverse effects.	Perform ER transactivation assay or binding assay with added metabolising system OR perform assay from levels 4 or 5.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. Consider route of exposures for UT assay and existing effects data and

Scenarios	Result of OECD TG 440 (UT) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						possible implications of ADME characteristics of the chemical. Check data on chemical analogues A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
F	+	-	Eq/0	Weak evidence for E/anti-E activity via ER. Acts via non-ER mechanism. Requires metabolic activation and metabolite has weak/equivocal activity.	Perform ER transactivation assay or binding assay with added metabolising system or perform assay from levels 4 or 5.	Check data on chemical analogues. Further mechanistic studies would help determine MoA. Upper level studies will provide hazard data. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
G	+	Eq/0	+	Moderate or strong evidence for E/anti-E activity via ER. May act via ER, metabolic activation is required. Has potential for adverse effects via ER mechanism. May acts via non-ER mechanism and may or may not require metabolic activation	For the "0" scenario, perform ER transactivation assay or binding assay. For the "Eq" scenario perform ER transactivation assay or binding assay with added metabolising	If existing data are from level 4 or 5 (or less sensitive assays) then may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Check data on chemical analogues. Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis

Scenarios	Result of OECD TG 440 (UT) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					system.	
H	+	Eq/0	-	Weak evidence for E/anti-E activity. May act via ER, metabolic activation is required E/anti-E activity does not result in adverse effects.	For the "0" scenario, perform ER transactivation assay or binding assay. For the "Eq" scenario perform ER transactivation assay or binding assay with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. Consider route of exposures for UT assay and existing effects data and possible implications of ADME characteristics of the chemical Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	E/anti-E activity of unknown potency. May act via ER, metabolic activation is required. Unknown potential for adverse effects.	For the "0" scenario, perform ER transactivation assay or binding assay. For the "Eq" scenario perform ER transactivation assay or binding	Check data on chemical analogues. Further mechanistic studies would help determine MoA.

Scenarios	Result of OECD TG 440 (UT) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					assay with added metabolising system, or level 4 or 5 assay if existing data indicates this is needed	
J	-	+	+	No evidence for E/anti-E activity *in vivo* via ER. Route of exposure, metabolic differences or potency explain differences between UT assay and existing *in vitro*/*in vivo* studies Effects seen in existing studies are via non-ER mechanism.	Perform ER transactivation assay or binding assay with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposure for UT assay and possible implications of ADME characteristics of the chemical. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
K	-	+	-	No evidence for E/anti-E activity *in vivo* via ER. Metabolic differences or potency explain *in vitro*/*in vivo* differences.	Perform ER transactivation assay or binding assay with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less

Scenarios	Result of OECD TG 440 (UT) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						sensitive assay then a higher level test may be required. Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* activity is not realised. Consider possible routes of exposure implications of metabolism.
L	-	+	Eq/0	No evidence for E/anti-E activity *in vivo* via ER. Metabolic differences or potency explain *in vitro*/*in vivo* difference. Unknown potential for adverse effects.	Perform ER transactivation assay or binding assay with added metabolising system OR perform assay from levels 4 or 5.	Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* activity is not realised. Consider possible routes of exposure implications of metabolism.
M	-	-	+	No evidence for E/anti-E activity *in vivo* or *in vitro* via ER. Metabolic differences or route of exposure explain *in vitro*/*in vivo* existing differences. Effects seen in existing studies are via non-ER mechanism.	Perform *in vitro* assays with added metabolising system.	Effects seen in existing studies may be in a more sensitive life stage. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
N	-	-	-	No evidence for E/anti-E activity *in vivo* or *in vitro* via ER. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the

Scenarios	Result of OECD TG 440 (UT) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					perform assay from level 4.	most information). Check data on chemical analogues.
O	**-**	**-**	Eq/0	No evidence for E/anti-E activity *in vivo* or *in vitro* via ER. Unknown potential for adverse effects via other non-ER mechanisms.	Perform assay from levels 4 or 5.	Consider route of exposure for UT assay and possible implications for ADME characteristics of the chemical in follow up assay.
P	-	Eq/0	+	No evidence for E/anti-E activity *in vivo* via ER. Unknown potential for adverse effects via other mechanism.	For the "0" scenario perform *in vitro* EATS assays, otherwise Eq result available.	Consider route of exposure for UT assay and possible implications for differences from existing assay. Effects seen in existing studies may be in a more sensitive life stage. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
Q	-	Eq/0	-	No evidence for E/anti-E activity *in vivo* via ER. No evidence of adverse effects.	For the "0" scenario perform *in vitro* EATS assays, otherwise Eq result available.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
R	-	Eq/0	Eq/0	No evidence for E/anti-E activity *in vivo* via ER.	For the "0" scenario perform *in vitro* EATS	Consider route of exposure for UT assay and possible implications for differences from existing assay.

Scenarios	Result of OECD TG 440 (UT) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					assays, otherwise Eq result available.	Check data on chemical analogues. Further mechanistic studies would help determine MoA.

C.4.2 OECD TG 441: Hershberger Bioassay in Rats (H Assay) (including OECD GD for Weanling Hershberger Bioassay)

C.4.2.1 Modality detected/endpoints: Androgens (weights of ventral prostate, seminal vesicles, LABC, cowpers glands, glans penis ↑); Anti-androgens (weights of testosterone stimulated ventral prostate, seminal vesicles, LABC, cowpers glands, glans penis ↓); (Optional others *e.g.* liver, paired kidney, paired adrenal and testis weights, changes in serum hormones). Note: weanling H assay does not include glans penis.

Background to the Assay

C.4.2.2 This assay is a short-term *in vivo* screening assay in male rodents for chemicals that interact with the AR. It is based on changes in weight of the accessory tissues of the male reproductive tract in response to androgens and antiandrogens in animal models where endogenous androgens are minimal as a result of castration or because the animals are immature. The surgically castrated peripubertal rat is the primary model validated for the assay and is described in OECD TG 441. This model is sensitive to androgens and antiandrogens. An alternative model – the intact (uncastrated) weanling rat; was also validated due to animal welfare concerns with the castration procedure but did not seem to consistently detect weak anti-androgenic chemicals at the doses tested, although androgenic chemicals were detected. The castrated peripubertal model is therefore more commonly used because both androgenic and antiandrogenic protocols can be run in the same experiment. The use of the weanling H assay is described in a guidance document (OECD, 2009c). The castrated peripubertal rat model utilises the weights of five androgen-dependent sex accessory tissues (ventral prostate, seminal vesicles, LABC, cowpers glands and glans penis) as the primary endpoints, whilst for the weanling rat model the list does not include the glans penis because the weanling male has not yet achieved preputial separation. Testis weight is an optional endpoint in the weanling model although it should be noted that the weight changes with androgens and antiandrogens are opposite to those seen with the other sex accessory tissues. Serum hormone levels are also optional for both models. These include the thyroid hormones (T3 and T4) so that additional information on thyroid effects may also be obtained, and LH, FSH and testosterone.

C.4.2.3 The castrated peripubertal rat does not have an intact HPG axis and therefore chemicals acting through this mechanism will not be detected. The HPG axis in the weanling rat is intact and therefore it is possible that such chemicals may be detected. In practice, this has not been tested and the immaturity of the animals, plus the co-administration of testosterone in the antiandrogen test, makes this unlikely.

C.4.2.4 Androgenic chemicals cause growth of the sex accessory tissues whilst antiandrogenic chemicals inhibit the growth caused by co-administration of testosterone. Antiandrogens may act either via AR antagonism (*e.g.* flutamide) or they may act via inhibition of the enzyme 5-alpha-reductase (*e.g.* finasteride) which converts testosterone to the more potent dihydrotestosterone. 5-Alpha-reductase inhibitors may be distinguished from AR antagonists in the H assay by a more pronounced effect on the ventral prostate. AR antagonists can also de distinguished from 5-alpha-reductase inhibitors by the use of *in vitro* assays as 5-alpha-reductase inhibitors do not generally interact with AR. At present there are no validated assays for 5-alpha-reductase inhibition although literature methods are available (Lo *et al*, 2007).

C.4.2.5 The growth of the sex accessory tissues may not always be entirely of androgenic origin. High doses of other hormones may give similar responses *e.g.* potent estrogens may increase the weight of seminal vesicles. Chemicals affecting steroid metabolism could also conceivably affect the antiandrogen assay.

When/Why the Assay May be Used

C.4.2.6 Although OECD TG 441 can be used at any stage in the hazard assessment process, the most likely use scenario will be following a positive result in an AR transactivation assay or AR binding assay, in order to determine whether the positive result *in vitro* is translated into a positive result *in vivo*. It may also be used as a screen in the absence of positive *in vitro* data, when a chemical that is negative in the *in vitro* AR-interaction screens is suspected of producing androgenic metabolites *in vitro*. In this case, the first option would be to use an additional metabolising system in the *in vitro* tests but the H assay as an *in vivo* assay will include all metabolising systems. Another possible scenario is following observation of effects in higher tier tests, for example delayed puberty onset in males, but which are not exclusively indicative of an effect on AR. In the EU, chemicals included in REACH, Plant Protection Products and Biocides legislation are likely to be tested in OECD TG 416 (Two Generation Reproductive Toxicity Study) and the H assay may then be used as a follow up. The H assay is also likely to be carried out as part of the US EPA EDSP Tier 1 screening battery. The castrated peripubertal rat assay (as described in OECD TG 441) is mandatory for the US EPA EDSP Tier 1 screening battery and is most likely to be the assay of choice in other testing strategies. Selection of the most appropriate assays has to be on a case-by-case basis but also considering the need to minimise animal testing.

C.4.2.7 It should be noted that the H assay was designed to be sensitive and will detect weak and strong AR modulators and 5-alpha-reductase inhibitors. In the validation of the H assay trenbolone acetate and testosterone were defined as "potent" androgens whilst finasteride was a "potent" antiandrogen. Linuron and vinclozolin were defined as "weak" antiandrogens (OECD, 2008b) but no weak androgens were tested. Weakly acting chemicals may not always be detected as EDs when tested in higher level tests because the endocrine system in intact/adult animals has a greater ability to compensate than in the H assay where the HPG axis is disrupted/immature and in the case of repeat dose studies dose levels may need adjustment to lower doses in order to cope with general toxicity.

C.4.2.8 The route of exposure is also an important consideration for the H assay. OECD TG 441 states that the test substance may be administered by oral or subcutaneous (sc) routes but suggests that the route most relevant for human exposure should be used. The route will have consequences for absorption, distribution, metabolism and excretion and is an important consideration when interpreting results.

Introduction to the Table of Scenarios

C.4.2.9 Table C.4.2 gives guidance on a further step to take in the event of a positive (+) or negative (-) result in the H assay and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.4.2.10 The results of OECD TG 441 are given in the second column. Criteria for positive results in OECD TG 441 are given in the test guideline itself *i.e.* a statistically significant increase (agonism) or decrease (antagonism or 5-alpha reductase inhibition) in weights of two or more of the sex accessory tissues compared to the relevant control and all target tissues showing some change in the relevant direction. In the case of agonists the control is only treated with vehicle for the test substance whilst for antagonists and 5-alpha reductase inhibitors the control is treated with testosterone plus vehicle for the test substance. Negative results are no (statistically significant) changes in weights of the sex accessory tissues compared to the relevant control. Single, isolated changes, would also be considered negative. The guideline suggests that combined evaluation of all sex accessory tissue responses could

be achieved using appropriate multivariate data analysis. It is important that quality criteria (Coeffecients of Variation) for the weights of control sex accessory tissues are demonstrated. Details are given in the test guideline. Note that in the weanling assay, testis weight decreases with agonists and increases with antagonists. Details of the criteria for positive results in this assay are given in the GD (OECD, 2009c)

C.4.2.11 Optional endpoints may include measurement of serum LH, FSH or testosterone. These endpoints should supplement the sex accessory tissue weights and the assay should not be considered to be positive result if changes in these endpoints occur in the absence of weight changes in the primary tissues. In addition, serum T3 and T4 levels may provide useful information on possible effects on the thyroid although measurement of thyroid weight and serum TSH levels would be also useful in this case. They are not considered further here as this is not the primary use of the assay. Measurement of serum testosterone may be useful if induction of liver xenobiotic metabolising enzymes is suspected. The optional endpoint of liver weight would also be very useful. In these cases, increased clearance of testosterone may lead to an apparent anti-androgenic effect on the sex accessory tissues that does not result from interaction with AR.

C.4.2.12 Equivocal results for the guideline are not included in the table because these data require further interrogation about the result itself. This assay is a screen and therefore a clear positive or negative result should be obtained. In the event of an equivocal result, the considerations mentioned above about control sex accessory tissue weights, non AR-related changes should be taken into account and further investigations made.

Existing Data to be Considered

C.4.2.13 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR, and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. As noted above, there is no validated assay available for 5-alpha reductase inhibitors at present and although 5-alpha reductase is present in H295R cells used in the steroidogenesis assay, the assay does not include the required endpoint for this (dihydrotestosterone). Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available (Jacobs *et al*,, 2008) and an OECD Detailed Review Paper (OECD, 2008a). These methods, however, have not yet been validated.

C.4.2.14 Existing "Effects" data refer to *in vivo* effects that may come from varied sources and will depend upon the type of substance (*e.g.* new chemicals, HPV chemical, pesticide). Thus, available data may range from repeated dose toxicity studies (28-day, 90-day) or combined repeat dose/reproductive screening assays to chronic toxicity studies and multi-generation reproductive tests. Some studies fail to identify EDs that weakly affect oestrogen or androgen receptors as was demonstrated on the basis of data generated in the validation process of OECD TG 407 assay with endocrine endpoints. In this validation only moderate EDs, such as nonylphenol and DDE, and strong EDs such as ethinylestradiol and flutamide (acting via ER and AR respectively) were detected. Thus OECD TG 407 cannot be regarded as a screening assay for endocrine activity. This means that when a relatively insensitive test is positive for both endocrine-specific and apical endpoints, this should be taken as an indication that the substance is a possible ED. Caution should be exercised, however, because endocrine endpoints may be impacted secondary to non-endocrine toxicity (*e.g.* hepatotoxicity) and *in vivo* apical endpoints can be affected by all modes of action including endocrine modalities. If data are available from single or multi-generation studies that are adequately conducted with updated guidelines that include endpoints sensitive to EDs, then there should be no reason to conduct a H assay as the higher tier test will provide stronger evidence for hazard and risk assessment. Multi-generation studies conducted prior to the introduction of these endpoints will still provide valuable information on reproductive and endocrine organ toxicity, reproduction and

development, but may not be sufficiently sensitive to EDs, in which case the H assay would provide further valuable information. A decision about whether to conduct further animal tests would, however, need to consider whether sufficient supplementary data may be provided by *in vitro* tests. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.4.2.15 When considering the results of the H assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.4.2.16 The scenarios (A to R) presented in Table C.4.2 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the current two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study. Further considerations, specific to each scenario are given in the Table.

C.4.2.17 Scenarios A to C represent positive results in the H assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. A positive result in AR-based assays in combination with a positive H assay is strong evidence for (anti)androgenic activity that may or may not be supported by the *in vivo* effects data. There may be sufficient evidence to conclude concern for endocrine disruption and therefore no need for further screening. Negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. The possibility of other mechanisms should also not be overlooked *e.g.* positive ER-based assays and a positive result H assay may indicate (anti)estrogenic effects. Other (non-EATS) mechanisms may also be considered *e.g.* involving other receptors or endocrine axes.

C.4.2.18 Scenarios D to F represent positive results in the H assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results in the *in vitro* assays should be viewed with caution in case a metabolite is responsible for the positive H assay. These scenarios may also occur if enhanced metabolism or clearance of testosterone is responsible for the positive H assay. Unless, the metabolic profile of the test substance is known then the first option should be to conduct these assays with an added metabolising system. If the metabolic profile is known, then a higher level *in vivo* test may be advisable. The choice of tests will depend upon the available *in vivo* effects data. As in scenarios A to C, negative *in vivo* existing effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption.

C.4.2.19 Scenarios G to I represent positive results in the H assay in the presence of various combinations of missing or equivocal data. The next step to take in these eventualities will depend

upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.2.20 Scenarios J to L represent negative results in the H assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. The *in vitro* mechanistic data given in the table could be any of the EATS tests *e.g.* the ER binding or steroidogenesis assay. A weak aromatase inhibitor for example could give Scenario J from a positive result in the steroidogenesis assay and a positive result in the female PP assay. All three scenarios could also arise from a chemical that binds to AR but is metabolised to a non-androgenic metabolite leading to negative results in the H assay and this possibility should be investigated first when considering the next step. Endocrine active potency may also explain differences between *in vitro* and *in vivo* results, *e.g.* a weak chemical may give a positive result *in vitro* but may be negative *in vivo*. Positive *in vivo* effects data may involve other EATS, non-EATS mechanisms (*e.g.* involving other receptors or endocrine axes), more sensitive endpoints, greater statistical power or life stages that are more sensitive to the substance than castrated/immature animals in the H assay.

C.4.2.21 Scenarios M to O represent negative results in the H assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend upon the weight of evidence and may not be possible. Where there are positive *in vivo* effects data, there could still be an androgen-related mechanism. The effects may be related to length of exposure, route of exposure or exposure at different life stages. Other EATS or non-EATS mechanisms may also be involved.

C.4.2.22 Scenarios P to R represent negative results in the H assay in the presence of various combinations of missing or equivocal data. As with the positive result scenarios above (paragraph C.2.2.19) the next step to take in these eventualities will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed directly to *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.2.23 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.4.2 is meant to provide a succinct guide and may not cover all circumstances or possibilities.

Table C.4.2. OECD TG 441: Hershberger Bioassay (H assay). Guidance for scenarios of combinations of results with existing data

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter. These may be repeated dose toxicity tests (*e.g.*OECD TG 407,OECD TG 408 28 and 90-day studies), reproductive tests (*e.g.* reproduction screening assays or 2-generation studies) or read across from chemical analogues.

Scenarios	Result of OECD TG 441 (H assay)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	Strong evidence for A/anti-A activity with (potential for) adverse effects via AR mechanism. 5-Alpha reductase inhibitor with (potential for) adverse effects.	Perform assay from upper levels *e.g.* male pubertal assay (level 4) OR ext-1 or 2-gen assay. (level 5).	Check pattern of change across sex tissues for possible 5-alpha reductase inhibition. If existing data are from level 4 or 5 (or less sensitive assays) then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Consider route of exposures for H assay and existing effects data and possible implications of ADME

Scenarios	Result of OECD TG 441 (H assay)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						characteristics of the chemical. The chemical itself may give positive *in vitro* results and *in vivo.* results but may also be metabolised to a metabolite that also has positive results *in vitro* and *in vivo*. A positive result could have arisen from other (EATS or non-EATS) mechanisms
B	+	+	-	Strong evidence for A/anti-A activity via AR but effects not detected in other *in vivo* studies in intact animals. 5-Alpha reductase inhibitor with (potential for) adverse effects but effects not detected in other *in vivo* studies in intact animals.	Perform assay from level 4 *e.g.* male pubertal assay OR level 5 *e.g.* ext-1 or 2-gen assay.	Check pattern of change across sex tissues for possible 5-alpha reductase inhibition. If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. Consider route of exposures for H assay and existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues A positive result could have arisen from other (EATS or non-EATS) mechanisms.

Scenarios	Result of OECD TG 441 (H assay)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
C	+	+	Eq/0	Strong evidence for A/anti-A activity via AR, but no or equivocal data from other *in vivo* studies. 5-Alpha reductase inhibitor with (potential for) adverse effects but no or equivocal data from other *in vivo* studies.	Perform assay from levels 4 or 5 *e.g.* ext)1 or 2-gen assay.	Check pattern of change across sex tissues for possible 5-alpha reductase inhibition. Check data on chemical analogues. Consider route of exposures for H assay and existing effects data and possible implications of ADME characteristics of the chemical. Depending on route/kinetic and existing data considerations, may perform assay from upper levels. (levels 4 or 5). A positive result could have arisen from other (EATS or non-EATS) mechanisms
D	+	-	+	Strong evidence for A/anti-A activity. Acts via AR mechanism, but requires metabolic activation. 5-Alpha reductase inhibitor but requires metabolic activation. Acts via non-AR mechanism and may or may not require metabolic activation.	Perform AR transactivation assay or binding assay with added metabolising system	If existing data are from level 4 or 5 (or less sensitive assays) then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Check pattern of change across sex tissues for possible 5-alpha reductase inhibition. Further mechanistic studies would help determine MoA. A positive result could have arisen

Scenarios	Result of OECD TG 441 (H assay)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
E	+	-	-	Weak evidence for A/anti-A activity via AR but requires metabolic activation. 5-Alpha reductase inhibitor but requires metabolic activation. Chemical requires metabolic activation and metabolite has weak activity. Weak A/anti-A activity/5-Alpha reductase inhibition does not result in adverse effects. Acts via non-AR mechanism.	Perform AR transactivation assay or binding assay with added metabolising system OR perform assay from levels 4 or 5.	Check pattern of change across sex tissues for possible 5-alpha reductase inhibition. If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. Consider route of exposures for H assay and existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis or liver enzyme induction.
F	+	-	Eq/0	Weak evidence for A/anti-A activity via AR but requires metabolic activation. 5-Alpha reductase inhibitor but requires metabolic activation. Requires metabolic activation and metabolite has weak/equivocal activity.	Perform AR transactivation assay or binding assay with added metabolising system	Check pattern of change across sex tissues for possible 5-alpha reductase inhibition. Check data on chemical analogues. Further mechanistic studies would help determine MoA.

Scenarios	Result of OECD TG 441 (H assay)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				Acts via non-AR mechanism.	OR perform assay from levels 4 or 5.	Upper level studies will provide hazard data. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis or liver enzyme induction.
G	+	Eq/0	+	Moderate or strong evidence for A/anti-A activity via AR. May require metabolic activation. 5-Alpha reductase inhibitor. May require metabolic activation. Has potential for adverse effects via AR mechanism or 5-alpha reductase inhibition. May act via non-AR mechanism and may or may not require metabolic activation.	For the "0" scenario, perform AR transactivation assay or binding assay. For the "Eq" scenario perform AR transactivation assay or binding assay with added metabolising system.	Check pattern of change across sex tissues for possible 5-alpha reductase inhibition. If existing data are from level 4 or 5 (or less sensitive assays) then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Check data on chemical analogues. Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
H	+	Eq/0	**-**	Weak evidence for A/anti-A activity. May act via AR, metabolic activation is required. 5-Alpha reductase inhibitor with (potential for) adverse effects but effects not detected in other *in vivo* studies in intact animals. A/anti-A activity/5-Alpha reductase does not	For the "0" scenario, perform AR transactivation assay or binding assay. For the "Eq"	Check pattern of change across sex tissues for possible 5-alpha reductase inhibition. If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption

Scenarios	Result of OECD TG 441 (H assay)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				result in adverse effects.	scenario perform ER transactivation assay or binding assay with added metabolising system.	(the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. Consider route of exposures for H assay and existing effects data and possible implications of ADME characteristics of the chemical Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	A/anti-A activity of unknown potency. May act via AR, metabolic activation is required. 5-Alpha reductase inhibitor of unknown potency. Unknown potential for adverse effects.	For the "0" scenario, perform AR transactivation assay or binding assay. For the "Eq" scenario perform AR transactivation assay or binding assay with added metabolising system, or level 4 or 5 assay if existing data indicates this is	Check pattern of change across sex tissues for possible 5-alpha reductase inhibition. Check data on chemical analogues. Further mechanistic studies would help determine MoA.

Scenarios	Result of OECD TG 441 (H assay)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					needed.	
J	-	+	+	No evidence for A/anti-A activity via AR or 5-Alpha reductase inhibition *in vivo*. Route of exposure, metabolic differences or potency explain differences between H assay and existing *in vitro*/*in vivo* studies Effects seen in existing studies are via non-AR/5-Alpha reductase mechanism.	Perform AR transactivation assay or binding assay with added metabolising system.	If existing data are from an adequate level 5 assay (or less sensitive assays) there may be sufficient information to conclude concern for endocrine disruption (the ext-1 gen assay provides the most information). Consider route of exposure for H assay and possible implications of ADME characteristics of the chemical. Effects seen in existing studies may be in a more sensitive life stage. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
K	-	+	-	No evidence for A/anti-A activity via AR or 5-Alpha reductase inhibition *in vivo*. Metabolic differences or potency explain *in vitro*/*in vivo* differences.	Perform AR transactivation assay or binding assay with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. Metabolic deactivation of chemical may occur *in vivo* so that potential *in*

Scenarios	Result of OECD TG 441 (H assay)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						vitro activity is not realised. Consider possible routes of exposure, implications of metabolism.
L	-	+	Eq/0	No evidence for A/anti-A activity via AR or 5-Alpha reductase inhibition *in vivo*. Metabolic differences or potency explain *in vitro*/*in vivo* difference. Unknown potential for adverse effects.	Perform AR transactivation assay or binding assay with added metabolising system OR perform assay from levels 4 or 5.	Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* activity is not realised. Consider possible routes of exposure, implications of metabolism.
M	-	-	+	No evidence for A/anti-A activity via AR or 5-Alpha reductase inhibition in H assay or *in vitro*. Metabolic differences or route of exposure explain *in vitro*/*in vivo* existing differences. Effects seen in existing studies are via non-AR or non-endocrine mechanism.	Perform *in vitro* assays with added metabolising system.	Effects seen in existing studies may be in a more sensitive life stage. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
N	-	-	-	No evidence for A/anti-A activity via AR or 5-Alpha reductase inhibition *in vivo* or *in vitro*. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may perform assay from levels 4 or 5.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Check data on chemical analogues.

Scenarios	Result of OECD TG 441 (H assay)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
O	**-**	**-**	Eq/0	No evidence for A/anti-A activity via AR or 5-Alpha reductase inhibition *in vivo* or *in vitro*. Unknown potential for adverse effects via other non-AR mechanisms.	Perform assay from levels 4 or 5.	Consider route of exposure for H assay and possible implications for ADME characteristics of the chemical in follow up assay.
P	-	Eq/0	+	No evidence for A/anti-A activity via AR or 5-Alpha reductase inhibition *in vivo*. Unknown potential for adverse effects via other mechanism.	For the "0" scenario perform *in vitro* EATS assays, otherwise Eq result available.	Consider route of exposure for H assay and possible implications for differences from existing assay. Effects seen in existing studies may be in a more sensitive life stage. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
Q	-	Eq/0	-	No evidence for A/anti-A activity via AR or 5-Alpha reductase inhibition *in vivo*. No evidence of adverse effects.	For the "0" scenario perform *in vitro* EATS assays, otherwise Eq result available.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
R	-	Eq/0	Eq/0	No evidence for A/anti-A activity via AR or 5-Alpha reductase inhibition activity *in vivo*.	For the "0" scenario perform *in vitro* EATS	Consider route of exposure for H assay and possible implications for differences from existing assay.

Scenarios	Result of OECD TG 441 (H assay)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					assays, otherwise Eq result available, OR perform level 5 assay.	Check data on chemical analogues. Further mechanistic studies would help determine MoA.

C.4.3 Pubertal Development and Thyroid Function Assay in Peripubertal Male Rats (Male PP Assay) (US EPA OPPTS 890.1500)

C.4.3.1 Modalities detected: (Anti)-Androgen, thyroid, steroidogenesis.

Endpoints: Age and body weight at preputial separation (PPS). Weight of seminal vesicles (+ coagulating gland), ventral prostate, dorsolateral prostate, LABC, epididymides, testes, thyroid, pituitary, adrenals. Histopathologic changes in epididymis, testis, thyroid. Serum testosterone, T4 and TSH.

Background to the Assay

C.4.3.2 This assay is designed to identify chemicals that have the potential to interact with AR-mediated modalities, thyroid hormone mediated modalities and interference with steroidogenesis. It will also detect chemicals that alter pubertal development via changes in the HPG axis. It will also detect ER-mediated effects but the accuracy of this is unknown. The principle of the assay is that male rats are dosed with chemical during the period of sexual maturation, starting at post-natal day 23. The prepubertal period is a very sensitive age for exposure to agents which alter the endocrine system (USEPA, 2007b). Serum androgens in male rats change dramatically during puberty and reproductive organ weights grow rapidly during puberty (Stoker *et al*. 2000). Preputial separation (PPS) is an apical measure of the progression of puberty and it has been used as the primary endpoint of puberty onset in the rat. It is an androgen dependent event. The assay has its female counterpart in the peripubertal female rat assay. Male rats achieve sexual maturity at a later age than females (vaginal opening) and therefore the male assay is of longer duration than the female assay (31 days *c.f.* 21 days) and this should be taken into account when comparing the severity of effects obtained in the two assays.

C.4.3.3 The male PP assay was designed to be one of the suite of assays comprising US EPA's "Tier 1" and has been validated in that context (ref). There is no OECD test guideline for the assay but the US EPA (OPPTS) guideline is available (USEPA, 2009d)). Male and female PP assays are considered to be apical assays *i.e.* they contain endpoints that may be changed by a number of different modes of action and may not be specific to EDs. The animals have intact hypothalamus-pituitary-gonadal/thyroid axes and therefore are a relevant model for human health although the sensitivity of the assays for ER/AR agonists and antagonists are less than that of the UT and H assays. A strength of the PP assays is that (unlike the H and UT assays) they will detect multiple modes of action although it may not be possible to isolate the mechanism of action. The male PP assay is likely to detect ER (ant)agonists in addition to ATS modalities. The estrogen agonist methoxychlor was included in the validation studies of the male assay and gave a weak positive response for some endpoints. Published studies have also demonstrated that the assay responds to strong estrogens such as diethylstilbestrol (Ashby and Lefevre, 2000) and weak estrogens such as nonylphenol (Tan *et al*, 2003). The validation of the male PP assay indicated that sensitivity was high and although it has not been extensively investigated, it showed that the male pubertal assay can be sensitive to dose levels that are near the LOEL in a developmental toxicity study on the androgen antagonist vinclozolin (USEPA, 2007b).

C.4.3.4 A limitation of the validation is that no chemical was shown to be completely negative in the assay. Chloronitrobenzene was included in the validation as a chemical that was expected to be toxic but without endocrine activity, but when tested was positive, delaying PPS, decreasing serum testosterone, decreasing growth of androgen dependent tissues, and reducing T4 levels. It is not known whether these effects were due to non-specificity of the assay or a real effect on endocrine systems. Other chemicals, however, that were positive for one endocrine system were not necessarily positive on others *e.g.* perchlorate altered thyroid hormones and thyroid weight but caused no effects on any of the reproductive tract weights or puberty onset. Another possible limitation is the inability to detect specific aromatase inhibitors. Although more general inhibitors of steroidogenesis (including

aromatase inhibition), such as ketoconazole, are detected in the assay, specific inhibitors of aromatase only, such as fadrozole, were not detected (Marty *et al*, 2001).

When/Why the Assay May be Used

C.4.3.5 As mentioned above, the male PP assay is likely to be used as part of the US EPA Tier 1 screening battery as an apical assay to detect interaction with multiple endocrine systems. In this context its use is primarily for hazard determination. In addition to this specific regulatory application, it may also be used as a follow-up assay following positive results in *in vitro* assays *e.g.* a positive result in the steroidogenesis assay. Positive results in an AR *in vitro* assay would preferably be followed by an H assay for reasons of animal welfare – H assays require fewer animals than the male PP assays and are of shorter duration.If there is a need to test in an apical assay then the PP assay may be chosen, realising the caveat that there is some uncertainty regarding the specificity of the PP assay. Depending upon the number of doses used, the PP assay may be used for hazard assessment (when one or two doses are used) or may contribute to risk assessment if a more detailed dose response curve is available. The assay could potentially also be used to investigate or supplement higher tier data. One scenario could be if only limited reproductive data are available *e.g.* a study not conducted to modern standards or not containing endpoints for sexual development. Data from female and male PP assays could then be used to investigate the occurrence of endocrine effects. A decision about whether to conduct further animal tests would, however, need to consider whether sufficient supplementary data may be provided by *in vitro* tests.

Introduction to the Table of Scenarios

C.4.3.6 Table C 4.3 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.4.3.7 The results of the male PP assay are given in the second column. The assay contains multiple endpoints and it is not possible to provide alternative scenarios for all combinations, therefore some discrimination has been attempted by dividing the endpoints into "apical" and "indicators of hormonal activity". The terminology used has been chosen to be consistent between both the wildlife and mammalian tests. Both groups have similar biological importance, although the "indicators of hormonal activity" in the mammalian assays are serum hormones and are generally, but not always, more variable than "apical endpoints". "Apical endpoints" are age/body weight at PPS weights of seminal vesicles, prostate, LABC, epididymides, testes, thyroid, pituitary and adrenals; histopathologic changes in epididymis, testis, thyroid. "Indicators of hormonal activity" are hormones (testosterone, T4 and TSH).

Three possible outcomes for a positive result are therefore envisaged in Table 4.3:

1) Indicators of hormonal activity and apical endpoints positive
2) Indicators of hormonal activity positive and apical endpoints negative
3) Indicators of hormonal activity negative and apical endpoints positive

C.4.3.8 A positive result for apical endpoints could be delayed puberty (PPS) or statistically significant reductions in weights of the epididymides, prostate and seminal vesicles accompanied by treatment-related histopathologic changes. A positive result for indicators of hormonal activity could

be statistically significant changes in thyroid hormone profiles. The multiple endpoints in this assay means that there is some redundancy in the assay but this is useful as not all chemicals may affect all endpoints associated with a mechanism of action and there may be site-specific differences in response.

C.4.3.9 Single isolated changes may be indicative of spurious results but robust dose response information may not be available as the TG only requires two dose levels. The guidance on histopathologic changes in endocrine tests (OECD, 2009a) may be helpful in interpretation. Such results should be considered with caution although it is possible that weak effects have been detected which may then be seen in longer-term studies.

C.4.3.10 A negative result for the male PP assay is taken to be absence of changes in indicators of hormonal activity and apical endpoints. In the absence of other pertinent lines of evidence, negative results in this test alone cannot be taken as evidence that the substance is not an ED. Further studies will be required as confirmation.

C.4.3.11 Equivocal results for the guideline are not considered in the table, partly for brevity but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative or whether the result must be put to one side and the test repeated (using the same or a different test). Factors which may have interfered with the result *e.g.* composition of the diet used, environmental influences; should be considered. Performance criteria (CVs for the test endpoints) should be checked for compliance with those in the TG. The assay does not include concurrent positive controls but attempts have been made to mitigate this by including the performance criteria.

Existing Data to be Considered

C.4.3.12 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs *et al.* (2008) and an OECD Detailed Review Paper (OECD, 2008a). These methods, however, have not yet been validated.

C.4.3.13 Existing "Effects" data refer to *in vivo* effects that may come from H assays where a non-physiological animal model is used. In these cases, it should be remembered that these assays are specifically designed to be sensitive to EDs. Another possibility is that repeat oral toxicity studies, reproduction/developmental toxicity screen tests or read across from analogues, may be available. It is unlikely that the male PP assay will be performed if data from robust higher tier reproductive studies are already available as the PP assay offers no advantage over these assays. It is possible though that the PP assay has been performed to supplement non-robust higher tier data for the reasons given above. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.4.3.14 When considering the results of the male PP assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.4.3.15 The scenarios (A to R) presented in Table C.4.3 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study. Further considerations, specific to each scenario are given in the Table.

C.4.3.16 Scenarios A to C represent positive results in the male PP assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive male PP result scenario is divided into the three possible outcomes given above. A positive result in the *in vitro* assays in combination with a positive male PP assay is moderate or strong evidence for EATS-mediated activity that may or may not be supported by the *in vivo* effects data. In the absence of robust upper level data, the next step may be to conduct an upper level test. In the presence of robust level 5 data then there may be sufficient evidence to conclude concern for endocrine disruption and therefore no need for further screening. Negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. The possibility of other (non-EATS) mechanisms should also not be overlooked *e.g.* involving other receptors or endocrine axes.

C.4.3.17 Scenarios D to F represent positive results in the male PP assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive male PP result scenario is divided into the three possible outcomes given above. Negative results in the *in vitro* assays should be viewed with caution in case a metabolite is responsible for the positive male PP assay. Unless the metabolic profile of the test substance is known, the first option should be to conduct these assays with an added metabolising system. If the metabolic profile is known, then a higher level *in vivo* test may be advisable. The choice of tests will depend upon the available *in vivo* effects data. As in scenarios A to C, negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption.

C.4.3.18 Scenarios G to I represent positive results in the male PP assay in the presence of various combinations of missing or equivocal data. Each positive male PP result scenario is divided into the three possible outcomes given above. The next step to take in these eventualities will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects.. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.3.19 Scenarios J to L represent negative results in the male PP assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. As a negative result for the male PP is taken to be negative findings for both indicators of hormonal activity and apical

endpoints then (unlike the situation with positive outcomes) there is only one possible negative outcome. All three scenarios could also arise from a chemical that is positive in *in vitro* assays, but is metabolised to a non-active metabolite leading to negative results in the male PP assay. This possibility should be investigated first when considering the next step. Endocrine active potency may also explain differences between *in vitro* and *in vivo* results, *e.g.* a weak chemical may give a positive result *in vitro* but may be negative *in vivo*. Positive *in vivo* effects data may involve other EATS, non-EATS mechanisms (*e.g.* involving other receptors or endocrine axes), more sensitive endpoints, greater statistical power or life stages that are more sensitive to the substance than the young adult animals in the male PP assay.

C.4.3.20 Scenarios M to O represent negative results in the male PP assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend upon the weight of evidence and may not be possible. Where there are positive *in vivo* effects data there could still be an EATS-related mechanism, the effects may be related to length of exposure, route of exposure or exposure at different life stages. Other EATS or non-EATS mechanisms may also be involved.

C.4.3.21 Scenarios P to R represent negative results in the male PP assay in the presence of various combinations of missing or equivocal data. As with the positive result scenarios above (paragraph C.2.3.18) the next step to take in these eventualities will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.3.22 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.4.3 is meant to provide a succinct guide and may not cover all circumstances or possibilities.

Table C.4.3. Pubertal Development and Thyroid Function Assay in Peripubertal Male Rats (Male PP Assay) (OPPTS 890.1500). Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter. These may be repeated dose toxicity tests (*e.g.* OECD TG 407, OECD TG 408 28 and 90-day studies) or read across from chemical analogues.

***Note: three possible outcomes for a positive result are given:

1) Indicators of hormonal activity and apical endpoints positive

2) Indicators of hormonal activity positive and apical endpoints negative

3) Indicators of hormonal activity negative *and apical endpoints positive*

"Apical endpoints" are age/body weight at PPS weights of seminal vesicles, prostate, LABC, epididymides, testes, thyroid, pituitary and adrenals; histopathologic changes in epididymis, testis, thyroid.

"Indicators of hormonal activity" are hormones (testosterone, T4 and TSH).

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+ ***	+	+	1) Increased evidence of (anti)-ATS activity (weak, moderate or strong). 2) Possible evidence of (anti)-ATS activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-ATS activity. 3) Moderate or strong (anti)-ATS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-ATS activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay..	If existing data are from level 5 then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Possible effects on E modality should also be considered. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
B	+	+	-	1) Increased evidence of (anti)-ATS activity (weak, moderate or strong). 2) Possible evidence of (anti)-ATS	Perform assay from level 5 *e.g.*ext-1 or 2-	Question why difference from existing data. Consider route of exposures and

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-ATS activity. 3) Moderate or strong (anti)-ATS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-ATS activity.	gen assay.	possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ATS modalities or other mechanisms. Possible effects on E modality should also be considered.
C	+	+	Eq/0	1) Increased evidence of (anti)-ATS activity (weak, moderate or strong). 2) Possible evidence of (anti)-ATS activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-ATS activity. 3) Moderate or strong (anti)-ATS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-ATS activity.	Perform assay from level 5 *e.g.*ext)1 or 2-gen assay.	Check data on chemical analogues. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ATS modalities or other mechanisms. Possible effects on E modality should also be considered. Consider route of exposure for

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						female PP assay and follow-up assay. Possible implications of ADME characteristics of the chemical.
D	+	-	+	1) Increased evidence of (anti)-ATS activity (weak, moderate or strong). Acts via non-AR, TR, S mechanism or requires metabolic activation for activity. 2) Possible evidence of (anti)-ATS activity, apical endpoints may be less sensitive or unaffected. Acts via non-AR,TR, S mechanism or requires metabolic activation for activity. 3) Moderate or strong (anti)-ATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-AR,TR, S mechanism or requires metabolic activation for activity.	Perform *in vitro* ER, AR,TR, S assays with added metabolising system.	If existing data are from an adequate level 5 assay then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ATS modalities or other mechanisms. Possible effects on E modality should also be considered. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						Further mechanistic studies would help determine MoA.
E	+	-	-	1) Increased evidence of (anti)-ATS activity (weak, moderate or strong). Acts via non-AR, TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data 2) Possible evidence of (anti)-ATS activity, apical endpoints may be less sensitive or unaffected. Acts via non-AR, TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data. 3) Possible evidence of (anti)-ATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-AR, TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.* ext-1 or 2-gen assay.	Question why difference from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate other mechanisms.
F	+	-	Eq/0	1) Increased evidence of (anti)-ATS activity (weak, moderate or strong). Acts via non-AR, TR, S mechanism or	Perform *in vitro* ER, AR, TR, S assays with	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				requires metabolic activation for activity. 2) Possible evidence of (anti)-ATS activity, apical endpoints may be less sensitive or unaffected. Acts via non-AR,TR, S mechanism or requires metabolic activation for activity 3) Moderate (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-AR, TR, S mechanism or requires metabolic activation for activity.	added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay..	endpoints. Effects on apical endpoints alone may indicate other mechanisms. Check data on chemical analogues. Further mechanistic studies may help determine MoA. Upper level studies will provide hazard data.
G	+	Eq/0	+	1) Increased evidence of (anti)-ATS activity (weak, moderate or strong). May act via AR, TR, S mechanism. (metabolic activation may be needed) 2) Possible evidence of (anti)-ATS activity, apical endpoints may be less sensitive or unaffected. May act via AR,TR, S mechanism (metabolic activation may be needed). 3) Moderate or strong (anti)- ATS activity, indicators of hormonal activity may be less sensitive or unaffected.	Perform *in vitro* ER, AR, TR, S assays. (for the "0" scenario, otherwise Eq result available) OR Perform *in vitro* ER, AR, TR, S assays with added metabolising	If existing data are from level 5 then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ATS modalities or

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				Increased evidence of (anti)-ATS activity. May act via AR, TR, S mechanism. (metabolic activation needed)	system.	other mechanisms. Possible effects on E modality should also be considered. Check data on chemical analogues. Further mechanistic studies would help determine MoA. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
H	+	Eq/0	-	1) Increased evidence of (anti)-ATS activity (weak, moderate or strong). Acts via unknown mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data 2) Possible evidence of (anti)-ATS activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism Route of exposure may account for the differences from existing data 3) Moderate (anti)- ATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via	For the "0" scenario, perform *in vitro* ER, AR, TR, S assays with added metabolising system. (otherwise Eq result available)	Question why difference from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of subtle changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ATS modalities or other mechanisms.

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				unknown mechanism. Route of exposure may account for the differences from existing data		Possible effects on E modality should also be considered. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	1) Increased evidence of (anti)-ATS activity (weak, moderate or strong). Acts via unknown mechanism. Unknown potential for adverse effects. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. 3) Moderate or strong (anti)-ATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. There may be a need for metabolic activation.	Perform *in vitro* ER, AR, TR, S assays. Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR perform assay from level 5 *e.g.*(ext)-1 or 2-gen assay.	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ATS modalities or other mechanisms. Possible effects on E modality should also be considered. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
J	-	+	+	No evidence for ATS activity in male PP	Perform *in vitro*	If existing data are from an

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				assay. Metabolism or potency explains the difference from existing *in vitro* and *in vivo* data. Effects seen in existing studies are via non-ATS mechanism.	ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*(ext)-1 or 2-gen assay.	adequate level 5 assay then question why differences. If data are from H assay then this may be more sensitive than male PP assay. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical. Further mechanistic studies would help determine MoA.
K	-	+	-	No evidence for ATS activity in male PP assay. Metabolism or potency explains *in vitro*/*in vivo* differences.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. If data are from H assay then need

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						to conduct higher tier assay to conclude absence of concern for endocrine disruption. Further mechanistic studies with metabolism may help determine MoA.
L	-	+	Eq/0	No evidence for ATS activity in male PP assay. Metabolism or potency explains *in vitro*/*in vivo* differences. Unknown potential for adverse effects.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* activity is not realised. Consider possible routes of exposure implications of metabolism.
M	-	-	+	No evidence for ATS activity in male PP assay. Effects seen in existing studies are via non-ATS mechanism.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay then question why differences (the ext-1 gen assay provides the most information).

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						If data are from H assay then this may be more sensitive than male PP assay. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical.
N	-	-	-	No evidence for ATS activity in male PP assay. No evidence for (anti)-ATS activity *in vitro*. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may perform assay from level 5 *e.g.* ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information).
O	-	-	Eq/0	No evidence for ATS activity in male PP assay. No evidence for (anti)-ATS activity *in vitro*. Unknown potential for adverse effects.	Perform assay from level 5 *e.g.*(ext)-1 or 2-gen assay.	Consider route of exposures and possible implications for ADME characteristics of the chemical in follow-up assay.
P	-	Eq/0	+	No evidence for ATS activity in female PP assay.	Perform *in vitro* ER, AR,TR, S	Consider route of exposure and possible implications for

Scenarios	Result of male PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				Potential for adverse effects via unknown mechanism.	assays	differences from existing assay. If data are from H assay then this may be more sensitive than male PP assay. Effects seen in existing studies may be in a more sensitive life stage.
Q	-	Eq/0	-	No evidence for ATS activity in male PP assay. No evidence of adverse effects.	Perform *in vitro* ER, AR,TR, S assays	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Further mechanistic studies would increase evidence.
R	-	Eq/0	Eq/0	No evidence for ATS activity in male PP assay.	Perform *in vitro* ER, AR,TR, S assays, otherwise Eq result available.	Further mechanistic studies would increase evidence. Check data on chemical analogues.

C.4.4 Pubertal Development and Thyroid Function Assay in Peripubertal Female Rats (Female PP Assay) (US EPA OPPTS 890.1450)

C.4.4.1 Modalities detected: (Anti)-Estrogen, thyroid, steroidogenesis.

Endpoints: Age and body weight at vaginal opening (VO). Weight of ovaries, uterus, thyroid, pituitary, adrenals. Histopathologic changes in ovaries, uterus, thyroid. Serum T4 and TSH. Age at first vaginal estrus after VO, estrus cyclicity parameters.

Background to the Assay

C.4.4.2 This assay is designed to identify chemicals that have the potential to interact with ER-mediated modalities, thyroid hormone mediated modalities and interference with steroidogenesis. It will also detect chemicals that alter pubertal development via changes in the HPG axis. The principle of the assay is that female rats are dosed with chemical during period of sexual maturation, starting at post-natal day 22. The prepubertal period is a very sensitive age for exposure to agents which alter the endocrine system (Goldman *et al*., 2000). Sexual maturation is determined in females as vaginal opening (VO) (or patency) and is an estrogen-dependent event that follows the first period of ovarian follicular growth (Goldman *et al*., 2000). The assay has its male counterpart in the peripubertal male rat assay. Female rats achieve sexual maturity at an earlier age than males (PPS) and therefore the female assay is of shorter duration than the male assay (21 days *c.f.* 31 days) and this should be taken into account when comparing the severity of effects obtained in the two assays.

C.4.4.3 The female PP assay was designed to be one of the suite of assays comprising US EPA's "Tier 1" and has been validated in that context (USEPA, 2007c). There is no OECD test guideline for the assay but the US EPA (OPPTS) guideline is available (USEPA, 2009e).). Male and female PP assays are considered to be apical assays *i.e.* they contain endpoints that may be changed by a number of different modes of action and may not be specific to EDs. The animals have intact hypothalamus-pituitary-gonadal/thyroid axes and therefore are a relevant model for human health although the sensitivity of the assays for ER/AR agonists and antagonists are less than that of the UT and H assays. A strength of the PP assays is that (unlike the H and UT assays) they will detect multiple modes of action although it may not be possible to isolate the mechanism of action. The female PP assay is likely to detect AR (ant)agonists in addition to ETS modalities although androgen agonists and antagonists were not included in the validation studies of the female assay. The validation of the female PP assay indicated that sensitivity was high and although it has not been extensively investigated, it appeared to provide a good estimate of the NOEL/LOELs obtained in studies of similar or longer duration *e.g.* the LOAEL for ethinylestradiol in the female PP assay was similar to that for reproductive effects in a multigenerational study (USEPA, 2007c).

C.4.4.3a A limitation of the validation is that no chemical was shown to be completely negative in the assay. Chloronitrobenzene was included in the validation as a chemical that was expected to be toxic but without endocrine activity, but when tested was positive in the assay, delaying VO, reducing uterine weight, reducing T4 levels and increasing TSH levels. It is not known whether these effects were due to non-specificity of the assay or a real effect on endocrine systems. Other chemicals, however, that were positive for one endocrine system were not necessarily positive on others *e.g.* propylthiouracil altered thyroid hormones and thyroid weight but caused no effects on any of the reproductive tract weights or puberty onset.

When/Why the Assay May be Used

C.4.4.4 As mentioned above, the female PP assay is likely to be used as part of the US EPA Tier 1 screening battery as an apical assay to detect interaction with multiple endocrine systems. In this

context its use is primarily for hazard determination. In addition to this specific regulatory application, it may also be used as a follow-up assay following positive results in *in vitro* assays *e.g.* a positive result in the steroidogenesis assay. Positive results in an ER *in vitro* assay would preferably be followed by a UT assay for reasons of animal welfare – UT assays require fewer animals than the female PP assays and are of shorter duration. If there is a need to test in an apical assay then the PP assay may be chosen, realising the caveat that there is some uncertainty regarding the specificity of the PP assay. Depending upon the number of doses used, the PP assay may be used for hazard assessment (when one or two doses are used) or may contribute to risk assessment if a more detailed dose-response curve is available. The assay could potentially also be used to investigate or supplement higher tier data. One scenario could be if only limited reproductive data are available *e.g.* a study not conducted to modern standards or not containing endpoints for sexual development. Data from female and male PP assays could then be used to investigate the occurrence of endocrine effects. A decision about whether to conduct further animal tests would, however, need to consider whether sufficient supplementary data may be provided by *in vitro* tests.

Introduction to the Table of Scenarios

C.4.4.5 Table C 4.4 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.4.4.6 The results of the female PP assay are given in the second column. The assay contains multiple endpoints and it is not possible to provide alternative scenarios for all combinations, therefore some discrimination has been attempted by dividing the endpoints into "apical" and "indicators of hormonal activity". The terminology used has been chosen to be consistent between both the wildlife and mammalian tests. Both groups have similar biological importance, although the "indicators of hormonal activity" in the mammalian assays are serum hormones and are generally, but not always, more variable than "apical endpoints". "Apical endpoints" are age/body weight at VO, estrus cyclicity parameters, weights of ovaries, uterus, thyroid, pituitary and adrenals; histopathologic changes in ovaries and uterus. "Indicators of hormonal activity" are hormones (T4 and TSH).

Three possible outcomes for a positive result are therefore envisaged in Table 4.4:

1) Indicators of hormonal activity and apical endpoints positive
2) Indicators of hormonal activity positive and apical endpoints negative
3) Indicators of hormonal activity negative and apical endpoints positive

C.4.4.7 A positive result for apical endpoints could be delayed puberty (VO) or statistically significant reductions in uterine weights, accompanied by treatment-related histopathologic changes. A positive result for indicators of hormonal activity could be statistically significant changes in hormone profiles. The multiple endpoints in this assay means that there is some redundancy in the assay but this is useful as not all chemicals may affect all endpoints associated with a mechanism of action and there may be site-specific differences in response.

C.4.4.8 Single isolated changes may be indicative of spurious results but robust dose response information may not be available as the TG only requires two dose levels. The guidance on histopathologic changes in endocrine tests (OECD, 2009a) may be helpful in interpretation. Such results should be considered with caution although it is possible that these endpoints may have

detected weak effects that were not detected by the apical endpoints in this study but may then be detected in longer-term studies.

C.4.4.9 A negative result for the female PP assay is taken to be absence of changes in both endocrine relevant indicators of hormonal activity and apical endpoints. In the absence of other pertinent lines of evidence negative results in this test alone cannot be taken as evidence that the substance is not an ED. Further studies will be required as confirmation.

C.4.4.10 Equivocal results for the guideline are not considered in the table, partly for brevity but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative or whether the result must be put to one side and the test repeated (using the same or a different test). Factors which may have interfered with the result *e.g.* composition of the diet used, environmental influences; should be considered. Performance criteria (CVs for the test endpoints) should be checked for compliance with those in the TG. The assay does not include concurrent positive controls but attempts have been made to mitigate this by including the performance criteria.

Existing Data to be Considered

C.4.4.11 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs *et al.* (2008) and an OECD Detailed Review Paper (OECD, 2008a). These methods, however, have not yet been validated.

C.4.4.12 Existing "Effects" data refer to *in vivo* effects that may come from UT assays where a non-physiological animal model is used. In these cases, it should be remembered that these assays are specifically designed to be sensitive to EDs. The immature rodent UT assay is also sensitive to activities other than ER (ant)agonism, including changes resulting from energy intake (Odum *et al*, 2004). Another possibility is that repeat oral toxicity studies, reproduction/developmental toxicity screen tests or read across from analogues, may be available. It is unlikely that the female PP assay will be performed if data from robust higher tier reproductive studies are already available as the PP assay offers no advantage over these assays. It is possible though that the PP assay has been performed to supplement non-robust higher tier data for the reasons given above. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.4.4.13 When considering the results of the female PP assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.4.4.14 The scenarios (A to R) presented in Table C.4.4 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level

tests should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study. Further considerations, specific to each scenario are given in the Table.

C.4.4.15 Scenarios A to C represent positive results in the female PP assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive female PP result scenario is divided into the three possible outcomes given above. A positive result in the *in vitro* assays in combination with a positive female PP assay is moderate or strong evidence for EATS-mediated activity that may or may not be supported by the *in vivo* effects data. In the absence of robust upper level data, the next step may be to conduct an upper level test. In the presence of robust level 5 data then there may be sufficient evidence to conclude concern for endocrine disruption and therefore no need for further screening. Negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. The possibility of other (non-EATS) mechanisms should also not be overlooked *e.g.* involving other receptors or endocrine axes.

C.4.4.16 Scenarios D to F represent positive results in the female PP assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive female PP result scenario is divided into the three possible outcomes given above. Negative results in the *in vitro* assays should be viewed with caution in case a metabolite is responsible for the positive female PP assay. Unless the metabolic profile of the test substance is known, the first option should be to conduct these assays with an added metabolising system. If the metabolic profile is known, then a higher level *in vivo* test may be advisable. The choice of tests will depend upon the available *in vivo* effects data. As in scenarios A to C negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption.

C.4.4.17 Scenarios G to I represent positive results in the female PP assay in the presence of various combinations of missing or equivocal data. Each positive female PP result scenario is divided into the three possible outcomes given above. The next step to take in these eventualities will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects.. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.4.18 Scenarios J to L represent negative results in the female PP assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. As a negative result for the female PP is taken to be negative findings for both indicators of hormonal activity and apical endpoints then (unlike the situation with positive outcomes) there is only one possible negative outcome. All three scenarios could also arise from a chemical that is positive in *in vitro* assays, but is metabolised to a non-active metabolite leading to negative results in the female PP assay. This possibility should be investigated first when considering the next step. Endocrine active potency may also explain differences between *in vitro* and *in vivo* results, *e.g.* a weak chemical may give a positive result *in vitro* but may be negative *in vivo*. Positive *in vivo* effects data may involve other EATS, non-

EATS mechanisms (*e.g.* involving other receptors or endocrine axes), more sensitive endpoints, greater statistical power or life stages that are more sensitive to the substance than the young adult animals in the female PP assay.

C.4.4.19 Scenarios M to O represent negative results in the female PP assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend upon the weight of evidence and may not be possible. Where there are positive *in vivo* effects data there could still be an EATS-related mechanism, the effects may be related to length of exposure, route of exposure or exposure at different life stages. Other EATS or non-EATS mechanisms may also be involved.

C.4.4.20 Scenarios P to R represent negative results in the female PP assay in the presence of various combinations of missing or equivocal data. As with the positive result scenarios above (paragraph C.2.4.17) the next step to take in these eventualities will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.4.21 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.4.4 is meant to provide a succinct guide and may not cover all circumstances or possibilities.

Table C.4.4. Pubertal Development and Thyroid Function Assay in Peripubertal Female Rats (Female PP Assay) (US EPA OPPTS 890.1450). Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter. These may be repeated dose toxicity tests (*e.g.* OECD TG 407, OECD TG 408 28 and 90-day studies) or read across from chemical analogues.

***Note: three possible outcomes for a positive result are given:

1) Indicators of hormonal activity and apical endpoints positive

2) Indicators of hormonal activity positive and apical endpoints negative

3) Indicators of hormonal activity negative *and apical endpoints positive*

"Apical endpoints" are age/body weight at VO, estrus cyclicity parameters, weights of ovaries, uterus, thyroid, pituitary and adrenals; histopathologic changes in ovaries and uterus.

"Indicators of hormonal activity" are hormones (T4 and TSH).

Scenarios	Result of female PP assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
A	+ ***	+	+	1) Increased evidence of (anti)-ETS activity (weak, moderate or strong). 2) Possible evidence of (anti)-ETS activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-ETS activity. 3) Moderate or strong (anti)-ETS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-ETS activity.	Perform assay from level 5 *e.g.* ext-1 or 2-gen assay.	If existing data are from level 5 then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Possible effects on A modality should also be considered. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
B	+	+	-	1) Increased evidence of (anti)-ETS activity (weak, moderate or strong). 2) Possible evidence of (anti)-ETS activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-ETS	Perform assay from level 5 *e.g.* ext-1 or 2-gen assay	Question why difference from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical.

Scenarios	Result of female PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				activity. 3) Moderate or strong (anti)-ETS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-ETS activity.		If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ETS modalities or other mechanisms. Possible effects on A modality should also be considered.
C	+	+	Eq/0	1) Increased evidence of (anti)-ETS activity (weak, moderate or strong). 2) Possible evidence of (anti)-ETS activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-ETS activity. 3) Moderate or strong (anti)-ETS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-ETS activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	Check data on chemical analogues. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ETS modalities or other mechanisms. Possible effects on A modality should also be considered. Consider route of exposure for

Scenarios	Result of female PP assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						female PP assay and follow-up assay. Possible implications of ADME characteristics of the chemical.
D	+	-	+	1) Increased evidence of (anti)-ETS activity (weak, moderate or strong). Acts via non-ER, TR, S mechanism or requires metabolic activation for activity. 2) Possible evidence of (anti)-ETS activity, apical endpoints may be less sensitive or unaffected. Acts via non-ER,TR, S mechanism or requires metabolic activation for activity. 3) Moderate or strong (anti)-ETS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-ER,TR, S mechanism or requires metabolic activation for activity.	Perform *in vitro* ER, AR,TR, S assays with added metabolising system.	If existing data are from an adequate level 5 assay then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ETS modalities or other mechanisms. Possible effects on A modality should also be considered. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical. Further mechanistic studies would help determine MoA.
E	+	-	-	1) Increased evidence of (anti)-ETS activity	Perform *in vitro*	Question why difference from

Scenarios	Result of female PP assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				(weak, moderate or strong). Acts via non-ER, TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data 2) Possible evidence of (anti)-ETS activity, apical endpoints may be less sensitive or unaffected. Acts via non-ER, TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data. 3) Possible evidence of (anti)-ETS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-ER, TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data.	ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate other mechanisms.
F	+	-	Eq/0	1) Increased evidence of (anti)-ETS activity (weak, moderate or strong). Acts via non-ER, TR, S mechanism or requires metabolic activation for activity. 2) Possible evidence of (anti)-ETS activity, apical endpoints may be less sensitive or	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate other mechanisms.

Scenarios	Result of female PP assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				unaffected. Acts via non-ER,TR, S mechanism or requires metabolic activation for activity 3) Moderate (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-ER, TR, S mechanism or requires metabolic activation for activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Check data on chemical analogues. Further mechanistic studies may help determine MoA. Upper level studies will provide hazard data.
G	+	Eq/0	+	1) Increased evidence of (anti)-ETS activity (weak, moderate or strong). May act via ER, TR, S mechanism. (metabolic activation may be needed) 2) Possible evidence of (anti)-ETS activity, apical endpoints may be less sensitive or unaffected. May act via ER, TR, S mechanism (metabolic activation may be needed). 3) Moderate or strong (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-ETS activity. May act via ER, TR, S mechanism. (metabolic activation needed)	Perform *in vitro* ER, AR, TR, S assays. (for the "0" scenario, otherwise Eq result available) OR Perform *in vitro* ER, AR, TR, S assays with added metabolising system.	If existing data are from level 5 then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ETS modalities or other mechanisms. Possible effects on A modality should also be considered. Check data on chemical analogues. Further mechanistic studies would help determine MoA.

Scenarios	Result of female PP assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
H	+	Eq/0	-	1) Increased evidence of (anti)-ETS activity (weak, moderate or strong). Acts via unknown mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data 2) Possible evidence of (anti)-ETS activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism Weak activity does not result in adverse effects. 3) Moderate (anti)-ETS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via unknown mechanism. Weak activity does not result in adverse effects.	For the "0" scenario, perform *in vitro* ER, AR, TR, S assays with added metabolising system. (otherwise Eq result available)	Question why difference from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of subtle changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ETS modalities or other mechanisms. Possible effects on A modality should also be considered. Check data on chemical analogues. Further mechanistic studies would

Scenarios	Result of female PP assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						help determine MoA.
I	+	Eq/0	Eq/0	1) Increased evidence of (anti)-ETS activity (weak, moderate or strong). Acts via unknown mechanism. Unknown potential for adverse effects. 2) Possible evidence of (anti)-ETS activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. 3) Moderate or strong (anti)-ETS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. There may be a need for metabolic activation.	Perform *in vitro* ER, AR, TR, S assays. Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR perform assay from level 5 *e.g.*ext-1 or 2-gen assay	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate ETS modalities or other mechanisms. Possible effects on A modality should also be considered. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
J	-	+	+	No evidence for ETS activity in female PP assay. Metabolism or potency explains the difference from existing *in vitro* and *in vivo* data. Effects seen in existing studies are via non-ETS mechanism.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen	If existing data are from an adequate level 5 assay then question why differences. If data are from UT assay then this may be more sensitive than female PP assay. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and

Scenarios	Result of female PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					assay	possible implications of ADME characteristics of the chemical. Further mechanistic studies would help determine MoA.
K	-	+	-	No evidence for ETS activity in female PP assay. Metabolism or potency explains *in vitro*/*in vivo* differences.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. If data are from UT assay then need to conduct higher tier assay to conclude absence of concern for endocrine disruption. Further mechanistic studies with metabolism may help determine MoA.
L	-	+	Eq/0	No evidence for ETS activity in female PP	Perform *in vitro*	Metabolic deactivation of chemical

Scenarios	Result of female PP assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				assay. Metabolism or potency explains *in vitro*/*in vivo* differences. Unknown potential for adverse effects.	ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	may occur *in vivo* so that potential *in vitro* activity is not realised. Consider possible routes of exposure implications of metabolism.
M	-	-	+	No evidence for ETS activity in female PP assay. Effects seen in existing studies are via non-ETS mechanism.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	If existing data are from an adequate level 5 assay then question why differences. If data are from UT assay then this may be more sensitive than female PP assay. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical.
N	-	-	-	No evidence for ETS activity in female PP assay. No evidence for (anti)-ETS activity *in vitro*. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the

Scenarios	Result of female PP assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
					perform assay from level 5 *e.g.*ext-1 or 2-gen assay	ext-1 gen assay provides the most information).
O	-	-	Eq/0	No evidence for ETS activity in female PP assay. No evidence for (anti)-ETS activity *in vitro*. Unknown potential for adverse effects.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	Consider route of exposures and possible implications for ADME characteristics of the chemical in follow-up assay.
P	-	Eq/0	+	No evidence for ETS activity in female PP assay. Potential for adverse effects via unknown mechanism.	Perform *in vitro* ER, AR,TR, S assays	Consider route of exposure and possible implications for differences from existing assay. If data are from H assay then this may be more sensitive than female PP assay. Effects seen in existing studies may be in a more sensitive life stage.
Q	-	Eq/0	-	No evidence for ETS activity in female PP assay. No evidence of adverse effects.	Perform *in vitro* ER, AR,TR, S assays	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information).

Scenarios	Result of female PP assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						Further mechanistic studies would increase evidence.
R	-	Eq/0	Eq/0	No evidence for ETS activity in female PP assay.	Perform *in vitro* ER, AR,TR, S assays, otherwise Eq result available.	Further mechanistic studies would increase evidence. Check data on chemical analogues.

C.4.5 OECD TG 407: Repeated Dose 28-Day Oral Toxicity Study in Rodents

C.4.5.1 Modalities detected: (Anti)-Estrogen, (anti)-androgen, thyroid, steroidogenesis.

C.4.5.2 Endpoints: Mandatory: Weight of adrenals, testes, epididymides, prostate + seminal vesicles with coagulating glands. Histopathologic changes in testes, epididymides, prostate + seminal vesicles with coagulating glands, ovary, uterus/cervix, vagina, thyroid gland and adrenals.

C.4.5.3 Optional: Weight of uterus, ovaries. Estrous cyclicity. Histopathologic changes in mammary glands and pituitary. Circulating levels of T3, T4, TSH.

Background to the Assay

C.4.5.4 This assay determines the general toxicity of chemicals after 28 days of oral dosing *e.g.* effects on liver, kidneys, heart, lungs; it also provides information on effects on the nervous, immune and reproductive systems. This is the primary purpose of this assay. It underwent a validation study where more parameters suitable for the detection of EDs were included. Following the validation study many of the parameters were included in the updated guideline, as either mandatory or optional endpoints. It is important that the collection of endocrine endpoints does not interfere with the primary purpose *e.g.* collection of blood for hormones should ideally be carried out at a comparable time of day in case of diurnal variations but blood collection for clinical chemistry should take precedence. OECD TG 407 is considered to be an apical assay *i.e.* it contains endpoints that may be changed by a number of different modes of action and may not be specific to EDs. The animals are young adults with intact hypothalamus-pituitary-gonadal/thyroid axes and therefore are a relevant model for human health although the sensitivity of the assay for EDs is less than that of the UT and H assays. The validation of the assay for endocrine endpoints showed that this assay is relatively insensitive and would only detect chemicals that were moderate and strong EDs for (anti)-estrogenicity and (anti)-androgenicity (*e.g.* ethinylestradiol and flutamide). However, it did detect EDs that were weak and strong modulators of thyroid hormone-related effects (*e.g.* propylthiouracil and methyl testosterone). It may also detect steroidogenesis inhibition although only one (potent) chemical was used in the validation study (CGS 18320B) (OECD, 2006b). Endocrine modalities other than EATS may also be detected although these have not been validated.

When/Why the Assay May be Used

C.4.5.5 This assay is likely to be used as a preliminary study for longer-term studies *e.g.* 90-day studies or carcinogenicity studies, where the endocrine endpoints give additional information on the potential of the chemical to interact with the endocrine system. This assay is also necessary as a standard information requirement in certain chemical legislations (e.g. REACH for chemicals manufactured or imported in quantities of 10 tonnes or more). It may also be used for chemicals chronic exposure scenarios are not anticipated. Depending upon the number of doses used, the assay may be used for hazard assessment (when one or two doses are used) or for risk assessment if a more detailed dose response curve is available. It should be noted that, as this assay is not primarily designed to detect endocrine disruption, a higher degree of systemic toxicity is typically induced than is the case with the other level 3 and 4 assays. The possibly confounding effect of systemic toxicity on endocrine endpoints therefore needs to be considered.

Introduction to the Table of Scenarios

C.4.5.6 Table C.4.5 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.4.5.7 The results of OECD TG 407 are given in the second column. As OECD TG 407 is not a screening test where a yes/no (qualitative) answer is obtained, criteria for positive results for the endocrine endpoints are not given in the test guideline. Results for the endpoints would be considered both individually and as a whole. It is not possible to provide guidance on all endpoints individually and therefore the endpoints have been pragmatically divided into "apical" and "indicators of hormonal activity". The terminology used has been chosen to be consistent between both the wildlife and mammalian tests. Both groups have similar biological importance, although the "indicators of hormonal activity" in the mammalian assays are serum hormones and are generally, but not always, more variable than "apical endpoints".

"Apical" endpoints are weights of testes, epididymides, prostate (+ seminal vesicles with coagulating glands), ovary, uterus, histopathologic changes in testes, epididymides, prostate, seminal vesicles, coagulating glands, ovary, uterus/cervix, vagina, thyroid and estrous cyclicity. "Indicators of hormonal activity" are hormones (T3, T4 and TSH).

Three possible outcomes for a positive result are therefore envisaged in Table C.4.5:

1) Indicators of hormonal activity and apical endpoints positive
2) Indicators of hormonal activity positive and apical endpoints negative
3) Indicators of hormonal activity negative and apical endpoints positive

C.4.5.8 A positive result for apical endpoints could be dose-related reductions in reproductive organ weights, accompanied by treatment-related histopathologic changes. A positive result for indicators of hormonal activity could be statistically significant changes in thyroid hormone profiles. The indicators of hormonal activity are optional endpoints for this test guideline and therefore they may not be measured. Alternatively other endpoints not specified in the guideline *e.g.* reproductive hormones, may be measured and if positive would contribute to the overall assessment of a positive result. The apical endpoints for the detection of effects on male and female reproductive organs tended to be less sensitive than the indicators of hormonal activity in the validation of the OECD TG 407 and therefore changes are more likely to be indicative of an ED although the results in entirety should be considered rather than single isolated changes. This was not true for the thyroid though where changes in thyroid histopathology were always as sensitive, or more sensitive, than changes in thyroid hormone/TSH levels. The guidance on histopathologic changes in endocrine tests (OECD, 2009a) may be helpful in interpretation. A positive result for indicators of hormonal activity alone should be considered with caution although it is possible that these endpoints may have detected weak effects that were not detected by the apical endpoints in this study but may then be detected in longer-term studies.

C.4.5.9 A negative result for the OECD TG 407 is taken to be absence of changes in both endocrine relevant indicators of hormonal activity and apical endpoints. In the absence of other pertinent lines of evidence negative results in this test alone cannot be taken as evidence that the chemical is not an ED. Further studies will be required as confirmation.

C.4.5.10 Equivocal results for the guideline are not considered in Table C.4.2, partly for brevity but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative or whether the result must be put to one side and the test repeated (using the same or a different test guideline). Factors which may have

interfered with the result *e.g.* composition of the diet used, environmental influences; should be considered. Apparent equivocal results may arise because of the low sensitivity of the assay for (anti-)estrogens/androgens.

Existing Data to be Considered

C.4.5.11 Existing "Mechanism" *in vitro* data are assumed to be available from ER (ER binding and ER STTA), AR (AR binding and AR STTA) and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs *et al*. (2008) and an OECD Detailed Review Paper (OECD, 2008a). These methods, however, have not yet been validated.

C.4.5.12 Existing "Effects" data refer to *in vivo* effects that may come from UT or H assays where a non-physiological animal model is used. In these cases, it should be remembered that these assays are specifically designed to be sensitive to EDs, compared to OECD TG 407. Other data such as repeat oral toxicity studies, reproduction/developmental toxicity screen tests may be available although it is unlikely that the OECD TG 407 will be performed if higher tier data are already available as the OECD TG 407 offers no advantage over these assays. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.4.5.13 When considering the results of the OECD TG 407 assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.4.5.14 The scenarios (A to R) presented in Table C.4.5 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001 . It is recognised however, that some jurisdictions may require a two-generation study. Further considerations, specific to each scenario are given in the Table.

C.4.5.15 Scenarios A to C represent positive results in the OECD TG 407 assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive OECD TG 407 result scenario is divided into the three possible outcomes given above. A positive result in the *in vitro* assays in combination with a positive OECD TG 407 assay is moderate or strong evidence for EATS-mediated activity that may or may not be supported by the *in vivo* effects data. In the absence of robust upper level data, the next step may be to conduct an upper level test. In the presence of robust *in vivo* data then there may be sufficient evidence to conclude concern

for endocrine disruption and therefore no need for further testing. Negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption. The possibility of other (non-EATS) mechanisms should also not be overlooked *e.g.* involving other receptors or endocrine axes.

C.4.5.16 Scenarios D to F represent positive results in the OECD TG 407 assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive OECD TG 407 result scenario is divided into the three possible outcomes given above. Negative results in the *in vitro* assays should be viewed with caution in case a metabolite is responsible for the positive OECD TG 407 assay. Unless the metabolic profile of the test substance is known, the first option should be to conduct these assays with an added metabolising system. If the metabolic profile is known, then a higher level *in vivo* test may be advisable. The choice of tests will depend upon the available *in vivo* effects data. As in scenarios A to C negative existing *in vivo* effects data should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or alternatively that the effects do not present a concern for endocrine disruption.

C.4.5.17 Scenarios G to I represent positive results in the OECD TG 407 assay in the presence of various combinations of missing or equivocal data. Each positive OECD TG 407 result scenario is divided into the three possible outcomes given above. The next step to take in these eventualities will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects.. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.5.18 Scenarios J to L represent negative results in the OECD TG 407 assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. As a negative result for OECD TG 407 is taken to be negative findings for both indicators of hormonal activity and apical endpoints then (unlike the situation with positive outcomes) there is only one possible negative outcome. Negative outcomes in the OECD TG 407 should be viewed with caution because of the power of the assay to detect (anti)- estrogens and androgens. All three scenarios could also arise from a chemical that is positive in *in vitro* assays, but is metabolised to a non-active metabolite leading to negative results in the OECD TG 407 assay. This possibility should be investigated first when considering the next step. Endocrine active potency may also explain differences between *in vitro* and *in vivo* results, *e.g.* a weak chemical may give a positive result *in vitro* but may be negative *in vivo*. Positive *in vivo* effects data may involve other EATS, non-EATS mechanisms (*e.g.* involving other receptors or endocrine axes), more sensitive endpoints, greater statistical power or life stages that are more sensitive to the substance than the young adult animals in OECD TG 407.

C.4.5.19 Scenarios M to O represent negative results in the OECD TG 407 assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend upon the weight of evidence and may not be possible. Where there are positive *in vivo* effects data there could still be an EATS-related mechanism, the effects may be related to length of exposure, route of exposure or exposure at different life stages. Other EATS or non-EATS mechanisms may also be involved.

C.4.5.20 Scenarios P to R represent negative results in the OECD TG 407 assay in the presence of various combinations of missing or equivocal data. As with the positive result scenarios above (paragraph C.4.5.17) the next step to take in these eventualities will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.5.21 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.4.5 is meant to provide a succinct guide and may not cover all circumstances or possibilities.

Table C.4.5 OECD TG 407: Repeated Dose 28-Day Oral Toxicity Study in Rodents. Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter. These may be other repeated dose toxicity tests, UT and H assays or read across from chemical analogues.

***Note: three possible outcomes for a positive result are given:

1) Indicators of hormonal activity and apical endpoints positive

2) Indicators of hormonal activity positive and apical endpoints negative

3) Indicators of hormonal activity negative *and apical endpoints positive*

"Apical endpoints" are weights of testes, epididymides, prostate (+ seminal vesicles with coagulating glands), ovary, uterus, histopathologic changes in testes, epididymides, prostate, seminal vesicles, coagulating glands, ovary, uterus/cervix, vagina, thyroid and estrous cyclicity.

"Indicators of hormonal activity" are hormones (T3, T4 and TSH).

Scenarios	Result of OECD TG 407 (rodent 28 d) assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
A	+ ***	+	+	1) Moderate or strong (anti)-EATS activity. Increased evidence of (anti)-EATS activity. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-EATS activity. 3) Moderate or strong (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-EATS activity.	Perform assay from level 5 *e.g.* ext-1 or 2-gen assay.	If existing data are from level 5 then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
B	+	+	-	1) Moderate or strong (anti)-EATS activity. Increased evidence of (anti)-EATS activity. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay then question why differences. Consider route of exposures and possible implications of ADME

Scenarios	Result of OECD TG 407 (rodent 28 d) assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				EATS activity. 3) Moderate or strong (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-EATS activity.		characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms.
C	+	+	Eq/0	1) Moderate or strong (anti)-EATS activity. Increased evidence of (anti)-EATS activity. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-EATS activity. 3) Moderate or strong (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-EATS activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Check data on chemical analogues. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Consider route of exposure for OECD TG407 and follow-up assay. Possible implications of ADME characteristics of the chemical.
D	+	-	+	1) Moderate or strong (anti)-EATS activity. Acts via non-ER,AR,TR, S mechanism or	Perform *in vitro* ER, AR,TR, S	If existing data are from an adequate level 5 assay then there is sufficient

Scenarios	Result of OECD TG 407 (rodent 28 d) assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				requires metabolic activation for activity. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity. 3) Moderate or strong (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity.	assays with added metabolising system.	information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical. Further mechanistic studies would help determine MoA.
E	+	-	-	1) Moderate (anti)-EATS activity. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences between OECD TG407 and existing data	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR	If existing data are from an adequate level 5 assay then question why differences. Consider route of exposures and possible implications of ADME characteristics of the chemical.

Scenarios	Result of OECD TG 407 (rodent 28 d) assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via non-ER, AR,TR, S mechanism or requires metabolic activation for activity. Weak activity does not result in adverse effects. 3) Moderate (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-ER, AR,TR, S mechanism or requires metabolic activation for activity. Weak activity does not result in adverse effects.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate other mechanisms.
F	+	-	Eq/0	1) Moderate (anti)-EATS activity. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity 3) Moderate (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay..	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate other mechanisms. Check data on chemical analogues. Further mechanistic studies may help determine MoA. Upper level studies will provide hazard data.

Scenarios	Result of OECD TG 407 (rodent 28 d) assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
G	+	Eq/0	+	1) Moderate or strong (anti)-EATS activity. Increased evidence of (anti)-EATS activity. May act via ER, AR,TR, S mechanism. (metabolic activation needed) 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-EATS activity. May act via ER, AR,TR, S mechanism (metabolic activation needed). 3) Moderate or strong (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-EATS activity. May act via ER, AR,TR, S mechanism. (metabolic activation needed)	Perform *in vitro* ER, AR, TR, S assays. (for the "0" scenario, otherwise Eq result available) OR Perform *in vitro* ER, AR, TR, S assays with added metabolising system.	If existing data are from level 5 then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Check data on chemical analogues. Further mechanistic studies would help determine MoA. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
H	+	Eq/0	-	1) Moderate (anti)-EATS activity. Acts via unknown mechanism or requires metabolic activation for activity.	For the "0" scenario, perform *in vitro* ER, AR,	If existing data are from an adequate level 5 assay then question why differences.

Scenarios	Result of OECD TG 407 (rodent 28 d) assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				Route of exposure may account for the differences between OECD TG407 and existing data 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism Weak activity does not result in adverse effects. 3) Moderate (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via unknown mechanism. Weak activity does not result in adverse effects.	TR, S assays with added metabolising system. (otherwise Eq result available)	Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of subtle changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	1) Moderate or strong (anti)-EATS activity. Acts via unknown mechanism. Unknown potential for adverse effects. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. 3) Moderate or strong (anti)-EATS activity, indicators of hormonal activity may be less	Perform *in vitro* ER, AR, TR, S assays. Perform *in vitro* ER, AR, TR, S assays with added metabolising system	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Check data on chemical analogues. Further mechanistic studies would

Scenarios	Result of OECD TG 407 (rodent 28 d) assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. There may be a need for metabolic activation.	OR perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	help determine MoA.
J	-	+	+	No evidence for moderate or strong (anti)-EATS activity in OECD TG407. Weak (anti)-EAS activity not detected by this assay. Metabolism or potency explains the difference from existing *in vitro*/ and *in vivo* data.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay then question why differences. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical. Further mechanistic studies would help determine MoA.
K	-	+	-	No evidence for moderate or strong (anti)-EATS activity in OECD TG407. Weak (anti)-EAS activity not detected by this assay. Metabolism or potency explains *in vitro*/*in vivo* differences.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level

Scenarios	Result of OECD TG 407 (rodent 28 d) assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					*e.g.*ext-1 or 2-gen assay.	test may be required. Further mechanistic studies with metabolism may help determine MoA.
L	-	+	Eq/0	No evidence for moderate or strong (anti)-EATS activity in OECD TG407. Weak (anti)-EAS activity not detected by this assay. Metabolism or potency explains *in vitro*/*in vivo* differences. Unknown potential for adverse effects.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* activity is not realised. Consider possible routes of exposure implications of metabolism.
M	-	-	+	No evidence for moderate or strong (anti)-EATS activity in OECD TG407Weak (anti)-EAS activity not detected by this assay.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay then question why differences. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical.
N	-	-	-	No evidence for moderate or strong (anti)-EATS activity in OECD TG407. Weak (anti)-	Perform assay from level 5	If existing data are from an adequate level 5 assay there may be sufficient

Scenarios	Result of OECD TG 407 (rodent 28 d) assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				EAS activity not detected by this assay. No evidence for (anti)-EATS activity *in vitro*. No evidence of adverse effects.	*e.g.*ext-1 or 2-gen assay.	information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information).
O	-	-	Eq/0	No evidence for moderate or strong (anti)-EATS activity in OECD TG407. Weak (anti)-EAS activity not detected by this assay. No evidence for (anti)-EATS activity *in vitro*. Unknown potential for adverse effects.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Consider route of exposures and possible implications for ADME characteristics of the chemical in follow-up assay.
P	-	Eq/0	+	No evidence for moderate or strong (anti)-EATS activity in OECD TG407. Weak (anti)-EAS activity not detected by this assay. Potential for adverse effects via unknown mechanism.	Perform *in vitro* ER, AR,TR, S assays	Consider route of exposure for OECD TG 407 assay and possible implications for differences from existing assay. Effects seen in existing studies may be in a more sensitive life stage.
Q	-	Eq/0	-	No evidence for moderate or strong (anti)-EATS activity in OECD TG407. Weak (anti)-EAS activity not detected by this assay. No evidence of adverse effects.	Perform *in vitro* ER, AR,TR, S assays	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Further mechanistic studies would increase evidence.

Scenarios	Result of OECD TG 407 (rodent 28 d) assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
R	-	Eq/0	Eq/0	No evidence for moderate or strong (anti)-EATS activity in OECD TG407. Weak (anti)-EAS activity not detected by this assay.	Perform *in vitro* ER, AR,TR, S assays, otherwise Eq result available.	Further mechanistic studies would increase evidence. Check data on chemical analogues.

C.4.6 OECD TG 416: Two-Generation Reproduction Toxicity Study (including guidance on OECD TG 415: One-Generation Reproduction Toxicity Study)

C.4.6.1 Modalities detected: (Anti)-Estrogen, (anti)-androgen, thyroid, steroidogenesis.

Endpoints: (for OECD TG 416, note that endpoints for OECD TG 415 are not as extensive)

Time to mating, male fertility, female fertility, gestation length, number of implantations & corpora lutea, number of live births and post implantation loss, litter size, sex ratio (F1, F2), litter/pup weight, pup survival index

Estrus cyclicity (P, F1), sexual maturation (age at VO and PPS (F1)), AGD (F2, if triggered by changes in sex ratio or sexual maturation in F1), pup development (F1, F2).

Weights of: (P, F1) uterus, ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating glands) thyroid, adrenals.

Histopathologic changes in vagina, uterus (+ cervix), ovaries, testis, epididymis, prostate, seminal vesicles and coagulating glands.

Sperm numbers (testicular homog resistant spermatids & cauda epididymides sperm reserve), sperm motility, sperm morphology (P, F1).

Background to the Assay

C.4.6.2 The OECD two-generation and one-generation reproduction toxicity studies are apical assays designed to provide general information concerning the effects of a chemical on the male and female reproductive systems including gonadal function, the estrus cycle, mating, conception, gestation, parturition, lactation, weaning and growth and development of the offspring. The studies are not specifically designed to detect EDs but they have many endpoints relevant for the assessment of possible endocrine disruption and provide data on adverse effects related to reproduction and development. The one-generation study (OECD TG 415), adopted in 1983, only includes one cycle of mating and is much less prescriptive in both the performance of the study and the endpoints to be assessed. It has therefore been placed at Level 4 when the CF was revised in 2011 (Annex 1). In contrast the two-generation study (OECD TG 416) includes two cycles of mating and the original OECD TG was revised in 2001 to include a more comprehensive range of endpoints. These endpoints include sexual maturation (VO and PPS) which are particularly sensitive to EDs. One generation studies and two-generation studies conducted prior to the adoption of the revised OECD TG 416 are therefore unlikely to provide as much data as studies conducted to the revised OECD TG 416, particularly with respect to endocrine disruption. They do however provide a great deal of useful data, particularly on adverse effects on reproduction.

C.4.6.3 All versions of the TGs require that parental males be dosed for a period of time encompassing at least one spermatogenic cycle and that parental females be dosed for at least several estrus cycles. Dosing is continuous during mating and throughout production of subsequent generations. The exposure of the fetus (which is a sensitive life-stage for endocrine disruption effects), the long duration of dosing and the diversity of endpoints means that the revised OECD TG 416 may be considered to be more predictive for ED-mediated adverse effects via EATS modalities. As all the endpoints are apical, it is difficult to discern mechanism of action from this test alone. Information on mechanism of action needs to be obtained from *in vitro* EATS assays or *in vivo* lower tier tests such as UT and H assays.

C.4.6.4 Although formal validation of OECD TG 416 with EDs has not taken place, studies have been published showing that ER agonists [such as ethinylestradiol (NTP, 2010)], AR antagonists [such as vinclozolin (Matsuura *et al*, 2005)], steroidogenesis inhibitors [such as myclobutanil (Rockett *et al*, 2006)] and thyroid hormone modulators [such as propylthiouracil (Axelstad *et al*, 2008)] can all be detected by reproductive toxicity tests. Endocrine modalities other than EATS are also detected *e.g.* chemicals acting on the HPG axis or other hormone systems. Some chemicals interacting weakly with endocrine disrupting modalities in lower tier tests, designed to have greater sensitivity than specificity, may not have effects in this test as functional HPG axes in parents and offspring may allow compensation for weak effects. In these cases it could be interpreted that the weak effects do not lead to adverse outcomes in more comprehensive studies. Nonylphenol, for example, is a weak ER agonist in *in vitro* ER assays and in the *in vivo* UT assay, but has no effect on reproduction or development in reproductive tests (Tyl *et al*, 2006) although there were some effects on the offspring (slight changes in the oestrous cycle length, the timing of vaginal opening and possibly also in ovarian weight and sperm/spermatid count, although functional changes in reproduction were not induced at the dose levels tested). The observed perturbations in offspring were concluded (ECBI/48/99 HSE, UK) to be compatible with the predicable or hypothesised effects of exogenous oestrogenic activity. Octylphenol is a further example of a weak ER agonist in *in vitro* ER binding assays and in the UT assay, but did not reveal effects on reproduction or development in a good quality test conducted according to OECD TG 416 (Tyl et al., 1999).

The adequacy of the protocol in these studies where no endocrine-related effects are reported needs to be confirmed so that the absence of effects is not due to inadequacy of methods or reporting.

C.4.6.5 If the adequacy of the protocol is suspect, or the test was conducted before OECD TG 416 was revised, it may be possible to conduct or to use additional studies to support the reproductive toxicity test. For example, a one-generation reproduction toxicity study (OECD TG 415) not including data on sexual maturation could be supplemented by male and female peripubertal assays.

When/Why the Assay May be Used

C.4.6.6 This assay forms part of the package of studies required for registration of pesticides in many jurisdictions. It forms part of the standard information requirements in certain chemical legislations (*e.g.* REACH for chemicals which are manufactured or imported in quantities of 1000 tonnes or more). It may also be carried out for HPV chemicals of high concern, as well as being a more comprehensive test at level 5 of the CF. It is likely to have at least three dose levels and therefore may be used for both hazard and risk assessment.

Introduction to the Table of Scenarios

C.4.6.7 Table C.4.6 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.4.6.8 The results of OECD TG 416 are given in the second column. As this assay is not a screening test where a yes/no (qualitative) answer is obtained, criteria for positive results for the endocrine endpoints are not given in the test guideline. Results for the endpoints would be considered both individually and as a whole. It is not possible to provide guidance on all endpoints individually and for this test all endpoints are considered to be "apical". Serum hormone determinations are not included in OECD TG 416, therefore (unlike with the male and female PP assays and OECD TG 407) the division of the endpoints into "apical" and "indicators of hormonal activity" has not been possible

C.4.6.9 For the purpose of this guidance, a positive result is defined as a biologically significant change in any of the endocrine endpoints, *e.g.* statistically significant reductions in reproductive organ weight. Changes in related endpoints will increase their biological significance, *e.g.* abnormal estrous cyclicity combined with reduced fertility.

C.4.6.10 A negative result for the OECD TG 416 is taken to be the absence of biologically significant changes in all of the endocrine endpoints.. Studies conducted to current standards are considered to be more predictive for absence of reproductive and developmental effects.

C.4.6.11 Equivocal results for the guideline are not considered in the table, partly for brevity but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative or whether the result must be put to one side and the test repeated or supplemented by a different test. Factors which may have interfered with the result *e.g.* composition of the diet used, environmental influences; should be considered.

Existing Data to be Considered

C.4.6.12 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs *et al.* (2008) and an OECD Detailed Review Paper (OECD, 2008a). These methods, however, have not yet been validated.

C.4.6.13 Existing "Effects" data refer to *in vivo* effects that may come from lower level assays *e.g.* UT or H assays (level 3); PP assays or OECD TG 407 assays (level 4), or there may be longer term studies *e.g.* in the case of pesticide r*egis*tration packages where 90-day and carcinogenicity studies may be available. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.4.6.14 When considering the results of the OECD TG 416, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.4.6.15 The scenarios (A to R) presented in Table C.4.6 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. Further considerations, specific to each scenario are given in the Table.

C.4.6.16 Scenarios A to C represent positive results in OECD TG 416 in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. A positive result in the *in vitro* assays in combination with a positive OECD TG 416 assay is strong evidence of adverse effects on reproduction/development and/or endocrine organs via EATS mechanisms. Differential effects on the different endpoints may assist with interpretation. In all scenarios a robust OECD TG 416 study should provide sufficient information to conclude evidence of concern for reproductive toxicity via an endocrine disruption mechanism. If the study is not considered to be robust (for reasons given in paragraphs 4.6.2-4.6.5) then supplemental testing could be considered.

C.4.6.17 Scenarios D to F represent positive results in OECD TG 416 in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data.. A positive result in OECD TG 416 is strong evidence of adverse effects on reproduction/development and/or endocrine organs. Differential effects on the different endpoints may assist with interpretation. In all scenarios a robust OECD TG 416 study should provide sufficient information to conclude evidence of concern for reproductive toxicity via an endocrine disruption mechanism. If the study is not considered to be robust (for reasons given above) then supplemental testing could be considered. Negative results in the *in vitro* assays should be viewed with caution in case a metabolite is responsible for the positive OECD TG 416 study. If the metabolic profile of the test substance is not known then performing the *in vitro* assays with addition of a metabolising system may help to understand mechanism.

C.4.6.18 Scenarios G to I represent positive results in OECD TG 416 in the presence of various combinations of missing or equivocal data. The next step to take in these eventualities will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects.. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.6.19 Scenarios J to L represent negative results in OECD TG 416 in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. In all scenarios a robust OECD TG 416 study may provide sufficient information to conclude absence of concern for reproductive toxicity via an endocrine disruption mechanism. If the study is not considered to be robust (for reasons given in paragraphs 4.6.2-4.6.5) then supplemental testing could be considered. All three scenarios could fit a chemical that is positive in *in vitro* assays but is metabolised to a non-active metabolite, leading to negative results in OECD TG 416. This possibility may be investigated to help understand mechanism. Endocrine active potency may also explain differences between *in vitro* and *in vivo* results, *e.g.* a weak chemical may give a positive result *in vitro* but may be negative *in vivo*. Positive *in vivo* effects data may involve EATS or non-EATS mechanisms (*e.g.* involving other receptors or endocrine axes), more sensitive endpoints, greater statistical power but knowledge of ADME may help to explain differences from the OECD TG 416 data.

C.4.6.20 Scenarios M to O represent negative results in OECD TG 416 in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may provide sufficient information to conclude absence of concern for reproductive toxicity via an endocrine disruption mechanism. If the study is not considered to be robust (for reasons given above) then supplemental testing could be considered. Positive *in vivo* effects data may involve EATS or non-EATS mechanisms (*e.g.* involving other receptors or endocrine axes) but knowledge of ADME may help to explain differences from the OECD TG 416 data.

C.4.6.21 Scenarios P to R represent negative results in OECD TG 416 in the presence of various combinations of missing or equivocal data. As with the positive result scenarios above (paragraph

C.4.6.18) the next step to take in these eventualities will have to be decided on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.6.22 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.4.6 is meant to provide a succinct guide and may not cover all circumstances or possibilities.

Table C.4.6 OECD TG 416: Two-Generation Reproduction Toxicity Study (including guidance on OECD TG 415: One-Generation Reproduction Toxicity Study). Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter. These may be other repeated dose toxicity tests, UT and H assays, peripubertal assays or read across from chemical analogues.

***Note: a positive result is defined as a biologically significant change in any of the endocrine endpoints (all "apical endpoints").

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+ ***	+	+	Evidence for adverse effects via (anti)-EATS activity in TG 416.		Sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Effects on apical endpoints may

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether EATS mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
B	+	+	-	Evidence for adverse effects via (anti)-EATS activity in TG 416.		Sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Effects on apical endpoints may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether EATS mechanism is credible for

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
C	+	+	Eq/0	Evidence for adverse effects via (anti)-EATS activity in TG 416.		Sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption.Effects on apical endpoints may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether EATS mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
D	+	-	+	Evidence for adverse effects in TG 416 s but not via EATS mechanism or requires metabolic activation for activity.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Sufficient information to conclude evidence of concern for reproductive toxicity via possible endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Effects on apical endpoints may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
E	+	-	-	Evidence for adverse effects in TG 416 via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity.	To further discern mechanism could	Sufficient information to conclude evidence of concern for reproductive toxicity via unknown mechanism.

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Note that the ext-1 gen assay provides the most information on endocrine disruption.Effects on apical endpoints may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
F	+	-	Eq/0	Evidence for adverse effects in TG 416 via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. .	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Sufficient information to conclude evidence of concern for reproductive toxicity via unknown mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Effects on apical endpoints may indicate EATS modalities or other mechanisms. Consider potency of effects for

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
G	+	Eq/0	+	Evidence for adverse effects in TG 416, may act via EATS mechanism and may require metabolic activation for activity.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Sufficient information to conclude evidence of concern for reproductive toxicity via possible endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Effects on apical endpoints may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
H	+	Eq/0	-	Evidence for adverse effects in TG 416 via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Sufficient information to conclude evidence of concern for reproductive toxicity via unknown mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Effects on apical endpoints may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non- endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
I	+	Eq/0	Eq/0	Evidence for adverse effects in TG 416 via	To further discern	Sufficient information to

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				unknown mechanism.	mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	conclude evidence of concern for reproductive toxicity via unknown mechanism. Consider existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
J	-	+	+	No evidence of adverse effects on reproduction/development/endocrine organs. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay. Metabolism or potency explains the difference from existing *in vitro*/ and *in vivo* data.	If test is to current OECD TG 416 standards, no further testing needed. If not then consider supplemental testing, depending upon existing data. To further discern	If existing data are from other, adequate, apical studies, than question why differences. Note that the ext-1 gen assay provides the most information on endocrine disruption.Consider route of exposures and possible implications for ADME characteristics of the chemical with existing studies. Further mechanistic studies with metabolism may help determine

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	MoA.
K	-	+	-	No evidence of adverse effects on reproduction/development/endocrine organs. Metabolism or potency explains *in vitro*/*in vivo* differences.	If test is to current OECD TG 416 standards, no further testing needed. If not then consider supplemental testing, depending upon existing data. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption. Further mechanistic studies with metabolism may help determine MoA.
L	-	+	Eq/0	No evidence of adverse effects on reproduction/development/endocrine	If test is to current OECD TG 416	There may be sufficient information to conclude absence

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				organs. Metabolism or potency explains *in vitro*/*in vivo* differences. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay.	standards, no further testing needed. If not then consider supplemental testing, depending upon existing data. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	of concern for endocrine disruption. Note that the ext-1 gen assay provides the most information on endocrine disruption.Consider route of exposures and possible implications for ADME characteristics of the chemical with existing studies. Further mechanistic studies with metabolism may help determine MoA.
M	-	-	+	No evidence of adverse effects on reproduction/development/endocrine organs. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay.	If test is to current OECD TG 416 standards, no further testing needed. If not then consider	If existing data are from adequate in vivo studies such as 28d, 90d, chronic/carcinogenicity studies, than question why differences. Note that the ext-1 gen assay provides the most information on endocrine

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					supplemental testing, depending upon existing data.	disruption.Consider route of exposures and possible implications for ADME characteristics of the chemical with existing studies. Further mechanistic studies with metabolism may help determine MoA.
N	-	-	-	No evidence of adverse effects on reproduction/development/endocrine organs.	If test is to current OECD TG 416 standards, no further testing needed. If not then consider supplemental testing, depending upon existing data.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption. Note that the ext-1 gen assay provides the most information on endocrine disruption.
O	-	-	Eq/0	No evidence of adverse effects on reproduction/development/endocrine organs. No evidence for (anti)-EATS activity *in vitro*.	If test is to current OECD TG 416 standards, no further testing needed. If not then	There may be sufficient information to conclude absence of concern for endocrine disruption. Note that the ext-1 gen assay provides the most information on endocrine disruption.Further

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					consider supplemental testing, depending upon existing data. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	mechanistic studies with metabolism may help determine MoA. Consider route of exposures and possible implications for ADME characteristics of the chemical with existing studies. Check data on chemical analogues.
P	-	Eq/0	+	No evidence of adverse effects on reproduction/development/endocrine organs. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay. Effects seen in existing studies are via unknown mechanism.	If test is to current OECD TG 416 standards, no further testing needed. If not then consider supplemental testing, depending upon existing data.	If existing data are from adequate in vivo studies such as 28d, 90d, chronic/carcinogenicity studies, than question why differences. Note that the ext-1 gen assay provides the most information on endocrine disruption.Consider route of exposures and possible implications for ADME characteristics of the chemical with existing studies.

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Check data on chemical analogues.
Q	-	Eq/0	-	No evidence of adverse effects on reproduction/development/endocrine organs.	If test is to current OECD TG 416 standards, no further testing needed. If not then consider supplemental testing, depending upon existing data.	There may be sufficient information to conclude absence of concern for endocrine disruption. Check data on chemical analogues.
R	-	Eq/0	Eq/0	No evidence of adverse effects on reproduction/development/endocrine organs.	If test is to current OECD TG 416 standards, no further testing needed. If not then consider supplemental	There may be sufficient information to conclude absence of concern for endocrine disruption. Note that the ext-1 gen assay provides the most information on endocrine disruption.Further mechanistic studies would increase evidence.

Scenarios	Result of OECD TG 416 (2 gen study)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					testing, depending upon existing data.	Check data on chemical analogues.

C.4.7 OECD TG 443: Extended One-Generation Reproductive Toxicity Study

C.4.7.1 Modalities detected: (Anti)-Estrogen, (anti)-androgen, thyroid, steroidogenesis.

Endpoints: Time to mating, male fertility, female fertility, dystocia, gestation length, number of implantations & corpora lutea, number of ovarian follicles, number of live births and post implantation loss, viability index, litter size, sex ratio, litter/pup weight, pup survival index, placental weight

AGD, presence of nipples, pup development including genitals (and presence of abnormalities), sexual maturation (age at VO and PPS), (F1).

Weights and/or histopathologic analysis: uterus (with oviducts and cervix), ovaries, testes, epididymides, prostate, seminal vesicles (+ coagulating glands) thyroid, adrenals, pituitary, mammary gland (P& F1).

Estrus cyclicity (P& F1)

Sperm numbers (testicular homog resistant spermatids & cauda epididymides sperm reserve), sperm motility, sperm morphology (P& F1).

Hormones: T4, TSH (P& F1)

Apical endpoints from the developmental neuro- and immunotoxicity cohorts may be sensitive to endocrine modulation.

Background to the Assay

C.4.7.2 The extended one-generation reproduction toxicity study is an apical assay designed to evaluate specific life stages not covered by other tests and to test for effects that may occur as a result of pre- and post-natal exposure to chemicals. It is based on the proposal of Cooper *et al* (2006) and includes three possible cohorts of F1 animals:

1. To assess reproductive/developmental endpoints;
2. To assess effects on the developing nervous system;
3. To assess effects on the developing immune system.

The reproductive/developmental element of the study provides a thorough evaluation of systemic, reproductive and developmental toxicity including gonadal function, the estrus cycle, epididymal sperm maturation, mating, conception, gestation, parturition, lactation, weaning and growth and development of the offspring. Depending on the modules carried out in the test, effects on the developing nervous and immune systems are also assessed. These systems may also be sensitive to endocrine influences. The study uses fewer animals than OECD TG 416 (Two-generation Reproduction Toxicity Study), whilst increasing the number of pups studied in the F1 generation and the number of endpoints. Inclusion of an F2 generation is "triggered" depending upon results obtained in the F1 generation. Decisions on whether to assess the second generation, omit the developmental neurotoxicity or developmental immunotoxicity have to be taken on a case-by-case basis depending upon existing knowledge and regulatory purpose. The procedure and internal triggers for deciding whether a second generation should be produced are described in OECD GD 117 (2011c). As the second generation is "triggered", then at present the OECD TG 416 is the only OECD mammalian test that automatically covers two generations. OECD GD 151 (2010f) also supports OECD TG 443, providing advice on study design including the gathering of key data on the substance to be tested, endpoints and data interpretation issues not adequately covered in the TG.

C.4.7.3 The extended one-generation study is not specifically designed to detect EDs but Cohort 1 has many endpoints relevant for the assessment of possible ED, for example endpoints such as sexual maturation and estrous cyclicity are particularly sensitive to estrogens and androgens. Effects on the thyroid and thyroid hormones are also detected by serum T4 and TSH levels, thyroid weight and by histopathology in P and F1 generations. The assay also provides data on adverse effects related to reproduction and development which may or may not be related to ED. Cohorts 2 and 3 also have apical endpoints that may be sensitive to endocrine modulation. The developing brain, for example, is a classical target of thyroid hormones whilst interaction of chemicals with the hypothalamic-pituitary-adrenal axis may affect both the developing immune and nervous systems.

Dosing is continuous, prior to and during mating, and throughout production of the subsequent generation(s). The exposure of the fetus (which is a sensitive life-stage for endocrine disruption effects), the long duration of dosing and the diversity of endpoints means that the extended one-generation study (in addition to the revised OECD TG 416) may be considered to be the most predictive test for ED-mediated adverse effects via EATS modalities. As all the endpoints are apical, it is difficult to discern mechanism of action from this test alone. Information on mechanism of action needs to be obtained from *in vitro* EATS assays or *in vivo* lower tier tests such as UT and H assays.

When/Why the Assay May be Used

C.4.7.4 The extended one-generation study is likely to replace OECD TG 416 over a period of time. As an alternative to OECD TG 416, it may form part of the package of studies required for registration of pesticides in many jurisdictions. It may also be used as an alternative to OECD TG 416 for part of the standard information requirements in certain chemical legislations . It may also be carried out for HPV chemicals of high concern as well as being the most comprehensive test at level 5 of the CF. It is likely to have at least three dose levels and therefore may be used for both hazard and risk assessment.

Introduction to the Table of Scenarios

C.4.7.5 Table C.4.7 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. “Existing results” are subdivided into “Mechanism” and “Effects” data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

C.4.7.6 The results of the extended one-generation study are given in the second column. As this assay is not a screening test where a yes/no (qualitative) answer is obtained, criteria for positive results for the endocrine endpoints are not given in the test guideline. Results for the endpoints would be considered both individually and as a whole. It is not possible to provide guidance on all endpoints individually and therefore the endpoints have been pragmatically divided into “apical” and “indicators of hormonal activity”. The terminology used has been chosen to be consistent between both the wildlife and mammalian tests. Both groups have similar biological importance, although the “indicators of hormonal activity” in the mammalian assays are serum hormones and are generally, but not always, more variable than “apical endpoints”.

For this guideline “Apical” endpoints are reproductive and developmental parameters (including AGD, presence of nipples, genital abnormalities), sexual maturation, sperm parameters, estrous cyclicity, weights and histopathologic changes in testes, epididymides, prostate, seminal vesicles (with coagulating glands), ovary, uterus (with oviducts and cervix), thyroid. “Indicators of hormonal activity” are hormones (T4, TSH).

Three possible outcomes for a positive result are therefore envisaged in C.4.7:

1) Indicators of hormonal activity and apical endpoints positive
2) Indicators of hormonal activity positive and apical endpoints negative
3) Indicators of hormonal activity negative and apical endpoints positive

C.4.7.7 A positive result for apical endpoints could be statistically significant changes in pup AGD, accompanied by treatment-related histopathologic changes in parental reproductive organs or decreased fertility. A positive result for indicators of hormonal activity could be statistically significant changes in hormone profiles. A positive result for indicators of hormonal activity alone should be considered with caution although it is possible that these endpoints may have detected weak effects that were not detected by the apical endpoints.

C.4.7.8 A negative result for the extended one-generation study is taken to be absence of changes in both endocrine relevant indicators of hormonal activity and apical endpoints. A well conducted study is considered to be more predictive for absence of reproductive and developmental effects and for endocrine disruptive effects mediated through EATS modalities..

C.4.7.9 Equivocal results for the guideline are not considered in the table, partly for brevity but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative or whether the result must be put to one side and the test repeated or supplemented by a different test. Factors which may have interfered with the result *e.g.* composition of the diet used, environmental influences; should be considered.

Existing Data to be Considered

C.4.7.10 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs *et al.* (2008) and an OECD Detailed Review Paper (OECD, 2008a). These methods, however, have not yet been validated.

C.4.7.11 Existing "Effects" data refer to *in vivo* effects that may come from lower level assays *e.g.* UT or H assays (level 3); PP assays or OECD TG 407 assays (level 4), or there may be longer term studies *e.g.* in the case of pesticide registration packages where 90-day and carcinogenicity studies may be available. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

C.4.7.12 When considering the results of the H assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

C.4.7.13 A series of scenarios (A to R) are presented in Table C.4.7 and represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are

generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage. Further considerations, specific to each scenario are given in the Table.

C.4.7.14 Scenarios A to C represent positive results in the extended one-generation study in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive extended one-generation study result scenario is divided into the three possible outcomes given above. A positive result in the *in vitro* assays in combination with a positive extended one-generation study is strong evidence of adverse effects on reproduction/development and/or endocrine organs via EATS mechanisms. Differential effects on the apical endpoints or indicators of hormonal activity may assist with interpretation. In all scenarios a robust extended one-generation study should provide sufficient information to conclude evidence of concern for reproductive toxicity via an endocrine disruption mechanism. If the study is not considered to be robust, then supplemental testing could be considered.

C.4.7.15 Scenarios D to F represent positive results in the extended one-generation study in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal in vivo effects data. Each positive result scenario is divided into the three possible outcomes given above. A positive result in the extended one-generation study is strong evidence of adverse effects on reproduction/development and/or endocrine organs. Differential effects on the different apical endpoints or indicators of hormonal activity may assist with interpretation. In all scenarios a robust extended one-generation study should provide sufficient information to conclude evidence of concern for reproductive toxicity via an endocrine disruption mechanism. If the study is not considered to be robust then supplemental testing could be considered. Negative results in the *in vitro* assays should be viewed with caution in case a metabolite is responsible for the positive extended one-generation study. If the metabolic profile of the test substance is not known then performing the *in vitro* assays with addition of a metabolising system may help to understand mechanism.

C.4.7.16 Scenarios G to I represent positive results in the extended one-generation study in the presence of various combinations of missing or equivocal data. Each positive result scenario is divided into the three possible outcomes given above. The next step to take in these eventualities will depend upon the nature of the other available data and the jurisdiction in which it is being used. In some cases equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action in question and why the data are considered equivocal, *e.g.* a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.7.17 Scenarios J to L represent negative results in the extended one-generation study in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. As a negative result is taken to be negative findings for both indicators of hormonal activity and apical endpoints then (unlike the situation with positive outcomes) there is only one possible negative outcome. In all scenarios a robust extended one-generation study may provide sufficient information to conclude absence of concern for reproductive toxicity via an endocrine disruption mechanism. If the study is not considered to be robust then supplemental testing could be considered. All three scenarios could fit a chemical that is positive in *in vitro* assays but is metabolised to a non-active metabolite, leading to negative results in the extended one-generation study. This possibility may be investigated to help understand mechanism. Endocrine active potency may also explain differences

between *in vitro* and *in vivo* results, *e.g.* a weak chemical may give a positive result in vitro but may be negative *in vivo*. Positive *in vivo* effects data may involve EATS or non-EATS mechanisms (*e.g.* involving other receptors or endocrine axes), more sensitive endpoints, or greater statistical power but knowledge of ADME may help to explain differences from the extended one-generation study data.

C.4.7.18 Scenarios M to O represent negative results in the extended one-generation study in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may provide sufficient information to conclude absence of concern for reproductive toxicity via an endocrine disruption mechanism. If the study is not considered to be robust then supplemental testing could be considered. Positive *in vivo* effects data may involve EATS or non-EATS mechanisms (*e.g.* involving other receptors or endocrine axes) but knowledge of ADME may help to explain differences from the extended one-generation study data.

C.4.7.19 Scenarios P to R represent negative results in the extended one-generation study in the presence of various combinations of missing or equivocal data. As with the positive result scenarios above (paragraph C.4.7.16) the next step to take in these eventualities will have to be decided on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

C.4.7.20 In all scenarios (A to R), the next step to take to increase evidence will depend upon the existing information. Table C.4.7 is meant to provide a succinct guide and may not cover all circumstances or possibilities.

Table C.4.7. OECD TG 443: Extended One-Generation Reproductive Toxicity Study. Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter. These may be other repeated dose toxicity tests, UT and H assays, peripubertal assays or read across from chemical analogues.

***Note: three possible outcomes for a positive result are given:

1) Indicators of hormonal activity and apical endpoints positive

2) Indicators of hormonal activity positive and apical endpoints negative

3) Indicators of hormonal activity negative *and apical endpoints positive*

"Apical endpoints" are reproductive and developmental parameters (including AGD, presence of nipples, genital abnormalities), sexual maturation, sperm parameters, estrous cyclicity, weights and histopathologic changes in testes, epididymides, prostate, seminal vesicles (with coagulating glands), ovary, uterus (with oviducts and cervix), thyroid. Apical endpoints from the developmental neuro- and immunotoxicity cohorts may also be sensitive to endocrine modulation.

"Indicators of hormonal activity" are hormones (T4, TSH).

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+ ***	+	+	1) Evidence for adverse effects via (anti)-EATS activity in TG 443. 2) Evidence for adverse effects via (anti)-EATS activity in TG 443. Apical endpoints may be less sensitive or unaffected. 3) Evidence for adverse effects via (anti)-EATS activity in TG 443. Indicators of hormonal activity may be less sensitive or unaffected.		Sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether EATS mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
B	+	+	-	1) Evidence for adverse effects via (anti)-EATS activity in TG 443.		Sufficient information to conclude evidence of concern for

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				2) Evidence for adverse effects via (anti)-EATS activity in TG 443. Apical endpoints may be less sensitive or unaffected. 3) Evidence for adverse effects via (anti)-EATS activity in TG 443. Indicators of hormonal activity may be less sensitive or unaffected.		reproductive toxicity via endocrine disruption mechanism. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether EATS mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
C	+	+	Eq/0	1) Evidence for adverse effects via (anti)-EATS activity in TG 443. 2) Evidence for adverse effects via (anti)-EATS activity in TG 443. Apical		Sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism.

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				endpoints may be less sensitive or unaffected. 3) Evidence for adverse effects via (anti)-EATS activity in TG 443. Indicators of hormonal activity may be less sensitive or unaffected.		Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether EATS mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
D	+	-	+	1) Evidence for adverse effects in TG 443 but not via EATS mechanism or requires metabolic activation for activity. 2) Evidence for adverse effects in TG 443 but not via EATS mechanism or requires metabolic activation for activity. Apical endpoints may be less sensitive or unaffected.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising	Sufficient information to conclude evidence of concern for reproductive toxicity via possible endocrine disruption mechanism. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints.

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				3) Evidence for adverse effects in TG 443but not via EATS mechanism or requires metabolic activation for activity. Indicators of hormonal activity may be less sensitive or unaffected.	system.	Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
E	+	-	-	1) Evidence for adverse effects in TG 443 via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. 2) Evidence for adverse effects in TG 443 via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. Apical endpoints may be less sensitive or unaffected.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Sufficient information to conclude evidence of concern for reproductive toxicity via unknown mechanism. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				3 Evidence for adverse effects in TG 443 via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. Indicators of hormonal activity may be less sensitive or unaffected.		may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
F	+	-	Eq/0	1) Evidence for adverse effects in TG 443via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. 2) Evidence for adverse effects in TG 443via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. Apical endpoints may be less sensitive or unaffected. 3) Evidence for adverse effects in TG 443 via non-EATS/non endocrine disruption	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Sufficient information to conclude evidence of concern for reproductive toxicity via unknown mechanism. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Consider potency of effects for

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				mechanism or requires metabolic activation for activity. Indicators of hormonal activity may be less sensitive or unaffected.		existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
G	+	Eq/0	+	1) Evidence for adverse effects in TG 443, may act via EATS mechanism and may require metabolic activation for activity. 2) Evidence for adverse effects in TG 443, may act via EATS mechanism and may require metabolic activation for activity. Apical endpoints may be less sensitive or unaffected. 3) Evidence for adverse effects in TG 443, may act via EATS mechanism and may require metabolic activation for activity. Indicators of hormonal activity may be less sensitive or unaffected.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Sufficient information to conclude evidence of concern for reproductive toxicity via possible endocrine disruption mechanism. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
H	+	Eq/0	-	1) Evidence for adverse effects in TG 443via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. 2) Evidence for adverse effects in TG 443via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. Apical endpoints may be less sensitive or unaffected. 3 Evidence for adverse effects in TG 443 via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. Indicators of hormonal activity may be less sensitive or unaffected.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Sufficient information to conclude evidence of concern for reproductive toxicity via unknown mechanism. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
I	+	Eq/0	Eq/0	1) Evidence for adverse effects in TG 443via unknown mechanism. 2) Evidence for adverse effects in TG 443 via unknown mechanism. Apical endpoints may be less sensitive or unaffected. 3) Evidence for adverse effects in TG 443 via unknown mechanism. Indicators of hormonal activity may be less sensitive or unaffected.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Sufficient information to conclude evidence of concern for reproductive toxicity via unknown mechanism. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Consider existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						characteristics of the chemical.
J	-	+	+	No evidence of adverse effects on reproduction/development/endocrine organs. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay. Metabolism or potency explains the difference from existing *in vitro*/ and *in vivo* data.	If test is robust, no further testing needed. If not then consider supplemental testing, depending upon existing data. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	There may be sufficient information to conclude absence of concern for endocrine disruption. If existing data are from other, adequate, apical studies, then question why differences. Consider route of exposures and possible implications for ADME characteristics of the chemical with existing studies. Further mechanistic studies with metabolism may help determine MoA.
K	-	+	-	No evidence of adverse effects on reproduction/development/endocrine organs. Metabolism or potency explains *in vitro*/*in vivo* differences.	If test is robust, no further testing needed.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					If not then consider supplemental testing, depending upon existing data. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	endocrine disruption. Further mechanistic studies with metabolism may help determine MoA.
L	-	+	Eq/0	No evidence of adverse effects on reproduction/development/endocrine organs. Metabolism or potency explains *in vitro*/*in vivo* differences. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay.	If test is robust no further testing needed. If not then consider supplemental	There may be sufficient information to conclude absence of concern for endocrine disruption. Consider route of exposures and possible implications for ADME characteristics of the chemical

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					testing, depending upon existing data. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	with existing studies. Further mechanistic studies with metabolism may help determine MoA.
M	-	-	+	No evidence of adverse effects on reproduction/development/endocrine organs. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay.	If test is robust no further testing needed. If not then consider supplemental testing, depending upon existing data.	There may be sufficient information to conclude absence of concern for endocrine disruption. Consider route of exposures and possible implications for ADME characteristics of the chemical with existing studies. Further mechanistic studies with metabolism may help determine MoA.

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
N	-	-	-	No evidence of adverse effects on reproduction/development/endocrine organs.	If test is robust, no further testing needed. If not then consider supplemental testing, depending upon existing data.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption.
O	-	-	Eq/0	No evidence of adverse effects on reproduction/development/endocrine organs. No evidence for (anti)-EATS activity *in vitro*.	If test is robust no further testing needed. If not then consider supplemental testing, depending upon existing data. To further discern	There may be sufficient information to conclude absence of concern for endocrine disruption. Further mechanistic studies with metabolism may help determine MoA. Consider route of exposures and possible implications for ADME characteristics of the chemical with existing studies. Check data on chemical analogues.

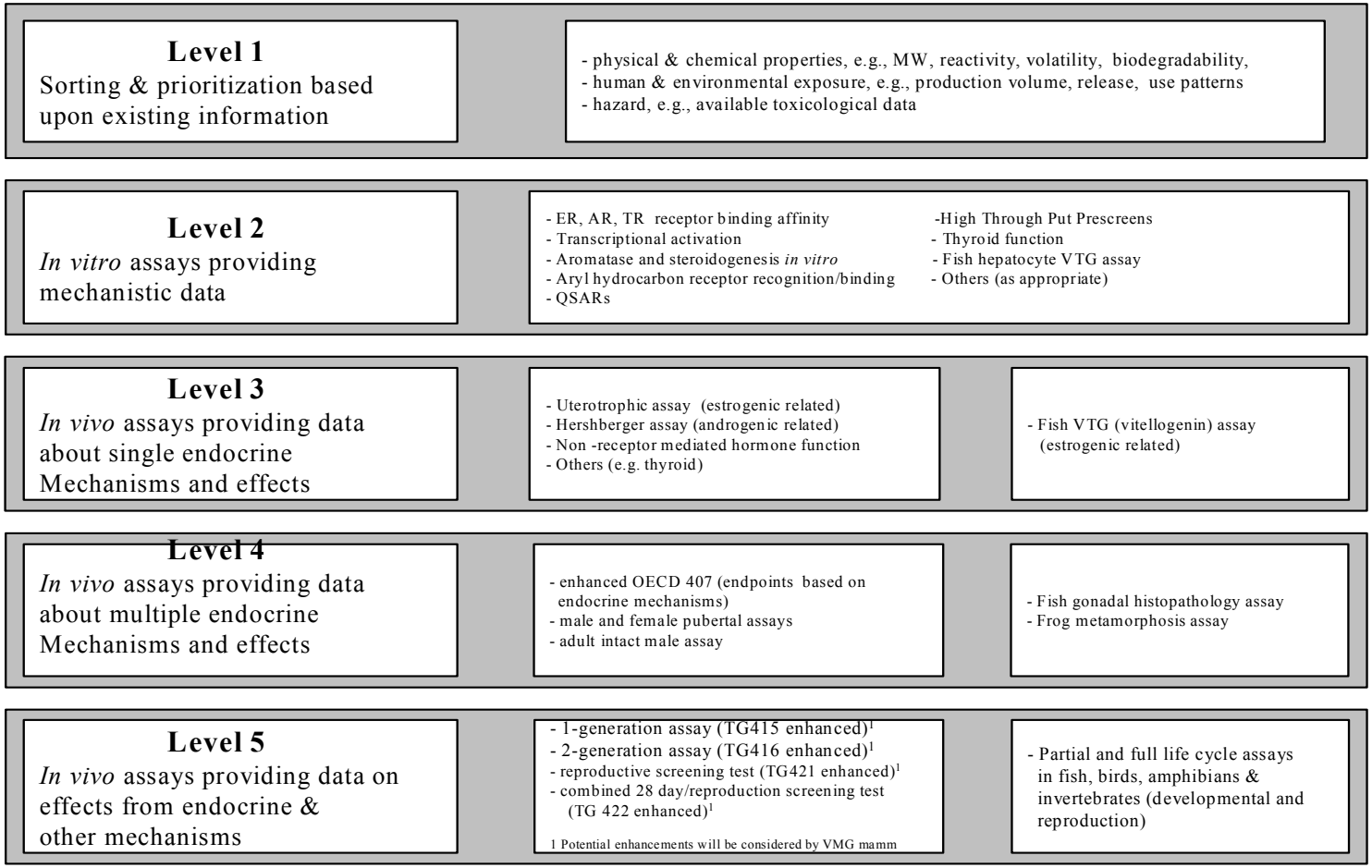
Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	
P	-	Eq/0	+	No evidence of adverse effects on reproduction/development/endocrine organs. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay. Effects seen in existing studies are via unknown mechanism.	If test is robust no further testing needed. If not then consider supplemental testing, depending upon existing data. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added	There may be sufficient information to conclude absence of concern for endocrine disruption. Further mechanistic studies with metabolism may help determine MoA. Consider route of exposures and possible implications for ADME characteristics of the chemical with existing studies. Check data on chemical analogues.

Scenarios	Result of ext 1-gen study	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					metabolising system.	
Q	-	Eq/0	-	No evidence of adverse effects on reproduction/development/endocrine organs.	If test is robust no further testing needed. If not then consider supplemental testing, depending upon existing data.	There may be sufficient information to conclude absence of concern for endocrine disruption. Check data on chemical analogues.
R	-	Eq/0	Eq/0	No evidence of adverse effects on reproduction/development/endocrine organs.	If test is robust, no further testing needed. If not then consider supplemental testing, depending upon existing data.	There may be sufficient information to conclude absence of concern for endocrine disruption. Further mechanistic studies would increase evidence. Check data on chemical analogues.

Annex 1. The OECD Conceptual Framework for Testing and Assessment of Endocrine Disrupting Chemicals with revisions.

Annex 1.1 The original OECD CF for testing and assessment of endocrine disrupting chemicals was agreed at the 6th meeting of the EDTA task force. A revised framework was discussed at the OECD workshop in Copenhagen (OECD, 2010) and finalised by the EDTA Advisory Group in 2011. Revisions were needed to include tests that were unavailable when the CF was first proposed. The original CF with notes and the revised CF are shown below.

Annex 1.2 Original OECD Conceptual Framework for Testing and Assessment of Endocrine Disrupting Chemicals (now superseded).



Annex 1.3 Notes to the Original Framework

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

Note 2: In level 5, ecotoxicology should include endpoints that indicate mechanisms of adverse effects, and potential population damage

Note 3: When a multimodal model covers several of the single endpoint assays, that model would replace the use of those single endpoint assays

Note 4: The assessment of each chemical should be based on a case by case basis, taking into account all available information, bearing in mind the function of the framework levels.

Note 5: The framework should not be considered as all inclusive at the present time. At levels 3,4 and 5 it includes assays that are either available or for which validation is under way. With respect to the latter, these are provisionally included. Once developed and validated, they will be formally added to the framework.

Note 6: Level 5 should not be considered as including definitive tests only. Tests included at that level are considered to contribute to general hazard and risk assessment.

Annex 1.4 2011 OECD Revised Conceptual Framework for Testing and Assessment of Endocrine Disrupters

The Conceptual Framework lists the OECD TGs and standardized test methods available, under development or proposed that can be used to evaluate chemicals for endocrine disruption. The Conceptual Framework is intended to provide a guide to the tests available which can provide information for endocrine disrupters assessment but is not intended to be a testing strategy. Furthermore, this Conceptual Framework does not include evaluation of exposure, however this should be included when deciding whether further testing is needed. Further information regarding the use and interpretation of these tests is available in GD 150 *(i.e.* this GD).

Mammalian and non mammalian Toxicology	
Level 1 Existing data and non-test information	• Physical & chemical properties, e.g., MW reactivity, volatility, biodegradability • All available (eco)toxicological data from standardized or non-standardized 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) (Mammalian and non mammalian methods)	• Estrogen or androgen receptor binding affinity • Estrogen receptor transactivation (OECD TG 455) • Androgen or thyroid transactivation (If/when TGs are available) • Steroidogenesis *in vitro* (OECD TG 456) • MCF-7 cell proliferation assays (ER ant/agonist) • Other assays as appropriate

	Mammalian Toxicology	**Non-Mammalian Toxicology**
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)	• Xenopus embryo thyroid signalling assay (When/if TG is available) • Amphibian metamorphosis assay (OECD TG 231) • Fish reproductive screening assay (OECD TG 229) • Fish screening assay (OECD TG 230) • Androgenized female stickleback screen (GD 140)

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) • 1-generation reproduction toxicity study (OECD TG 415) • Male pubertal assay (see GD 150 [*i.e.*this GD] Chapter C4.3)[3] • Female pubertal assay (see GD 150 [*i.e.*this GD] Chapter C4.4)[3] • Intact adult male endocrine screening assay (see GD 150 [*i.e.*this GD] Chapter Annex 2.5) • Prenatal developmental toxicity study (OECD TG 414) • Chronic toxicity and carcinogenicity studies (OECD TG 451-3) • Reproductive screening test (OECD TG 421 if enhanced) • Combined 28-day/reproductive screening assay (OECD TG 422 if enhanced) • Developmental neurotoxicity (OECD TG 426)	• Fish sexual development test (Draft OECD TG 234) • Fish reproduction Partial Lifecycle Test (when/If TG is Available) • Larval amphibian growth & development assay (when TG is available) • Avian reproduction assay (OECD TG 206) • Mollusc partial lifecycle assays (when TG is available) [4] • Chironomid toxicity test (TG 218-219) [4] • Daphnia reproduction test (with male induction) (OECD TG 211) [4] • Earthworm reproduction test (OECD TG 222, 2004) [4] • Enchytraeid reproduction test (OECD TG 220, 2004) [4] • Sediment water lumbriculus toxicity test using spiked sediment (OECD TG 225, 2007) [4] • Predatory mite reproduction test in soil (OECD TG 226, 2008) [4] • Collembolan reproduction test in soil (TG OECD 232, 2009) [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 (OECD TG 443)[5] • 2-Generation reproduction toxicity study (OECD TG 416 most recent update)	• Fish lifecycle toxicity test (FLCTT) (when TG is available) • Medaka multigeneration test (MMGT) (when TG is available) • Avian 2 generation reproductive toxicity assay (when TG is available) • Mysid lifecycle toxicity test (when TG is available)[4]

		• Copepod reproduction and development test (when TG is available)[4] • Sediment water chironomid life cycle toxicity test (OECD TG 233) [4] • Mollusc full lifecycle assays (when TG is available) [4] • Daphnia multigeneration assay (if TG is available) [4]

[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 endocrine disrupters 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. At levels 2, 3, 4 and 5 it includes assays that are either available or for which validation is under way. With respect to the latter, these are provisionally included.

Annex 2. Provisional Guidance on Assays Not Included in the Main Document.

Note that these assays have either not yet been fully validated or are not primarily designed for testing specifically for EDs, so the limited guidance offered in this annex may have to be amended in due course. See Section C.1 in main document for background to the "building blocks" here.

Endpoints which may eventually be reliably measured using these assays are listed in Table Annex 2(a).

Table Annex 2 (a). Endpoints relevant for endocrine disruption modalities in assays that have not yet received full validation for endocrine outcomes, or are test guidelines that are not primarily designed for testing specifically for EDs.

Probable direction of change is indicated where possible**.**

Note that for many assays, individual endpoints may not in themselves be diagnostic of an endocrine disruption modality. Such diagnosis often relies on a combination of endpoints or assays in a weight of evidence assessment.

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
***In vitro* screens**							
Stably Transfected Human Androgen Receptor Transactivation Assay (AR STTA) [Table Annex 2.1]	Nil	Nil	Activation of reporter gene linked to AR	Inhibition of activation of reporter gene linked to AR	Nil	Nil	Nil
Wildlife *in vivo* screens and tests							
Fish (Medaka) multi-generation test (MMGT) (draft OECD TG)	Female-biased phenotypic sex ratio VTG induction	?	Altered levels of estradiol and/or (keto) testosterone	?	Altered levels of thyroid hormones	VTG depression in females	Hatching success Weight

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
[Table Annex 2.2]	in males Altered levels of estradiol and/or (keto) testosterone						Length Behaviour Gross morphology Gonado-somatic index Multiple organ histopathology Time to maturity (time to first spawn) Fecundity Fertilisation success
Larval Amphibian Growth and Development Assay (LAGDA) (draft OECD TG) [Table Annex 2.3]	Not yet any validated endpoints with specific diagnostic properties, although some endpoints in the	Not yet any validated endpoints with specific diagnostic properties	Not yet any validated endpoints with specific diagnostic properties, although some endpoints in the right-hand column are expected to be specifically	Not yet any validated endpoints with specific diagnostic properties	Not yet any validated endpoints with specific diagnostic properties, although several endpoints in the	Not yet any validated endpoints with specific diagnostic properties	VTG T4 and TSH hormone titres Snout-vent length

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
	right-hand column are expected to be specifically responsive to estrogens		responsive to androgens		right-hand column are expected to be specifically responsive to thyroid disrupters.		Body weight Thyroid and gonad histopathology Time to metamorphosis (NF stage 62) Nuptial pad development Biased phenotypic sex ratio
Avian two generation test (ATGT) (draft OECD TG) [Table Annex 2.4]	Not yet any validated endpoints with specific diagnostic properties	Not yet any validated endpoints with specific diagnostic properties	Not yet any validated endpoints with specific diagnostic properties	Not yet any validated endpoints with specific diagnostic properties	Not yet any validated endpoints with specific diagnostic properties	Not yet any validated endpoints with specific diagnostic properties	As OECD TG 206, plus following indicators of hormonal activity: Weight of organs Histology of organs Testicular spermatid counts Gross anomalies of genital tract Feather

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
							dimorphism Cloacal gland size, 1st appearance of foam Time to first egg laid Sexual behaviour Steroid hormone titres Tibiotarsus length (F1)
Mammalian *in vivo* screens and tests							
Adult Male Assay [Table Annex 2.5]	The following changes may occur: Decreased weight of epididymides, prostate, seminal vesicles (+ coagulating gland), accessory sex	Assay is not designed to detect this modality and studies using pure antagonists are lacking. However, the following changes may occur in the following	Increased weight of epididymides, prostate, seminal vesicles (+ coagulating gland), accessory sex glands Decreased testis weight. Histopathologic changes in testes	Decreased weight of epididymides, prostate, seminal vesicles (+ coagulating gland), accessory sex glands Increased testis weight. Histopathologic changes in testes	Increased thyroid weight. Possible liver weight increase (in combination with other thyroid-related endpoints). Histopathologic changes in	Possible effects on: Weight of epididymides, prostate, seminal vesicles (+ coagulating gland), accessory sex glands Histopathologic	

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
	glands Increased testis weight. Histopathologic changes in testes epididymides. Changes in serum hormones including increased serum testosterone	endpoints: Weight of testes, epididymides, prostate, seminal vesicles (+ coagulating gland), accessory sex glands Histopathologic changes in testes epididymides. Changes in serum hormones	epididymides. Changes in serum hormones including decreased serum testosterone	epididymides. Changes in serum hormones including increased serum testosterone	thyroid. Serum T4, T3 decreased, TSH increased	changes in testes epididymides. Serum hormones	
OECD TG 408 Repeated dose 90-day oral toxicity study [Table Annex 2.6]	Increased uterus weight, decreased ovary weight. Histopathologic changes in ovary, uterus/cervix,	Studies using pure antagonists are lacking. However, changes may occur in the following: Uterus and	Decreased ovary weight. Histopathologic changes in ovary, uterus/cervix, vagina. Increased weight of epididymides,	Histopathologic changes in ovary, uterus/cervix, vagina. Decreased weight of epididymides, increased testes	Possible liver weight increase (in combination with other thyroid-related endpoints). Histopathologic changes in	Possible effects on: Uterus and ovary weight. Histopathologic changes in ovary,	Changes in adrenal weight. Histopathologic changes in adrenal, and pituitary glands.

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
	vagina and female mammary gland. Decrease in weight of epididymides, Histopathologic changes in testes, epididymides, male accessory sex organs and male mammary gland.	ovary weight.Histopathologic changes in ovary, uterus/cervix, vagina and female mammary gland. Testes and epididymides weights. Histopathologic changes in testes, epididymides, male accessory sex organs and male mammary gland	decreased testes weight. Histopathologic changes in testes, epididymides, male accessory sex organs	weight. Histopathologic changes in testes, epididymides, male accessory sex organs	thyroid gland.	uterus/cervix, vagina and female mammary gland. Weight of testes and epididymides, Histopathologic changes in testes, epididymides, male accessory sex organs and male mammary gland	
OECD TG 451-3: Combined Chronic Toxicity/Carcinogenicity	Increased uterus weight, decreased ovary	Studies using pure antagonists are lacking.	Decreased ovary weight.	Histopathologic changes in ovary, uterus/cervix,	Increased thyroid weight.	Possible effects on:	Changes in adrenal weight.

Test guideline or other test method [Reference to interpretation table within this document]	Endpoints for estrogen-mediated activity		Endpoints for androgen-mediated activity		Endpoints for thyroid-related activity	Endpoints for steroidogenesis-related activity	Endpoints potentially sensitive to, but not diagnostic of, EATS modalities
	Agonistic	Antagonistic	Agonistic	Antagonistic			
Studies [Table Annex 2.7]	weight. Histopathologic changes in ovary, uterus/cervix, vagina and female mammary gland. Decrease in weight of epididymides, Histopathologic changes in testes, epididymides, male accessory sex organs and male mammary gland.	However, changes may occur in the following: Uterus and ovary weight.Histopathologic changes in ovary, uterus/cervix, vagina and female mammary gland. Testes and epididymides weights. Histopathologic changes in testes, epididymides, male accessory sex organs and male mammary	Histopathologic changes in ovary, uterus/cervix, vagina. Increased weight of epididymides, decreased testes weight. Histopathologic changes in testes, epididymides, male accessory sex organs	vagina. Decreased weight of epididymides, increased testes weight. Histopathologic changes in testes, epididymides, male accessory sex organs	Possible liver weight increase (in combination with other thyroid-related endpoints). Histopathologic changes in thyroid gland.	Uterus and ovary weight. Histopathologic changes in ovary, uterus/cervix, vagina and female mammary gland. Weight of testes and epididymides, Histopathologic changes in testes, epididymides, male accessory sex organs and male mammary gland	Histopathologic changes in adrenal, and pituitary glands. Tumour types.

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
		gland					
OECD TG 421 and 422: Combined 28-day reproductive screening tests [Table Annex 2.8]	Histopathologic changes in ovary and uterus Decrease in weight of epididymides, Histopathologic changes in testes, epididymides and male accessory sex organs.	Studies using pure antagonists are lacking. However, changes may occur in the following: Uterus and ovary weight. Histopathologic changes in ovary and uterus Testes and epididymides weights. Histopathologic changes in testes, epididymides, male accessory	Histopathologic changes in ovary and uterus Increased weight of epididymides, decreased testes weight. Histopathologic changes in testes, epididymides, male accessory sex organs.	Histopathologic changes in ovary and uterus Decreased weight of epididymides, increased testes weight. Histopathologic changes in testes, epididymides, male accessory sex organs.	Possible liver weight increase (in combination with other thyroid-related endpoints). Histopathologic changes in thyroid gland.	Possible effects on: Histopathologic changes in ovary and uterus Weight of testes and epididymides, Histopathologic changes in testes, epididymides and male accessory sex organs.	Changes in adrenal weight. Histopathologic changes in adrenals. Changes in fertility, reproduction or fetal development. Gestation length Dystocia Gestation length Placental weight Number of implantations, corpora lutea Number of live births and pre and

Test guideline or other test method **[Reference to interpretation table within this document]**	**Endpoints for estrogen-mediated activity**		**Endpoints for androgen-mediated activity**		**Endpoints for thyroid-related activity**	**Endpoints for steroidogenesis-related activity**	**Endpoints potentially sensitive to, but not diagnostic of, EATS modalities**
	Agonistic	**Antagonistic**	**Agonistic**	**Antagonistic**			
		sex organs and male mammary gland.					post implantation loss

In Vitro Screens

Annex 2.1 The Human AR Transactivation Activation Assay for Detection of Androgen (Ant)agonist-Activity of Chemicals (AR STTA)

Annex 2.1.1 Modality detected/endpoints: Activation of reporter gene linked to AR (agonist assay). Inhibition of activation of reporter gene linked to AR (antagonist assay).

Note: No guidelines are available for this assay yet

Background to the Assay

Annex 2.1.2 The Stably Transfected AR Transcriptional Assay (AR STTA) is an *in vitro* screening assay to detect substances that bind to AR and activate the transcription of androgen responsive genes. It is an *in vitro* tool that provides mechanistic data. Several AR STTA assays in common use can be found in the literature (Hartig *et al*, 2002; Birkhøj *et al*, 2004), one of the first versions of this assay used was the "yeast androgen screen" (Sohoni and Sumpter, 1998) which is still widely used for screening of environmental samples. A number of AR STTA assays are being validated via OECD initiatives and a TG for this assay will be developed in the future.

Annex 2.1.3 The AR STTA provides a positive or negative result for the ability of a chemical to induce AR-mediated transactivation of gene expression (agonist assay) compared to a vehicle control. The antagonist assay determines whether a reduction in response occurs when cells are co-exposed to chemical and a potent androgen agonist compared to the potent androgen agonist alone. R1881 or dihydrotestosterone are commonly used as the co-administered agonist. The most common, currently used reporter gene is luciferase (*e.g.* in MDA-kb2 cells).

Annex 2.1.4 The AR STTA gives a positive or negative result for a test chemical when reporter gene activity is compared to controls. A measure of potency is also provided by the magnitude of the effect and the concentration at which it occurs.

Annex 2.1.5 Performance criteria are useful when evaluating results from this assay, although in the absence of a guideline they are not always used. The response with positive control chemicals (*e.g.* hydroxy-flutamide for antagonism and dihydrotestosterone for agonism) should be robust and cell viability should be above 80%.

Annex 2.1.6 Some cell lines used for the AR STTA also express the glucocorticoid receptor (GR) which may cause cross-talk interference with AR (Hartig *et al*, 2002). The level of GR expression in the cell-line and therefore potential for interference should be known.

Annex 2.1.7 The AR STTA will not detect substances that act by other mechanisms *e.g.* ER, TR and steroidogenesis interference. These chemicals will, however, be detected in ER, TR and steroidogenesis specific assays and therefore results from a suite of *in vitro* tests should be considered together. The assay will not detect substances that act by affecting the HPG as an *in vivo* intact axis is required for this.

When/Why the Assay May be Used

Annex 2.1.8 Although the AR STTA may be used at any stage in the hazard assessment process, the most likely use scenario is during initial assessment of chemicals for their ability to interact with endocrine systems *in vitro*, *i.e.* EATS modalities. The AR STTA is frequently conducted following a positive result in the AR binding assay. Assays for interaction with other modalities *e.g.* AR, ER and steroidogenesis, are likely to be conducted at the same time so that all results can be considered together. TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. AR STTAs do not include the use of a xenobiotic metabolising system but consideration should be given to the inclusion of this (OECD, 2008a; Jacobs *et al*, 2008) depending upon the circumstances *e.g.* if the metabolism of a chemical is unknown, although the methods for inclusion of xenobiotic metabolising systems are not yet validated (see paragraph B.18). Alternatively, for a chemical with known metabolites, these could also be tested in the AR STTA.

Annex 2.1.9 Another use scenario may be following effects obtained in higher tier tests, for example accelerated puberty onset in males, but which are not exclusively indicative of an effect on AR. Selection of the most appropriate tests has to be on a case-by-case basis but also considering the need to minimise animal testing.

Introduction to the Table of Scenarios

Annex 2.1.10 Table Annex 2.1 gives guidance on a further step to take in the event of a positive (+) or negative (-) result in the AR STTA and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

Annex 2.1.11 The results of the AR STTA are given in the second column. For the purposes of this guidance, a positive agonist response would be a statistically significant increase in the response compared to the vehicle control value whilst a positive antagonist response would be a statistically significant reduction in the agonist-stimulated response compared to the agonist-stimulated control value.

Annex 2.1.12 Equivocal results for the guideline are not included in the table because these data require further interrogation about the result itself. This assay is a screen and therefore a clear positive or negative result should be obtained. In the event of an equivocal result, the considerations mentioned above about control quality and proficiency criteria should be taken into account and further investigations made.

Existing Data to be Considered

Annex 2.1.13 Existing "Mechanism" *in vitro* data are assumed to be available from AR, ER and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform.

Annex 2.1.14 Existing "Effects" data refer to *in vivo* effects "of concern" *i.e.* data from level 3, 4 or 5 mammalian or wildlife assays/tests. These may come from varied sources and will depend upon the type of substance (*e.g.* new chemicals, HPV chemical, pesticide). Thus, available data may range from repeated dose toxicity studies (28-day, 90-day), or combined repeat dose/reproductive screening assays or fish screening assays, to chronic toxicity studies and multi-generation reproductive tests in

mammalian or wildlife species. Some studies fail to identify EDs that weakly affect oestrogen or androgen receptors as was demonstrated on the basis of data generated in the validation process of the OECD TG 407 assay with endocrine endpoints. In this validation only moderate EDs, such as nonylphenol and DDE, and strong EDs such as ethinylestradiol and flutamide (acting via ER and AR respectively) were detected. Thus OECD TG 407 cannot be regarded as a screening assay for endocrine activity. This means that when a relatively insensitive test is positive for both endocrine-specific and apical endpoints, this should be taken as an indication that the substance is a possible ED. Caution should be exercised, however, because endocrine endpoints may be impacted secondary to non-endocrine toxicity (*e.g.* hepatotoxicity) and *in vivo* apical endpoints can be affected by all modes of action including endocrine modalities. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

Annex 2.1.15 Data may also be available from level 3 tests (H and UT assays) although these tests may not give rise to "concern" as they are hazard screening tests only. The H assay is, however, more likely to be conducted after the AR STTA (to test whether a chemicals that is positive *in vitro* is also positive *in vivo*) rather than before. An AMA may also be available but as this test primarily detects thyroid disruption in amphibians it is unlikely to provide useful data for E-modalities.

Annex 2.1.16 When considering the results of the AR STTA, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may and read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

Annex 2.1.17 The scenarios (A to R) presented in table Annex 2.1 and represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

Annex 2.1.18 Scenarios A to I represent positive results in the AR STTA in the presence of positive and/or negative *in vitro* mechanistic data and positive and/or negative *in vivo* effects data. The possibilities of equivocal or missing existing data are given in scenarios C, F, G, H and I. Scenarios J to R represent negative results in the AR STTA in the presence of positive and/or negative *in vitro* mechanistic data and positive and/or negative *in vivo* effects data. The possibilities of equivocal or missing existing data are given in scenarios L, O, P, Q and R. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to taken on a case-by-case basis; for example, some equivocal data may be considered positive whilst in other cases no conclusions may be possible and therefore the situation is effectively "data not available".

Annex 2.1.19 The next step to take to increase evidence will depend upon the existing information. The table is meant to provide a succinct guide and may not cover all circumstances or possibilities.

Table Annex 2.1 The Human Androgen Receptor Transactivation Assay for Detection of Androgen (Ant)agonist-Activity of Chemicals (AR STTA). Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and S based assays (level 2). The AR binding assay is likely to be performed prior to the AR STTA. TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes various information, such as data from repeat oral toxicity studies, reproduction/developmental toxicity screen tests, read across from analogues, will be available.

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	AR (ant)agonism combined with effects on ER/T/S and potential for adverse effects via multiple mechanisms.	Perform assay from upper levels *e.g.* H assay or fish screen (AFSS) (level 3) or male PP assay (level 4) or ext-1 or 2-gen assays or partial/full fish life cycle tests (level 4/5).	If existing data are from level 5 then may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results and *in vivo.* results but may also be metabolised to a metabolite that also has positive results *in vitro* and *in vivo*.
B	+	+	-	AR (ant)agonism combined with effects on ER/T/S but effects not detected in *in vivo* studies. Weak AR (ant)agonism does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform AR STTA with added metabolising system OR assay from upper levels *e.g.* H assay or fish screen (AFSS) (level 3) or male PP assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species.
C	+	+	Eq/0	AR (ant)agonism combined with effects on ER/T/S but no or equivocal data from *in vivo* studies Weak AR (ant)agonism may not result in adverse effects.	Perform assay from upper levels *e.g.* H assay or fish screen (AFSS) (level 3) or male PP assay (level 4).	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
D	+	-	+	AR (ant)agonism and potential for adverse effects.	Perform assay from upper levels *e.g.* H assay or	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					fish screen (AFSS) (level 3) or male PP assay (level 4).	concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
E	+	-	-	AR (ant)agonism but effects not detected in *in vivo* studies. Weak AR (ant)agonism does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform AR STTA with added metabolising system or assay from upper levels *e.g.* H assay or fish screen (AFSS) (level 3) or male PP assay (level	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption. (the ext-1 gen assay provides the most information) If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					4).	endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species. Check data on chemical analogues.
F	+	-	Eq/0	AR (ant)agonism but no or equivocal data from *in vivo* studies. Weak AR (ant)agonism does not result in adverse effects.	Perform assay from upper levels *e.g.* H assay or fish screen (AFSS) (level 3), male PP assay (level 4)	Consider route of exposures for equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
G	+	Eq/0	+	AR (ant)agonism and potential for adverse	Perform assay	If existing data are from an adequate

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				effects via AR (ant)agonism or other ETS mechanisms. May act via EATS mechanism and may or may not require metabolic activation.	from upper levels *e.g.* H assay or fish screen (AFSS) (level 3) or male PP assay (level 4).	level 5 assay then there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Check data on chemical analogues. Further mechanistic studies would help determine MoA. A positive result could have arisen from other (EATS or non-EATS) mechanisms *e.g.* HPG axis.
H	+	Eq/0	**-**	AR (ant)agonism but effects not detected in *in vivo* studies. Weak AR (ant)agonism does not result in adverse effects. Metabolic differences explain *in vitro*/*in vivo* differences.	Perform AR STTA with added metabolising system	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information).

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Consider route of exposures for existing effects data and possible implications of ADME characteristics of the chemical. The chemical itself may give positive *in vitro* results but may not be absorbed or may be metabolised to an inactive metabolite *in vivo*. However, note that uptake and metabolism of chemicals can be different between wildlife species. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	AR (ant)agonism with unknown potential for	Perform AR	Consider route of exposures for

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				effects in *in vivo* studies. May act via AR mechanism and may or may not require metabolic activation. Unknown potential for adverse effects.	STTA with added metabolising system or H assay or fish screen (AFSS) (level 3) if existing data indicates this is needed	equivocal existing effects data and possible implications of ADME characteristics of the chemical. Check data on chemical analogues.
J	-	+	+	No evidence for AR (ant)agonism. Effects on ER/T/S and potential for adverse effects via EATS mechanisms.	Perform AR STTA with added metabolising system or Perform assay from upper levels (*e.g*. H assay or fish screen AFSS (level 3) or male PP assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Consider route of exposures for existing effects data and possible

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						implications of ADME characteristics of the chemical. Check data on chemical analogues. Further mechanistic studies may help determine MoA.
K	-	+	-	No evidence for AR (ant)agonism. Effects on ER/T/S but effects not detected in *in vivo* studies.	Perform assay from upper levels *e.g.* fish screen (AFSS) (level 3) or male or female PP assay or (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information) If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* EATS activity is not realised. Consider possible routes of

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						exposure, implications of metabolism.
L	-	+	Eq/0	No evidence for AR (ant)agonism. Effects on ER/T/S but effects not detected in *in vivo* studies. Unknown potential for adverse effects. Metabolic differences explain *in vitro*/*in vivo* EATS differences.	Perform assay from upper levels *e.g.* fish screen (AFSS) (level 3) or male or female PP assay (level 4).	Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* EATS activity is not realised. Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues.
M	-	-	+	No evidence for AR (ant)agonism. Metabolic differences or route of exposure explain *in vitro*/*in vivo* differences. Effects seen in existing studies are via non-EATS or non-endocrine mechanism.	Perform AR STTA with added metabolising system or Perform assay from upper levels *e.g.* fish screen (AFSS) (level 3) or male or female PP assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Metabolic activation of chemical may occur *in vivo*.

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						Consider possible routes of exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
N	-	-	-	No evidence for AR (ant)agonism. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may perform assay from upper levels *e.g.* fish screen (AFSS) (level 3) or male or female PP assay (level 4).	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Check data on chemical analogues.
O	-	-	Eq/0	No evidence for AR (ant)agonism. Unknown potential for adverse effects via	Perform AR STTA with added	Consider possible routes of exposure, implications of

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				other mechanisms.	metabolising system or fish screen (AFSS) (level 3) or male or female PP assay (level 4) if existing data indicates this is needed.	metabolism. Check data on chemical analogues.
P	-	Eq/0	+	No evidence for AR (ant)agonism. Unknown potential for adverse effects via other mechanism.	Perform AR STTA with added metabolising system	If existing data are from an adequate level 5 assay there may be sufficient information to conclude concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Consider possible routes of

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
Q	-	Eq/0	-	No evidence for AR (ant)agonism. No evidence of adverse effects.	Perform AR STTA with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from level 4 mammalian or wildlife assay then level 5 assay should provide more comprehensive data on adverse effects on endocrine relevant endpoints over more extensive parts of the life cycle of the organism. If existing data are from H assay or AFSS then level 4 mammalian assay or fish screen (OECD TG 229/230) will provide data on multiple modalities. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
R	-	Eq/0	Eq/0	No evidence for AR (ant)agonism.	For the "0"	Consider possible routes of

Scenarios	Result of AR STTA)	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				Unknown potential for adverse effects via other mechanism.	scenario perform AR STTA with added metabolising system or perform H assay or fish screen (AFSS) (level 3) if existing data indicates this is needed.	exposure, implications of metabolism. Check data on chemical analogues. Further mechanistic studies would help determine MoA.

Wildlife Screens and Tests

Annex 2.2 Fish (Medaka) Multi-Generation Test (MMGT).

Annex 2.2.1 Modality detected/endpoints: Estrogens (♂VTG ↑; phenotypic gonad sex reversal ♂→♀; gonadal histopathology; feminised secondary sex characteristics); Androgens (phenotypic gonad sex reversal ♀→♂; gonadal histopathology; masculinised secondary sex characteristics); Aromatase inhibitors (♀VTG↓; phenotypic gonad sex reversal ♀→♂; gonad histopathology; masculinised secondary sex characteristics).
In addition to sex reversal, the test also measures two other apical endpoints (fecundity and fertility) which respond both to EDs and to some non-EDs. Other endpoints which do not respond specifically to EDs include hatching rate, survival, growth, and non-gonad histopathology. Note that the MMGT has not yet completed validation.

Background to the Assay

Annex 2.2.2 This multi-generation assay runs for 24 weeks, from F0 reproducing adults to F2 pre-reproducing adults, and hence encompasses two complete generations. It is therefore expected to be responsive to most chemicals with EATS modalities, although the full extent of its applicability awaits further validation. It should be noted that if the assay gives a positive result, this may be due to a positive indicator of hormonal activity (*i.e*. VTG, secondary sex characteristics, gonad histopathology), a positive for apical endpoints (fecundity, fertility, gonad sex reversal[5]), or a positive for both types of endpoint. Each of these three possible combinations of positive response should be considered separately, (although the distinctions between indicators of hormonal activity and apical effects are not always clear) so they have been listed individually as points 1, 2 and 3 in the Possible Conclusions column of Table Annex 2.2.

When/Why the Assay May be Used

Annex 2.2.3 Although the MMGT could, in principle, be used at any stage in the hazard assessment process, the probable use scenario will be when there are already *in vitro* or *in vivo* screening data available about the possible endocrine disrupting properties of a chemical. In addition, there may be data from TG 234 (FSDT) and/or an FLCTT. It is highly unlikely that no other existing endocrine-relevant data will be available (*i.e.* if the MMGT has been used as a primary screen). Indeed, it is not recommended to perform this test in the absence of such data. However, in that case a positive result in the MMGT, especially if derived from an endocrine-sensitive apical endpoint, should ideally be followed up with relevant *in vitro* screening to confirm the suspected mode of action.

Existing Data to be Considered

Annex 2.2.4 Existing data available before deployment of the MMGT might include one or more of a range of *in silico* or *in vitro* results which suggest that the modalities indicated above may occur *in vivo,* plus *in vivo* results obtained with other vertebrates (*e.g.* a positive uterotrophic assay with

[5] Note that sex reversal will generally have been caused by exposure to an EDC.

rodents) or a positive result in the screening fish assay assays OECD TG 229 or 230, or in TG 234 (FSDT) or FLCTT). Indicators of possible *in vivo* activity might include (Q)SAR predictions of endocrine activity, 'read-across' from *in vivo* results obtained with chemically related chemicals, or positive results from an *in vitro* screen for estrogen or androgen receptor-mediated activity, or for effects on steroidogenesis (especially aromatase inhibition). Conduct of the MMGT would be particularly relevant if the test chemical has been shown to affect fish sexual development or reproduction in OECD TG 234 (FSDT) or OECD TG 229, although it may also be sensitive to thyroid-disrupting chemicals through their effects on growth and development (this remains to be firmly established). The MMGT will not only provide a reliable indication if the test chemical is an EATS ED, but should also allow calculation of a reliable long-term NOEC or ECx for use in environmental risk assessment.

Scenarios: Positive and Negative Results Combined with Existing Data

Annex 2.2.4a The scenarios (A to R) presented in Table Annex 2.2 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

Annex 2.2.5 Positive results obtained with one or more of the MMGT indicators of hormonal activity but not with apical endpoints (Table Annex 2.2, Scenarios A-I, sub-section 2) result in the conclusion that the test chemical is a possible, but almost certainly not an actual, ED *in vivo* (although biased sex ratio will generally have been caused by an ED). If both an indicator of hormonal activity and an apical endpoint give a response (Table Annex 2.2, Scenarios A-I, sub-section 1), this provides evidence that the chemical is almost certainly an actual ED (*i.e.* it causes adverse effects through an endocrine mechanism). If only an apical endpoint responds (Table Annex 2.2, Scenarios A-I, sub-section 3), it indicates that the chemical is a developmental, growth or reproductive toxicant, with a substantially reduced probability that it is an ED (although existing positive *in vitro* data, or positive *in vivo* data from other species, would have to be weighed against the latter conclusion). However, note that biased sex ratio would generally be expected to have resulted from exposure to an ED.

Annex 2.2.6 As indicated above, an effect on an apical endpoint in the MMGT shows that the test chemical causes adverse effects and is a developmental, growth or reproductive toxicant (assuming that the concentration giving this response is not sufficiently high to cause systemic toxicity). If these results are combined with positive indicators of hormonal activity and/or positive *in vitro* screening assay data, and a plausible causal relationship exists, it is reasonable to conclude that the chemical is an actual ED, and that the information could be used in a risk assessment.

Annex 2.2.7 A situation in which the MMGT gives a negative result needs careful consideration of any existing data. It is unlikely that these data would suggest that the chemical is endocrine-active both *in vitro* and *in vivo* (Table Annex 2.2, Scenario J), but if they do, then the probability is that the species used in the MMGT is different from those already tested positive. Little is known about inter-species variability in response of the MMGT as it has essentially only received validation with medaka. However, a negative high-tier test like the MMGT would generally be considered to trump a positive lower-tier test. One reason for this is the fact that some lower tiered tests are designed to be highly sensitive and may not represent normal physiological conditions.

Annex 2.2.8 If the MMGT and existing *in vivo* data are all negative, but *in vitro* data reveal some endocrine activity (Scenario K), the probability is that the test chemical is not sufficiently potent to produce endocrine effects *in vivo* in fish, or it may be rapidly metabolised. In such a situation, further testing is probably not necessary as the test chemical is almost certainly not an ED.

Annex 2.2.9 Furthermore, if the MMGT and the *in vitro* tests are negative, but there are positive existing *in vivo* data (Scenario M), the chemical is almost certainly not an ED in fish, although it may have endocrine properties in other species.

Annex 2.2.10 Finally, a negative MMGT, set against a background of negative *in vitro* and *in vivo* data (Scenario N), suggests that the test chemical is almost certainly not an ED, and no further testing for estrogenic, anti-estrogenic, androgenic or steroidogenic MOAs should generally be considered.

Annex 2.2.11 In each of the above scenarios, it is possible that existing data will be equivocal, or there may be no existing data (Scenarios C, F, G, H, I, L, O, P, Q and R). This will weaken the conclusions which can be drawn about a negative MMGT, although it will remain likely that the test chemical is not an ED in fish. On the other hand, if the MMGT is positive, no further *in vivo* testing in fish is needed. Again, however, it will always be desirable to obtain some mechanistic information before reaching a conclusion that a positive apical result in the MMGT has been caused by an ED.

Annex 2.2.12 The scenarios in which the results of an MMGT are themselves equivocal have not been dealt with in Table Annex 2.1, for reasons of brevity. In this context, an equivocal result might be an inconsistent concentration-response (*e.g.* no effect at a high concentration but effects at a lower), or a result which borders on statistical significance. Without knowing the exact circumstances, reliable advice cannot be given, but the opinions of an experienced ecotoxicologist should be sought. Clearly, however, such equivocal results do not necessarily rule out the existence of *in vivo* endocrine activity. For example, VTG induction in female medaka at a high concentration might be masked by any systemic toxicity, while an effect on fecundity might just fail to reach a statistically significant level due to the inherently high variability of this endpoint. If these or other possible reasons for false negatives are suspected with good reason, the test could be repeated (*e.g.* conduct it at lower concentrations which avoid systemic toxicity, providing such toxicity was present in all original treatments), or a more appropriate version of it (*e.g.* more fish per replicate) could be conducted. However, given that the scale of any equivocal effect is likely to be small, it may be considered that the high ethical and financial cost of a repeat test is not justified.

Annex 2.2.13 In summary, a positive apical response in the MMGT indicates that a chemical is a growth, developmental or reproductive toxicant which may or may not be an ED. A combination of a positive apical response and a positive endocrine-specific endpoint (*e.g.* vitellogenin) at similar concentrations is strong evidence that the chemical is an actual ED, especially if the two types of endpoint are causally related and if positive mechanistic data are also available. In this situation, further *in vivo* data from fish are unlikely to be required. On the other hand, negative results in the MMGT suggest that the test chemical is not an ED, at least not in the fish species under test – a judgement about the likelihood of endocrine effects in other species will have to be made in the light of existing *in vitro* and *in vivo* data.

Table Annex 2.2 Medaka Multi-Generation Test (MMGT). Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR, and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". There is no evidence at present that equivalent *in vitro* assays with systems derived from fish offer advantages over their mammalian counterparts.

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter.

The assay under discussion could either be positive for both apical endpoints and indicators of hormonal activity, or positive just for apical endpoints, or positive just for indicators of hormonal activity. For each scenario, each of these 3 possibilities is addressed separately in the Possible Conclusions column.

Scenario	**Result of MMGT**	**Existing Results**		**Possible conclusions**	**Next step which could be taken to increase evidence if necessary**	**Other considerations**
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****	**1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive**		
A	+	+	+	1) Strong evidence for	Probably no need for	-

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				adverse effects in fish and other organisms by an endocrine mechanism. 2) Strong evidence for endocrine effects, but they do not appear adverse in fish[6]. 3) Strong evidence for adverse effects in more than one organism, but mechanism may not be via direct interaction with ER or AR, or by aromatase inhibition.	additional data.	
B	+	+	-	1) Strong evidence for	Probably no need for	-

[6] Note however, that biased sex ratio should be considered as an adverse response.

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				adverse effects in fish by an endocrine mechanism. 2) Strong evidence for endocrine effects in fish, but they do not appear adverse. 3) Strong evidence for adverse effects in fish, but mechanism may not be via direct interaction with ER or AR, or by aromatase inhibition.	additional data.	
C	+	+	Eq/0**	1) Strong evidence for adverse effects in fish by an endocrine mechanism. 2) Strong evidence for endocrine effects in fish, but they do not	Probably no need for additional data.	-

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				appear adverse. 3) Strong evidence for adverse effects in fish, but mechanism may not be via direct interaction with ER or AR, or by aromatase inhibition.		
D	+	-	+	1) Strong evidence for adverse effects in fish and other organisms, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in fish, but they do not appear to be adverse. 3) Strong evidence for	Probably no need for additional data, but see column to the right.	Negative *in vitro* mechanistic data combined with positive endocrine-specific *in vivo* endpoints suggests that an unknown endocrine MOA is causing any adverse effects. Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in this scenario (sub-section 1) is an ED.

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				adverse effects in more than one organism, but mechanism may not be by endocrine disruption.		
E	+	-	-	1) Strong evidence for adverse effects in fish, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in fish, but they do not appear to be adverse. 3) Strong evidence for adverse effects in fish, but mechanism may not be by endocrine disruption.	Probably no need for additional data, but see column to the right.	Negative *in vitro* mechanistic data combined with positive endocrine-specific *in vivo* endpoints suggests that an unknown endocrine MOA is causing any adverse effects. Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in this scenario (sub-section 1) is an ED, especially as existing *in vivo* data are negative.
F	+	-	Eq/0	1) Strong evidence for adverse effects in fish,	Probably no need for additional data, but see	Negative *in vitro* mechanistic data combined with positive endocrine-specific *in vivo*

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in fish, but they do not appear to be adverse. 3) Strong evidence for adverse effects in fish, but mechanism may not be by endocrine disruption.	column to the right.	endpoints suggests that an unknown endocrine MOA is causing any adverse effects. Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in this scenario (sub-section 1) is an ED, especially as existing *in vivo* data are equivocal or absent. However, such a conclusion appears likely on the basis of the positive endocrine-sensitive endpoints *in vivo*.
G	+	Eq/0	+	1) Strong evidence for adverse effects in more than one organism, possibly by an unknown endocrine mechanism. 2) Medium-Strong	It would be desirable to obtain some clear mechanistic data before concluding that the chemical is an ED. See column to right.	Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in sub-section 1 is an ED. However, such a conclusion appears likely on the basis of the positive endocrine-sensitive endpoints *in vivo*.

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				evidence for endocrine effects, but they do not appear to be adverse in fish. 3) Strong evidence for adverse effects in more than one organism, but mechanism may not be by endocrine disruption.		
H	+	Eq/0	-	1) Strong evidence for adverse effects in fish, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in fish, but they do not appear to be adverse. 3) Strong evidence for adverse effects in fish,	It would be desirable to obtain some clear mechanistic data before concluding whether the chemical is an ED. See column to right.	Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in sub-section 1 is an ED. However, such a conclusion appears likely on the basis of the positive endocrine-sensitive endpoints *in vivo*.

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				but mechanism may not be by endocrine disruption.		
I	+	Eq/0	Eq/0	1) Strong evidence for adverse effects in fish, possibly by an unknown endocrine mechanism. 2) Moderate-Strong evidence for endocrine effects in fish, but they do not appear to be adverse. 3) Strong evidence for adverse effects in fish, but mechanism may not be by endocrine disruption.	It would be desirable to obtain some clear mechanistic data before concluding whether or not the chemical is likely to be an ED. See column to right.	Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in sub-section 1 is an ED. However, such a conclusion appears likely on the basis of the positive endocrine-sensitive endpoints *in vivo*.
J	-	+	+	The chemical is an ED	Regulatory authorities	The fact that the chemical has endocrine

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				in vivo in other species but does not appear to act on growth, sexual development or reproduction in fish. If any other fish tests are also negative, fish may not be responsive at all to the test chemical.	may consider that further evidence is not required.	properties *in vitro* and in other species *in vivo* suggests that it may be an ED, but probably not in fish . If the existing positive *in vivo* data are from a lower tier fish assay, note that it is generally considered that a negative higher tier test trumps a positive lower tier test.
K	-	+	-	Despite the *in vitro* mechanistic data for possible endocrine activity, there is no evidence for endocrine disruption *in vivo*. This may be because the chemical is d*egra*ded to an inactive metabolite, or because it only interacts very weakly	Regulatory authorities may consider that further evidence is not required.	-

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				with endocrine receptors.		
L	-	+	Eq/0	The chemical is not an ED in fish, but it may be active in other species as there is only one unequivocal *in vivo* test result (the negative MMGT).	Regulatory authorities may consider that further evidence is not required.	-
M	-	-	+	The chemical is probably not an ED acting on growth, sexual development or reproduction in fish, but it does have endocrine activity in other species. However, it may act through MOAs not	Regulatory authorities may consider that further evidence is not required.	The fact that the chemical has endocrine properties in other species *in vivo* suggests that it may be an ED, but probably not in fish . If the existing positive *in vivo* data are from a lower tier fish assay, note that it is generally considered that a negative higher tier test trumps a positive lower tier test.

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				covered by the available *in vitro* assays, or it may be more potent in a fish species other than that tested.		
N	-	-	-	The chemical is probably not an ED in fish or other species.	Regulatory authorities may consider that further evidence is not required.	-
O	-	-	Eq/0	The chemical is probably not an ED in fish.	Regulatory authorities may consider that further evidence is not required.	-
P	-	Eq/0	+	The chemical is probably not an ED acting on growth, sexual development or reproduction in fish, but it does have endocrine activity in other species.	Regulatory authorities may consider that further evidence is not required.	The fact that the chemical has endocrine properties in other species *in vivo* suggests that it may be an ED, but probably not in fish. If the existing positive *in vivo* data are from a lower tier fish assay, note that it is generally considered that a negative higher tier test trumps a positive lower tier test.

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				However, it may act through MOAs not covered by the available *in vitro* assays, or it may be more potent in a fish species other than that tested.		
Q	-	Eq/0	-	The chemical is probably not an ED acting on growth, sexual development or reproduction in fish, or *in vivo* on other species.	Regulatory authorities may consider that further evidence is not required, although negative *in vitro* data would strengthen the conclusion that the chemical is probably not an ED.	-
R	-	Eq/0	Eq/0	The chemical is probably not an ED	Regulatory authorities may consider that	-

Scenario	Result of MMGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				acting on growth, sexual development or reproduction in fish.	further evidence is not required, although negative *in vitro* data would strengthen the conclusion that the chemical is probably not an ED.	

Annex 2.3 Larval Amphibian Growth and Development Assay (LAGDA)

Annex 2.3.1 Modality detected/endpoints: This assay is undergoing validation, so it is not certain which endpoints are specifically responsive to thyroid disrupters and which to other EDs such as estrogens and androgens. However, the draft LAGDA TG (dated 20/10/10) indicates that there are 3 endpoints indicating generalised toxicity (mortality; abnormal behaviour; and growth), and several specifically indicative of endocrine disruption or impaired reproduction (histopathology of thyroid and gonads,; time to metamorphosis (NF stage 62); secondary sex characteristics (nuptial pads); blood hormones (T4 and TSH); vitellogenin; genetic and phenotypic sex ratio). It is presumed that most of these specific endocrine endpoints are likely to respond to interference with the HPG axis, while thyroid histopathology, thyroid hormones, and time to metamorphosis may respond to interference with the HPT axis (as may the 'generalised toxicity' indicator, growth).

Background to the Assay

Annex 2.3.2 This assay is a partial lifecycle test with the clawed frog *Xenopus laevis*. It starts with NF stage 8 F0 larvae and ends 10 weeks after the median time that controls take to reach NF stage 62 F0 juveniles. In essence, therefore, it covers the stages of larval/juvenile growth and sexual development, but not those of reproduction and embryonic development. It could therefore be thought of as the amphibian near-equivalent of the Fish Sexual Development Test (FSDT) (TG 234), although it also includes endpoints that are specifically responsive to thyroid disrupters. It does not include all processes which may respond to EATS EDs (especially reproduction), and it is currently unknown whether the LAGDA is therefore less responsive to some of these chemicals than a lifecycle test (a standardised protocol for which is not available). When validation data are available for the LAGDA, and when the responsiveness of the FLCTT or MMGT to thyroid disrupters is better known, it may be concluded that fish lifecycle tests are a more suitable alternative than the LAGDA or amphibian lifecycle testing.

Annex 2.3.3 The draft TG provides a table of endpoints (TG Table 1) which are referred to as 'apical', but many of these should more properly be considered as indicators of hormonal activity (*e.g*. nuptial pad development and thyroid hormone titres *etc*.). Probably the only true apical endpoints which could be used in a risk assessment (because they can be related directly to adverse effects on populations) are mortality, growth and phenotypic/genotypic sex ratio. The latter two are likely to be responsive to some EDs, but also to certain other chemicals. On the other hand, indicators of hormonal activity of use in diagnosing the effects of EDs, but probably not of value in risk assessments, include gonad and thyroid histopathology, time to metamorphosis, nuptial pad development, blood hormones, and vitellogenin. The endpoints will be grouped in this way for the purposes of this document, but it should be borne in mind that the true value of the indicators of hormonal activity for diagnosing EDs has not yet been validated.

Annex 2.3.4 Consequently, if the assay gives a positive result, this may be due to a combination of a positive indicator of hormonal activity and a positive apical endpoint, or a positive for an indicator of hormonal activity alone, or for an apical endpoint alone. Each of these possible combinations of positive response should be considered separately (although the distinctions between indicators of hormonal activity and apical effects are not always clear), so they have been listed individually as points 1, 2 and 3 in the Possible Conclusions column of Table Annex 2.2.

When/Why the Assay May be Used

Annex 2.3.5 Although the LAGDA could, in principle, be used at any stage in the hazard assessment process, the most likely use scenario will be when there are some data available about the possible

thyroid disrupting properties of a chemical, or if the chemical is suspected of having (anti)estrogenic or (anti)androgenic properties. Thus, there are likely to be data available from *in vitro* mechanistic screens, as well as *in vivo* wildlife screens such as OECD TG 229, 230 and/or 231. Furthermore, a number of mammalian (rat) assays (which may have been performed before any wildlife testing) are sensitive to thyroid disruption, including the pubertal assay (male or female), the enhanced repeat dose assay (OECD TG 407), and the intact male screening assay. Rodent screening assays (*e.g.* the Hershberger or Uterotrophic assays) with responsiveness to other EDs (*e.g.* androgens or estrogens) may also have been conducted.

Annex 2.3.6 It is unlikely that no endocrine-relevant data will available before the LAGDA is deployed (*i.e.* the LAGDA has been used as a primary screen), but in that case a positive result in the LAGDA should ideally be followed up with relevant *in vitro* screening to investigate the suspected mode of action. However, it should be noted that while *in vitro* screens are available for estrogens, androgens and steroidogenesis inhibitors, they additionally exist only for thyroid agonists and antagonists (*e.g.* GH_3 rat pituitary somatotroph cell proliferation; solid state thyroid receptor binding assays; transfected reporter gene assays in yeast or mammalian cell lines), while thyroid disruption can occur at other points in the endocrine system for which *in vitro* screens do not exist, or are still at the research stage (*e.g.* FRTL-5 rat cell lines sensitive to iodide uptake inhibitors). Furthermore, none of these thyroid screens have yet been validated and standardised at the international level.

Existing Data to be Considered

Annex 2.3.7 Existing data available before deployment of the LAGDA might include *in vivo* results obtained with other vertebrates (*e.g.* a positive *in vivo* assay with rats or fish – see above), or one or more of a range of *in silico* or *in vitro* results which suggest that estrogenic, androgenic or thyroid disruption may occur *in vivo* (but note the limitations of this approach for thyroid disrupters, as indicated above). Such indicators of possible endocrine activity might include (Q)SAR predictions, 'read-across' from *in vivo* results obtained with chemically related chemicals, or positive results from an *in vitro* screen.

Scenarios: Positive and Negative Results Combined with Existing Data

Annex 2.3.7a The scenarios (A to R) presented in Table Annex 2.3 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

Annex 2.3.8 Positive results obtained with an indicator of hormonal activity in the LAGDA but not with apical endpoints (Table Annex 2.3, Scenarios A-I, sub-section 2) result in the conclusion that the test chemical is probably a possible ED *in vivo*. If both an indicator of hormonal activity and an apical endpoint give a response (Table Annex 2.3, Scenarios A-I, sub-section 1), this provides evidence that one is dealing with an actual ED with adverse effects *in vivo*. If only an apical endpoint responds (Table Annex 2.3, Scenarios A-I, sub-section 3), it suggests that the chemical is harmful to growth or sexual development, but is not necessarily an ED (although existing positive *in vitro* data, or positive *in vivo* data from other species, would have to be weighed against this conclusion).

Annex 2.3.9 The situation in which a LAGDA gives a negative result (Table Annex 2.3, Scenarios J-R) needs careful consideration of any existing data. If these data suggest that the chemical is endocrine-active both *in vitro* and *in vivo* (Scenario J), then it is possible that the LAGDA is simply

insufficiently sensitive (perhaps because it does not include reproduction). Depending on the robustness of the existing data, it might therefore be appropriate to conduct an amphibian lifecycle test, although a protocol for one has not been standardised or validated.

Annex 2.3.10 If the LAGDA and existing *in vivo* data are all negative, but *in vitro* data reveal some endocrine activity (Scenario K), the probability is that the test chemical is not sufficiently potent to produce effects *in vivo* in amphibians or other organisms, or it may be rapidly metabolised. In such a situation, further testing is probably not necessary.

Annex 2.3.11 On the other hand, if the LAGDA and the *in vitro* tests are negative (Scenario M), but there are positive existing *in vivo* data, the nature of those existing data should be considered. Unless the existing data are from another amphibian, the chemical is probably not a possible ED acting on amphibian growth or development, but it may act via MOAs not covered by the *in vitro* screens, or it may be more potent in species or life-stages that have not been tested. In this situation, the existing *in vivo* data should be used to guide decisions about whether to conduct any further testing.

Annex 2.2.12 Finally, a negative LAGDA, set against a background of negative *in vitro* and *in vivo* data (Scenario N), suggests that the test chemical is not a possible EATS ED, and further action is unnecessary.

Annex 2.3.13 In each of the above scenarios, it is possible that existing data will be equivocal, or there may be no existing data. This will weaken the conclusions which can be drawn about a negative LAGDA, and this is reflected in Table Annex 2.3. However, a lack of *in vitro* mechanistic data should ideally be rectified before any further *in vivo* testing is finally rejected, although as indicated above, many thyroid modalities are not detectable in *in vitro* screens. On the other hand, if the LAGDA is positive, further *in vivo* testing would not generally be needed unless it is suspected that the chemical acts primarily on reproduction. Again, however, it may be useful to obtain some mechanistic information before conducting further *in vivo* testing, although note that a validated amphibian lifecycle protocol is unavailable. A possible substitute for the latter might be a fish lifecycle test (either the FLCTT or MMGT), although the responsiveness of such a procedure to thyroid disrupters is unknown.

Annex 2.3.14 The scenario in which the results of a LAGDA are themselves equivocal has not been dealt with in Table Annex 2.3, for reasons of brevity. In this context, an equivocal result might be an inconsistent concentration-response (*e.g.* no effect at a high concentration but effects at a lower), or a result which borders on statistical significance. Without knowing the exact circumstances, reliable advice cannot be given, but the opinions of an experienced ecotoxicologist should be sought. Clearly, however, such equivocal results do not necessarily rule out the existence of *in vivo* endocrine activity. For example, thyroid histopathology at a high concentration might be masked by any systemic toxicity, while growth measurements might just fail to reach a statistically significant level due to unexpectedly high variability. If these or other possible reasons for false negatives are suspected with good reason, the test could be repeated (*e.g.* conduct it at lower concentrations which avoid systemic toxicity), or a more appropriate version of it (*e.g.* more larvae per replicate) could be designed and conducted.

Annex 2.3.15 In summary, positive indicators of hormonal activity in the LAGDA indicate that a chemical is a possible ED via one of several modalities (not fully validated as yet), while a combination of positive indicators of hormonal activity and positive apical results suggest that it is an actual ED (especially if the two types of response are causally related). However, if an apical endpoint alone responds, the chemical may not be an ED (although existing data may help to inform this decision). Negative results in the LAGDA do not necessarily mean that the chemical is not an ED – a judgement about possible endocrine disruption and the possible need for additional testing will have to be made in the light of existing *in vitro* and *in vivo* data.

Table Annex 2.3 Larval Amphibian Growth and Development Assay (LAGDA). Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from available from ER and AR based assays and the steroidogenesis assay (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an thyroid disrupter.

The assay under discussion could either be positive for both apical endpoints and indicators of hormonal activity, or positive just for an apical endpoints or indicators of hormonal activity. For each scenario, each of these 3 possibilities is addressed separately in the Possible Conclusions column.

Scenarios	Result of LAGDA	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**	1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive		
A	+	+	+	1) Strong evidence for *in vivo* endocrine activity with adverse effects (on growth or sexual development) in amphibians, and effects in other species 2) Strong evidence for *in vivo* endocrine activity in amphibians and other species 3) Strong evidence for adverse effects on growth or sexual development in amphibians, and effects in other species, but possibly not	Regulatory authorities may consider that further data from amphibians are not required. However, see column to the right.	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction is not currently available

Scenarios	Result of LAGDA	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				via an endocrine mechanism.		
B	+	+	-	1) Strong evidence for *in vivo* endocrine activity with adverse effects (on growth or sexual development) in amphibians 2) Strong evidence for *in vivo* endocrine activity in amphibians 3) Strong evidence for adverse effects on growth or sexual development in amphibians, but possibly not via an endocrine mechanism.	Regulatory authorities may consider that further data from amphibians are not required. However, see column to the right.	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction is not currently available
C	+	+	Eq/0	1) Strong evidence for *in vivo* endocrine activity with adverse effects (on growth or sexual development) in amphibians 2) Strong evidence for *in vivo* endocrine activity in amphibians 3) Strong evidence for adverse effects on growth or sexual development in amphibians, but possibly not via an endocrine mechanism.	Regulatory authorities may consider that further data from amphibians are not required. However, see column to the right.	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction is not currently available
D	+	-	+	1) Strong evidence for *in vivo* endocrine activity with adverse effects (on growth or sexual development) in amphibians and other species, but possibly not via an EATS mechanism.	Regulatory authorities may consider that further data from amphibians are	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction

Scenarios	Result of LAGDA	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				2) Strong evidence for *in vivo* endocrine activity in amphibians and other species, but possibly not via an EATS mechanism. 3) Strong evidence for adverse effects on growth or sexual development in amphibians and other species, but probably not via an endocrine mechanism.	not required. However, see column to the right.	is not currently available
E	+	-	-	1) Strong evidence for *in vivo* endocrine activity with adverse effects (on growth or sexual development) in amphibians, but possibly not via an EATS mechanism. 2) Strong evidence for *in vivo* endocrine activity in amphibians, but possibly not via an EATS mechanism. 3) Strong evidence for adverse effects on growth or sexual development in amphibians, but probably not via an endocrine mechanism.	Regulatory authorities may consider that further data from amphibians are not required. However, see column to the right.	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction is not currently available
F	+	-	Eq/0	1) Strong evidence for *in vivo* endocrine activity with adverse effects (on growth or sexual development) in amphibians, but possibly not via an EATS mechanism. 2) Strong evidence for *in vivo* endocrine activity in amphibians, but possibly not via an EATS mechanism. 3) Strong evidence for adverse effects on growth or sexual development in amphibians,	Regulatory authorities may consider that further data from amphibians are not required. However, see column to the right.	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction is not currently available

Scenarios	Result of LAGDA	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
				but probably not via an endocrine mechanism.		
G	+	Eq/0	+	1) Strong evidence for *in vivo* endocrine activity with adverse effects (on growth or sexual development) in amphibians and other species, but possibly not via an EATS mechanism. 2) Strong evidence for *in vivo* endocrine activity in amphibians and other species, but possibly not via an EATS mechanism. 3) Strong evidence for adverse effects on growth or sexual development in amphibians and other species, but probably not via an endocrine mechanism.	It would be desirable to obtain some unequivocal mechanistic data to confirm whether or not an EATS mechanism is operating. Regulatory authorities may consider that further data from amphibians are not required. However, see column to the right.	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction is not currently available
H	+	Eq/0	-	1) Strong evidence for *in vivo* endocrine activity with adverse effects (on growth or sexual development) in amphibians, but possibly not via an EATS mechanism. 2) Strong evidence for *in vivo* endocrine	It would be desirable to obtain some unequivocal mechanistic data to confirm	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction

Scenarios	Result of LAGDA	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				activity in amphibians, but possibly not via an EATS mechanism. 3) Strong evidence for adverse effects on growth or sexual development in amphibians, but possibly not via an endocrine mechanism.	whether or not an EATS mechanism is operating. Regulatory authorities may consider that further data from amphibians are not required. However, see column to the right.	is not currently available
I	+	Eq/0	Eq/0	1) Strong evidence for *in vivo* endocrine activity with adverse effects (on growth or sexual development) in amphibians, but possibly not via an EATS mechanism. 2) Strong evidence for *in vivo* endocrine activity in amphibians, but possibly not via an EATS mechanism. 3) Strong evidence for adverse effects on growth or sexual development in amphibians, but possibly not via an endocrine mechanism.	It would be desirable to obtain some unequivocal mechanistic data to confirm whether or not an EATS mechanism is operating. Regulatory authorities may consider that further data from	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction is not currently available

Scenarios	Result of LAGDA	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
					amphibians are not required. However, see column to the right.	
J	-	+	+	The test chemical has EATS activity in other species but not apparently in amphibians, although it is possible that *Xenopus tropicalis* has responded atypically in this case (*e.g.* if *X. laevis* responded positively in OECD TG 231).	Regulatory authorities may consider that further data from amphibians are not required. However, see column to the right.	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction is not currently available
K	-	+	-	The test chemical has EATS activity *in vitro*, but no apparent activity *in vivo* in amphibians or other species, possibly due to metabolism or failure to reach the active site.	Regulatory authorities may consider that further testing is unnecessary.	-
L	-	+	Eq/0	The test chemical has EATS activity *in vitro*, but no apparent activity *in vivo* in amphibians, possibly due to metabolism or failure to reach the active site.	Regulatory authorities may consider that further testing is unnecessary, but	Given the presence of EATS activity *in vitro*, and the absence of reliable *in vivo* data from other species, it might be desirable to run an *in vivo* endocrine screen

Scenarios	Result of LAGDA	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
					see column to right	with fish or mammals
M	-	-	+	The test chemical does not apparently have EATS activity in amphibians, but endocrine activity is present in other species.	Regulatory authorities may consider that further data from amphibians are not required. However, see column to the right.	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction is not currently available
N	-	-	-	The test chemical does not have EATS activity in amphibians or other species.	No further action is necessary.	-
O	-	-	Eq/0	The test chemical does not have EATS activity in amphibians.	No further action is necessary.	-
P	-	Eq/0	+	The test chemical probably does not have EATS activity in amphibians, but the uncertain mechanistic data and the presence of endocrine activity in other species reduces confidence in this conclusion. It is possible that *Xenopus tropicalis* has responded atypically in this case (*e.g.* if *X. laevis* responded positively in OECD TG 231).	Regulatory authorities may consider that further data from amphibians are not required. However, see column to the right. Also, if clear *in*	The LAGDA does not cover the reproductive phase of the lifecycle, but a lifecycle test which could be used to address any concerns about reproduction is not currently available

Scenarios	Result of LAGDA	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive 3) Apical endpoint positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					vitro mechanistic data are missing, it might be desirable to obtain some.	
Q	-	Eq/0	-	The test chemical is probably without endocrine activity in amphibians or other taxa, but this conclusion is tentative given the lack of supporting mechanistic data.	If clear *in vitro* mechanistic data are missing, it might be desirable to obtain some.	-
R	-	Eq/0	Eq/0	The test chemical is probably without endocrine activity in amphibians, but this conclusion is tentative given the lack of supporting data.	Some regulatory authorities may conclude that no further evidence is required, but see column to right.	If clear *in vitro* mechanistic data are missing, it may be desirable to obtain some. If these data reveal EATS activity, it might then be desirable to conduct a fish or rodent screen.

Annex 2.4 Avian Two Generation Test (ATGT).

Annex 2.4.1 Modality detected/endpoints: It is unclear (either from validation data currently available, or from the draft avian 2-generation toxicity test TG) which endpoints respond specifically to which endocrine modalities, although estradiol treatment has been shown to cause changes in testis histology, feminisation of male plumage, altered female sexual maturation, and altered female cloacal gland size (Battelle, 2005)[7]. However, the list of indicators of hormonal activity considered in the draft TG to be responsive to some EATS EDs includes:

Weight of testes, thyroid, adrenals, oviduct, cloacal gland, liver
Histology of thyroid, adrenals, gonads, brain
Testicular spermatid counts and morphology
Gross anomalies of the genital tract
Feather dimorphism
Cloacal gland size, 1st appearance of foam
Time to first egg laid
Sexual behaviour
Faecal/urate steroid hormone titres (estradiol, testosterone)
Egg steroid content (estradiol, testosterone)
Tibiotarsus length (F1)

Annex 2.4.2 Further guidance on other possible indicators of hormonal activity (*e.g.* vitellogenin; thyroid hormones; gonadotropin releasing hormone) can be found in the Detailed Review Paper on the avian two-generation test (OECD, 2007b). In addition to the endocrine-specific endpoints, the test also measures several apical endpoints (egg production, embryo viability, fertility, hatchability, body weight, general toxic signs, mortality, shell thickness/cracking/strength, number of offspring, sex ratio of F1 offspring) which may respond both to EDs and to some non-EDs with reproductive toxicity.
Note that the ATGT is still being validated, so some of the advice given below may be subject to amendment.

Background to the Assay

Annex 2.4.3 This life cycle assay with the Japanese quail *Coturnix japonica* runs for 21 weeks, from 4 week old F0 reproducing adults to 2 week old F2 chicks, and hence encompasses more than one complete generation. It is therefore expected to be responsive to most chemicals with EATS modalities, although the full extent of its applicability awaits further validation. It should be noted that if the assay gives a positive result, this may be due to a positive indicator of hormonal activity (*i.e.* endocrine organ weights; endocrine organ gross pathology and histopathology; feather dimorphism; time to first egg lay; sexual behaviour; sex hormones; tibiotarsus length), a positive for apical endpoints (egg production; general health and toxic signs; sex ratio; body weight; shell thickness/cracking/strength; fertility; embryo viability; hatchability), or a positive for both types of endpoint. Each of these three possible combinations of positive response should be considered separately (although the distinctions between indicators of hormonal activity and apical effects are not always clear), so they have been listed individually as points 1, 2 and 3 in the Possible Conclusions column of Table Annex 2.4.

When/Why the Assay May be Used

[7] Note that a subsequent statistical re-evaluation of this report by the USEPA showed that some additional endpoints which were claimed to have responded to estradiol (male to female sex ratio; male sexual maturation) had not in fact given a statistically significant response.

Annex 2.4.4 Although the ATGT could, in principle, be used at any stage in the hazard assessment process, the probable use scenario will be when there are already *in vitro* or *in vivo* screening data available about the possible endocrine disrupting properties of a chemical. In addition, there may be data from the Avian Reproduction Test (OECD TG 206). It is highly unlikely that no other existing endocrine-relevant data will be available (*i.e.* if the ATGT has been used as a primary screen). Indeed, it is not recommended to perform this test in the absence of such data. However, in that case a positive result in the ATGT, especially if derived from an endocrine-sensitive apical endpoint, should ideally be followed up with relevant *in vitro* screening to confirm the suspected mode of action.

Existing Data to be Considered

Annex 2.4.5 Existing data available before deployment of the ATGT might include one or more of a range of *in silico* or *in vitro* results which suggest that EATS modalities may occur *in vivo,* plus *in vivo* results obtained with other vertebrates (*e.g.* a positive uterotrophic assay with rodents or a positive result in OECD TG 206). Indicators of possible *in vivo* activity might include (Q)SAR predictions of endocrine activity, 'read-across' from *in vivo* results obtained with chemically related chemicals, or positive results from an *in vitro* screen for ER or AR-mediated activity, or for effects on steroidogenesis (especially aromatase inhibition). Conduct of the ATGT would be particularly relevant if the test chemical has been shown to affect avian reproduction or growth in OECD TG 206. The ATGT will not only provide a reliable indication if the test chemical is an EATS ED, but should also allow calculation of a reliable long-term NOEC or ECx for use in environmental risk assessment.

Scenarios: Positive and Negative Results Combined with Existing Data

Annex 2.4.5a The scenarios (A to R) presented in Table Annex 2.4 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. Further considerations, specific to each scenario are given in the Table.

Annex 2.4.6 Positive results obtained with one or more of the ATGT indicators of hormonal activity but not with apical endpoints (Table Annex 2.4, Scenarios A-I, sub-section 2) result in the conclusion that the test chemical is a possible, but almost certainly not an actual, ED *in vivo*. If both an indicator of hormonal activity and an apical endpoint give a response (Table Annex 2.4, Scenarios A-I, sub-section 1), and there is a plausible causal relationship between them, this provides evidence that the chemical is almost certainly an actual ED (*i.e.* it causes adverse effects through an endocrine mechanism). If only an apical endpoint responds (Table Annex 2.4, Scenarios A-I, sub-section 3), it indicates that the chemical is a developmental, growth or reproductive toxicant, with a substantially reduced probability that it is an ED (although existing positive *in vitro* data, or positive *in vivo* data from other species, would have to be weighed against the latter conclusion).

Annex 2.4.7 As indicated above, an effect on an apical endpoint in the ATGT shows that the test chemical causes adverse effects and is a developmental, growth or reproductive toxicant (assuming that the concentration giving this response is not sufficiently high to cause systemic toxicity). If these results are combined with a positive indicator of hormonal activity and/or positive *in vitro* screening assay data, it is reasonable to conclude that the chemical is an actual ED, and that the information could be used in a risk assessment.

Annex 2.4.8 A situation in which the ATGT gives a negative result needs careful consideration of any existing data. It is unlikely that these data would suggest that the chemical is endocrine-active both *in vitro* and *in vivo* (Table Annex 2.4, Scenario J), but if they do, then the probability is that the species used in the ATGT is different from those already tested. Little is known about inter-species variability in response of the ATGT, although variations of it have been operated with the more slowly maturing bobwhite quail *Colinus virginianus*. However, a negative high-tier test like the ATGT would generally be considered to trump a positive lower-tier test. If positive *in vivo* data are available from OECD TG 206, they may be sufficient to permit categorization or risk assessment.

Annex 2.4.9 If the ATGT and existing *in vivo* data are all negative, but *in vitro* data reveal some endocrine activity (Scenario K), the probability is that the test chemical is not sufficiently potent to produce endocrine effects *in vivo* in birds, or it may be rapidly metabolised. In such a situation, further testing is probably not necessary as the test chemical is almost certainly not an ED.

Annex 2.4.10 Furthermore, if the ATGT and the *in vitro* tests are negative, but there are positive existing *in vivo* data (Scenario M), the chemical is almost certainly not an ED in birds, although it may have endocrine properties in other species.

Annex 2.4.11 Finally, a negative ATGT, set against a background of negative *in vitro* and *in vivo* data (Scenario N), suggests that the test chemical is almost certainly not an ED, and no further testing for estrogenic, anti-estrogenic, androgenic or steroidogenic MOAs should generally be considered.

Annex 2.4.12 In each of the above scenarios, it is possible that existing data will be equivocal, or there may be no existing data (Scenarios L, O, P, Q, R). This will weaken the conclusions which can be drawn about a negative ATGT, although it will remain likely that the test chemical is not an ED in birds. On the other hand, if the ATGT is positive, no further *in vivo* testing in birds is needed. Again, however, it will always be desirable to obtain some mechanistic information before reaching a conclusion that a positive apical result in the ATGT has been caused by an ED.

Annex 2.4.13 The scenarios in which the results of an ATGT are themselves equivocal have not been dealt with in Table Annex 2.3, for reasons of brevity. In this context, an equivocal result might be an inconsistent concentration-response (*e.g.* no effect at a high concentration but effects at a lower), or a result which borders on statistical significance. Without knowing the exact circumstances, reliable advice cannot be given, but the opinions of an experienced ecotoxicologist should be sought. Clearly, however, such equivocal results do not necessarily rule out the existence of *in vivo* endocrine activity. For example, an effect on egg production might just fail to reach a statistically significant level due to the inherently high variability of this endpoint. If this or other possible reasons for false negatives are suspected with good reason, the test could be repeated (*e.g.* with more birds per replicate). However, given that the scale of any equivocal effect is likely to be small, it may be considered that the high cost of a repeat test is not justified.

Annex 2.4.14 In summary, a positive apical response in the ATGT indicates that a chemical is a growth, developmental or reproductive toxicant which may or may not be an ED. A combination of a positive apical response and a positive endocrine-specific endpoint (*e.g.* feather dimorphism) is strong evidence that the chemical is an actual ED, especially if the two types of endpoint are causally related and if positive mechanistic data are also available. In this situation, further *in vivo* data from birds are unlikely to be required. On the other hand, negative results in the ATGT suggest that the test chemical is not an ED, at least not in Japanese quail – a judgement about the likelihood of endocrine effects in other species will have to be made in the light of existing *in vitro* and *in vivo* data.

Table Annex 2.4 Avian Two-Generation Test (ATGT). Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR, and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". There is no evidence at present that equivalent *in vitro* assays with systems derived from fish offer advantages over their mammalian counterparts.

Existing Results: ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter.

The assay under discussion could either be positive for both apical endpoints and indicators of hormonal activity, or positive just for apical endpoints, or positive just for indicators of hormonal activity. For each scenario, each of these 3 possibilities is addressed separately in the Possible Conclusions column.

Scenario	Result of ATGT	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**	1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive		

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+	+	+	1) Strong evidence for adverse effects in birds and other organisms by an endocrine mechanism. 2) Strong evidence for endocrine effects, but they do not appear adverse in birds. 3) Strong evidence for adverse effects in more than one organism, but mechanism may not be via direct interaction with ER or AR, or by aromatase inhibition or thyroid disruption.	Probably no need for additional data.	-
B	+	+	-	1) Strong evidence for adverse effects in birds by an endocrine	Probably no need for additional data.	-

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				mechanism. 2) Strong evidence for endocrine effects in birds, but they do not appear adverse. 3) Strong evidence for adverse effects in birds, but mechanism may not be via direct interaction with ER or AR, or by aromatase inhibition or thyroid disruption.		
C	+	+	Eq/0**	1) Strong evidence for adverse effects in birds by an endocrine mechanism. 2) Strong evidence for endocrine effects in	Probably no need for additional data.	-

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				birds, but they do not appear adverse. 3) Strong evidence for adverse effects in birds, but mechanism may not be via direct interaction with ER or AR, or by aromatase inhibition or thyroid disruption.		
D	+	-	+	1) Strong evidence for adverse effects in birds and other organisms, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in birds, but they do not appear to be adverse. 3) Strong evidence for	Probably no need for additional data, but see column to the right.	Negative *in vitro* mechanistic data combined with positive endocrine-specific *in vivo* endpoints suggests that an unknown endocrine MOA is causing any adverse effects. Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in this scenario (sub-section 1) is an ED.

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				adverse effects in more than one organism, but mechanism may not be by endocrine disruption.		
E	+	-	-	1) Strong evidence for adverse effects in birds, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in birds, but they do not appear to be adverse. 3) Strong evidence for adverse effects in birds, but mechanism may not be by endocrine disruption.	Probably no need for additional data, but see column to the right.	Negative *in vitro* mechanistic data combined with positive endocrine-specific *in vivo* endpoints suggests that an unknown endocrine MOA is causing any adverse effects. Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in this scenario (sub-section 1) is an ED, especially as existing *in vivo* data are negative.

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
F	+	-	Eq/0	1) Strong evidence for adverse effects in birds, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in birds, but they do not appear to be adverse. 3) Strong evidence for adverse effects in birds, but mechanism may not be by endocrine disruption.	Probably no need for additional data, but see column to the right.	Negative *in vitro* mechanistic data combined with positive endocrine-specific *in vivo* endpoints suggests that an unknown endocrine MOA is causing any adverse effects. Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in this scenario (sub-section 1) is an ED, especially as existing *in vivo* data are equivocal or absent. However, such a conclusion appears likely on the basis of the positive endocrine-sensitive endpoints *in vivo*.
G	+	Eq/0	+	1) Strong evidence for adverse effects in more than one organism, possibly by an unknown endocrine mechanism. 2) Medium-Strong	It would be desirable to obtain some clear mechanistic data before concluding that the chemical is an ED. See column to right.	Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in sub-section 1 is an ED. However, such a conclusion appears likely on the basis of the positive endocrine-sensitive endpoints *in vivo*.

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				evidence for endocrine effects, but they do not appear to be adverse in birds. 3) Strong evidence for adverse effects in more than one organism, but mechanism may not be by endocrine disruption.		
H	+	Eq/0	-	1) Strong evidence for adverse effects in birds, possibly by an unknown endocrine mechanism. 2) Medium-Strong evidence for endocrine effects in birds, but they do not appear to be adverse.	It would be desirable to obtain some clear mechanistic data before concluding whether the chemical is an ED. See column to right.	Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in sub-section 1 is an ED. However, such a conclusion appears likely on the basis of the positive endocrine-sensitive endpoints *in vivo*.

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				3) Strong evidence for adverse effects in birds, but mechanism may not be by endocrine disruption.		
I	+	Eq/0	Eq/0	1) Strong evidence for adverse effects in birds, possibly by an unknown endocrine mechanism. 2) Moderate-Strong evidence for endocrine effects in birds, but they do not appear to be adverse. 3) Strong evidence for adverse effects in birds, but mechanism may not be by endocrine disruption.	It would be desirable to obtain some clear mechanistic data before concluding whether the chemical is an ED. See column to right.	Some regulatory authorities may consider that the MOA needs further investigation before it can be concluded that the chemical in sub-section 1 is an ED. However, such a conclusion appears likely on the basis of the positive endocrine-sensitive endpoints *in vivo*.
J	-	+	+	The chemical is an ED	Regulatory authorities	The fact that the chemical has endocrine

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				in vivo in other species but does not appear to act on growth, sexual development or reproduction in birds. If any other bird tests are also negative, birds may not be responsive at all to the test chemical.	may consider that further evidence is not required.	properties *in vitro* and in other species *in vivo* suggests that it may be an ED, but probably not in birds. If the existing positive *in vivo* data are from a lower tier bird assay, note that it is generally considered that a negative higher tier test trumps a positive lower tier test.
K	-	+	-	Despite the *in vitro* mechanistic data for possible endocrine activity, there is no evidence for endocrine disruption *in vivo*. This may be because the chemical is d*egra*ded to an inactive metabolite,	Regulatory authorities may consider that further evidence is not required.	-

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				or because it only interacts very weakly with endocrine receptors.		
L	-	+	Eq/0	The chemical is not an ED in birds, but it may be active in other species as there is only one unequivocal *in vivo* test result (a negative).	Regulatory authorities may consider that further evidence is not required.	-
M	-	-	+	The chemical is probably not an ED acting on growth, sexual development or reproduction in birds, but it does have endocrine activity in other species. However, it may act through MOAs not covered by	Regulatory authorities may consider that further evidence is not required.	The fact that the chemical has endocrine properties in other species *in vivo* suggests that it may be an ED, but probably not in birds . If the existing positive *in vivo* data are from a lower tier bird assay, note that it is generally considered that a negative higher tier test trumps a positive lower tier test.

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				the available *in vitro* assays, or it may be more potent in a bird species other than that tested.		
N	-	-	-	The chemical is probably not an ED in birds or other species.	Regulatory authorities may consider that further evidence is not required.	-
O	-	-	Eq/0	The chemical is probably not an ED in birds.	Regulatory authorities may consider that further evidence is not required.	-
P	-	Eq/0	+	The chemical is probably not an ED acting on growth, sexual development or reproduction in birds,	Regulatory authorities may consider that further evidence is not required.	The fact that the chemical has endocrine properties in other species *in vivo* suggests that it may be an ED, but probably not in birds. If the existing positive *in vivo* data are from a lower tier bird assay, note that it is

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				but it does have endocrine activity in other species. However, it may act through MOAs not covered by the available *in vitro* assays, or it may be more potent in a bird species other than that tested.		generally considered that a negative higher tier test trumps a positive lower tier test.
Q	-	Eq/0	-	The chemical is probably not an ED acting on growth, sexual development or reproduction in birds, or *in vivo* on other species.	Regulatory authorities may consider that further evidence is not required, although negative *in vitro* data would strengthen the conclusion that the chemical is probably not an ED.	-
R	-	Eq/0	Eq/0	The chemical is	Regulatory authorities	-

Scenario	Result of ATGT	Existing Results		Possible conclusions 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				probably not an ED acting on growth, sexual development or reproduction in birds.	may consider that further evidence is not required, although negative *in vitro* data would strengthen the conclusion that the chemical is probably not an ED.	

Mammalian Screens and Tests

Annex 2.5 Adult Male Assay (no guideline)

Annex 2.5.1 Modalities detected: (Anti)-Androgen, thyroid, steroidogenesis.

Endpoints: Weight of testes, epididymides, prostate, seminal vesicles (+ coagulating gland), accessory sex glands (prostate + seminal vesicles + coagulating gland) thyroid. Histopathologic changes in testes, epididymides, thyroid.

Serum testosterone, dihydrotestosterone, estradiol, FSH, LH, prolactin, T4, T3 and TSH.

Background to the Assay

Annex 2.5.2 This assay is designed to identify chemicals that have the potential to interact with AR-mediated modalities, thyroid hormone mediated modalities and interference with steroidogenesis in rats. It will also detect chemicals that act directly or indirectly through changes in the HPG and HPT axes. It will also detect ER-mediated effects (USEPA, 2007d). The assay has similar endpoints and targets similar modalities to the male rat peripuberal (PP) assay and OECD TG 407 (28-day study). It has some advantages over the male PP assay *e.g.* duration of dosing is shorter (15 days compared to 31 days) and body weights of adult rats are more stable compared to peripubertal rats; but validation studies to date have indicated that it is less sensitive to weak androgenic compounds and the hormone results are rather variable.

Annex 2.5.3 The 15-day intact adult male rat assay was initially developed by the chemical industry to identify MOAs of chemicals (O'Connor *et al*, 2002a&b). Historically it has successfully identified a range of chemicals acting via EATS modalities. The adult male assay was originally proposed to be one of the suite of assays comprising US EPA's "Tier 1" but is currently not included as the male PP assay was more sensitive in the validation studies. Validation is, however, continuing and it is a potentially useful assay for the detection of EDs. Its strengths are the short dosing period, simple design, use of animals with mature HPG and HPT axes and multiple and complimentary male reproductive organs as primary endpoints with secondary hormonal endpoints.

When/Why the Assay May be Used

Annex 2.5.4 The adult male assay is likely to be used as a screen following positive results in ER, AR, TR and steroidogenesis disruption based assays in order to establish whether intrinsic interaction with these systems results in *in vivo* effects. It may also be used following positive results in UT and H assays (level 3) to establish whether effects in these models are also seen in animals with an intact HPG axis. The decision as to which level 4 assay to use (male PP, female PP, adult male or OECD TG 407) needs to be taken on a case-by-case basis. For example, the adult male assay may be used if a shorter dosing period is required.

Introduction to the Table of Scenarios

Annex 2.5.5 Table Annex 2.5 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

Annex 2.5.6 The results of the adult male assay are given in the second column. The assay contains multiple endpoints and it is not possible to provide alternative scenarios for all combinations, therefore some discrimination has been attempted by dividing the endpoints into "apical" and "indicators of hormonal activity". The terminology used has been chosen to be consistent between both the wildlife and mammalian tests. Both groups have similar biological importance, although the "indicators of hormonal activity" in the mammalian assays are serum hormones and are generally, but not always, more variable than "apical endpoints". "Apical endpoints" are weight of testes, epididymides, prostate, seminal vesicles, accessory sex glands, thyroid, histopathologic changes in testes, epididymides, thyroid. "Indicators of hormonal activity" are hormones (testosterone, dihydrotestosterone, estradiol, FSH, LH, prolactin, T4, T3 and TSH).

Three possible outcomes for a positive result are therefore envisaged in Table Annex 2.5:

1) Indicators of hormonal activity and apical endpoints positive
2) Indicators of hormonal activity positive and apical endpoints negative
3) Indicators of hormonal activity negative and apical endpoints positive

Annex 2.5.7 A positive result for apical endpoints could be statistically significant reductions in weights of the epididymides and prostate, accompanied by treatment-related histopathologic changes. A positive result for indicators of hormonal activity could be statistically significant changes in testosterone and LH. The multiple endpoints in this assay means that there is some redundancy in the assay but this is useful as not all chemicals may affect all endpoints associated with a mechanism of action and there may be site-specific differences in response.

Annex 2.5.8 Single isolated changes may be indicative of spurious results. The guidance on histopathologic changes in endocrine tests (OECD, 2009a) may be helpful in interpretation. Such results should be considered with caution although it is possible that these endpoints may have detected weak effects that were not detected by the apical endpoints in this study but may then be detected in longer-term studies.

Annex 2.5.9 A negative result for the adult male assay is taken to be absence of changes in both endocrine relevant indicators of hormonal activity and apical endpoints. In the absence of other pertinent lines of evidence, negative results in this test alone cannot be taken as evidence that the substance is not an ED. Further studies will be required as confirmation.

Annex 2.5.10 Equivocal results for the guideline are not considered in the table, partly for brevity but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative or whether the result must be put to one side and the test repeated (using the same or a different test). Factors which may have interfered with the result *e.g.* composition of the diet used, environmental influences; should be considered.

Existing Data to be Considered

Annex 2.5.11 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs *et al.* (2008) and an OECD Detailed Review Paper (OECD, 2008a). These methods, however, have not yet been validated.

Annex 2.5.12 Existing "Effects" data refer to *in vivo* effects that may come from H or UT assays. Another possibility is that repeat oral toxicity studies, reproduction/developmental toxicity screen tests or read across from analogues, may be available. It is unlikely that the adult male assay will be performed if data from robust higher tier reproductive studies are already available as it offers no advantage over these assays. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

Annex 2.5.13 When considering the results of the adult male assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

Annex 2.5.14 A series of scenarios (A to R) are presented in table Annex 2.5 and represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study. Further considerations, specific to each scenario are given in the Table.

Annex 2.5.15 Scenarios A to I represent positive results in the adult male assay in the presence of positive and/or negative *in vitro* mechanistic data and positive and/or negative *in vivo* effects data. Each positive adult male assay result scenario is divided into the three possible outcomes given above.

Annex 2.5.16 The possibilities of equivocal or missing existing data are given in scenarios C, F, G, H and I. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to taken on a case-by-case basis; for example, some equivocal data may be considered positive whilst in other cases no conclusions may be possible and therefore the situation is effectively "data not available.

Annex 2.5.17 Scenarios J to R represent negative results in the adult male assay in the presence of positive and/or negative *in vitro* mechanistic data and positive and/or negative *in vivo* effects data. As a negative result for the adult male assay is taken to be negative findings for both indicators of hormonal activity and apical endpoints then (unlike the situation with positive outcomes) there is only one possible negative outcome.

Annex 2.5.18 The possibilities of equivocal or missing existing data are given in scenarios L, O, P, Q and R.

Annex 2.5.19 The next step to take to increase evidence will depend upon the existing information. The table is meant to provide a succinct guide and may not cover all circumstances or possibilities.

Table Annex 2.5. Adult male assay. Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter. These may be repeated dose toxicity tests (*e.g.* OECD TG 407, OECD TG 408 28 and 90-day studies), UT and H assays or read across from chemical analogues.

***Note: three possible outcomes for a positive result are given:

1) Indicators of hormonal activity and apical endpoints positive

2) Indicators of hormonal activity positive and apical endpoints negative

3) Indicators of hormonal activity negative *and apical endpoints positive*

"Apical endpoints" are weight of testes, epididymides, prostate, seminal vesicles, accessory sex glands, thyroid, histopathologic changes in testes, epididymides, thyroid.

"Indicators of hormonal activity" are hormones (testosterone, dihydrotestosterone, estradiol, FSH, LH, prolactin, T4, T3 and TSH).

Scenarios	Result of adult male assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
A	+ ***	+	+	1) Increased evidence of (anti)-EATS activity. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. 3) Increased evidence of (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from level 5 then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Possible effects on E modality should also be considered. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
B	+	+	-	1) Increased evidence of (anti)-EATS activity. 2) Possible evidence of (anti)-EATS	Perform assay from level 5 *e.g.*ext-1 or 2-	Question why difference from existing data. Consider route of exposures and

Scenarios	Result of adult male assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				activity, apical endpoints may be less sensitive or unaffected.. 3) Increased evidence of (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected.	gen assay	possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Possible effects on E modality should also be considered.
C	+	+	Eq/0	1) Increased evidence of (anti)-EATS activity. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. 3) Increased evidence of (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	Check data on chemical analogues. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Possible effects on E modality should also be considered. Consider route of exposure for

Scenarios	Result of adult male assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						female PP assay and follow-up assay. Possible implications of ADME characteristics of the chemical.
D	+	-	+	1) Increased evidence of (anti)-EATS activity. Acts via non-ER, AR, TR, S mechanism or requires metabolic activation for activity. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via non-AR,TR, S mechanism or requires metabolic activation for activity. 3) Increased evidence of (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-AR,TR, S mechanism or requires metabolic activation for activity.	Perform *in vitro* ER, AR,TR, S assays with added metabolising system.	If existing data are from an adequate level 5 assay then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Possible effects on E modality should also be considered. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.

Scenarios	Result of adult male assay	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Indicators of hormonal activity and apical endpoints positive** **2) Indicators of hormonal activity positive and apical endpoints negative** **3) Indicators of hormonal activity negative and apical endpoints positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern)****			
						Further mechanistic studies would help determine MoA.
E	+	-	-	1) Increased evidence of (anti)-EATS activity. Acts via non-ER, AR, TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via non-AR, TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data. 3) Increased evidence of (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-AR, TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g*(ext-1 or 2-gen assay	Question why difference from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate other mechanisms.
F	+	-	Eq/0	1) Increased evidence of (anti)-EATS activity. Acts via non-ER, AR, TR, S mechanism or requires metabolic	Perform *in vitro* ER, AR, TR, S assays with	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical

Scenarios	Result of adult male assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				activation for activity. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via non-ER, AR,TR, S mechanism or requires metabolic activation for activity 3) Increased evidence of (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-ER, AR, TR, S mechanism or requires metabolic activation for activity.	added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	endpoints. Effects on apical endpoints alone may indicate other mechanisms. Check data on chemical analogues. Further mechanistic studies may help determine MoA. Upper level studies will provide hazard data.
G	+	Eq/0	+	1) Increased evidence of (anti)-EATS activity. May act via ER, AR, TR, S mechanism. (metabolic activation may be needed) 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. May act via ER, AR,TR, S mechanism (metabolic activation may be needed). 3) Increased evidence of (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. May	Perform *in vitro* ER, AR, TR, S assays. (for the "0" scenario, otherwise Eq result available) OR Perform *in vitro* ER, AR, TR, S assays with added metabolising	If existing data are from level 5 then there is sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or

Scenarios	Result of adult male assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				act via AR, TR, S mechanism. (metabolic activation needed)	system.	other mechanisms. Possible effects on E modality should also be considered. Check data on chemical analogues. Further mechanistic studies would help determine MoA. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
H	+	Eq/0	-	1) Increased evidence of (anti)-EATS activity. Acts via unknown mechanism or requires metabolic activation for activity. Route of exposure may account for the differences from existing data 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism. Weak activity does not result in adverse effects. 3) Increased evidence of (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via unknown mechanism. Weak activity does not result in adverse effects.	For the "0" scenario, perform *in vitro* ER, AR, TR, S assays with added metabolising system. (otherwise Eq result available)	Question why difference from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of subtle changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms.

Scenarios	Result of adult male assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						Possible effects on E modality should also be considered. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	1) Increased evidence of (anti)-EATS activity . Acts via unknown mechanism. Unknown potential for adverse effects. 2) Possible evidence of (anti)-EATS activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. 3) Increased evidence of (anti)-EATS activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. There may be a need for metabolic activation.	Perform *in vitro* ER, AR, TR, S assays. Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR perform assay from level 5 *e.g.*ext-1 or 2-gen assay	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate EATS modalities or other mechanisms. Possible effects on E modality should also be considered. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
J	-	+	+	No evidence for EATS activity in adult male assay. Metabolism or potency explains the	Perform *in vitro* ER, AR, TR, S assays with	If existing data are from an adequate level 5 assay then question why differences.

Scenarios	Result of adult male assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				difference from existing *in vitro* and *in vivo* data. Effects seen in existing studies are via non-EATS mechanism.	added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	If data are from H assay then this may be more sensitive than adult male assay. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical. Further mechanistic studies would help determine MoA.
K	-	+	-	No evidence for EATS activity in adult male assay. Metabolism or potency explains *in vitro*/*in vivo* differences.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*(ext)-1 or 2-gen assay	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less sensitive assay then a higher level test may be required. If data are from H assay then need to conduct higher tier assay to conclude absence of concern for

Scenarios	Result of adult male assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						endocrine disruption. Further mechanistic studies with metabolism may help determine MoA.
L	-	+	Eq/0	No evidence for EATS activity in adult male assay. Metabolism or potency explains *in vitro*/*in vivo* differences. Unknown potential for adverse effects.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.* ext-1 or 2-gen assay	Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* activity is not realised. Consider possible routes of exposure implications of metabolism.
M	-	-	+	No evidence for EATS activity in adult male assay. Effects seen in existing studies are via non-EATS mechanism.	Perform assay from level 5 *e.g.* ext-1 or 2-gen assay	If existing data are from an adequate level 5 assay then question why differences. If data are from H assay then this may be more sensitive than adult male assay. Effects seen in existing studies

Scenarios	Result of adult male assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical.
N	-	-	-	No evidence for EATS activity in adult male assay. No evidence for (anti)-EATS activity *in vitro*. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may perform assay from level 5 *e.g.*ext-1 or 2-gen assay	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If data are from H assay then need to conduct higher tier assay to conclude absence of concern for endocrine disruption.
O	-	-	Eq/0	No evidence for EATS activity in adult male assay. No evidence for (anti)-EATS activity *in vitro*. Unknown potential for adverse effects.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	Consider route of exposures and possible implications for ADME characteristics of the chemical in follow-up assay.
P	-	Eq/0	+	No evidence for EATS activity in adult male assay. Potential for adverse effects via unknown mechanism.	Perform *in vitro* ER, AR,TR, S assays	Consider route of exposure and possible implications for differences from existing assay. If data are from H assay then this

Scenarios	Result of adult male assay	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						may be more sensitive than adult male assay. Effects seen in existing studies may be in a more sensitive life stage.
Q	-	Eq/0	-	No evidence for EATS activity in adult male assay. No evidence of adverse effects.	Perform *in vitro* ER, AR,TR, S assays	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Further mechanistic studies would increase evidence.
R	-	Eq/0	Eq/0	No evidence for EATS activity in adult male assay.	Perform *in vitro* ER, AR,TR, S assays, otherwise Eq result available.	Further mechanistic studies would increase evidence. Check data on chemical analogues.

Annex 2.6. OECD TG 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents

Annex 2.6.1 Modalities detected: (Anti)-Estrogen, (anti)-androgen, thyroid, steroidogenesis.

Endpoints: Weight of adrenals, testes, epididymides, uterus, ovaries.

Histopathologic changes in pituitary, thyroid gland, gonads, uterus, accessory sex organs, female mammary gland, testes and adrenals.

Background to the Assay

Annex 2.6.2 This assay determines the general toxicity of chemicals in rodents after 90 days of oral dosing. It provides information on major toxic effects and target organ toxicity likely to arise from the post weaning period until well into adulthood. Although it has not been validated for the detection of EDs it contains many endpoints that are suitable for the determination of endocrine effects. A comparison can be made with validation of the OECD TG 407 (28-day oral toxicity study) for endocrine endpoints where substances that were moderate and strong EDs for (anti)-estrogenicity and (anti)-androgenicity (*e.g.* ethinylestradiol and flutamide) and weak and strong modulators of thyroid hormone-related effects (*e.g.* propylthiouracil and methyl testosterone) were detected (OECD, 2006b). Steroidogenesis inhibition was also detected although only one (potent) chemical was used in the validation study (CGS 18320B). The OECD TG 408 is likely to be more sensitive than the OECD TG 407 because of the extended dosing period and the larger number of animals per group (10 male and 10 female per group compared with 5 in OECD TG 407). The OECD TG 408, however, does not contain some sensitive endpoints (*e.g.* thyroid hormones, estrous cyclicity) that may be included in OECD TG 407.

When/Why the Assay May be Used

Annex 2.6.3 This assay is likely to be used as part of a pesticide submission package and forms part of the standard information requirements in certain chemical legislations (*e.g*. REACH for chemicals which are manufactured or imported in quantities of 100 tonnes or more). At least three dose levels are included so that an estimate of no-adverse-effect-level can be determined and the assay used for both hazard and risk assessment. It should be noted that, as this assay is not primarily designed to detect endocrine disruption, a higher degree of systemic toxicity is typically induced than is the case with the other level 3 and 4 assays. The possibly confounding effect of systemic toxicity on endocrine endpoints therefore needs to be considered.

Introduction to the Table of Scenarios

Annex 2.6.4 Table Annex 2.6 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. “Existing results” are subdivided into “Mechanism” and “Effects” data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

Annex 2.6.5 The results of OECD TG 408 are given in the second column. As OECD TG 408 is not a screening test where a yes/no (qualitative) answer is obtained for the test as a whole, positive results would generally be assessed for individual endpoints. For the purposes of this guidance, however, a positive result is defined as a biologically significant change in any of the endocrine endpoints listed above *e.g.* statistically significant reductions in reproductive organ weights. Changes in related endpoints will increase their biological significance *e.g.* changes in the weights of testes and epididymides accompanied by histopathological changes. A negative result for the OECD TG 408 is taken to be absence of biologically significant changes in all endocrine endpoints.

Annex 2.6.6 In the absence of other pertinent lines of evidence negative results in this test alone cannot be taken as evidence that the substance is not an ED. Further studies will be required as confirmation.

Annex 2.6.7 Equivocal results for the guideline are not considered in the table, partly for brevity but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative or whether the result must be put to one side and the test repeated (using the same or a different test guideline). Factors which may have interfered with the result *e.g.* composition of the diet used, environmental influences; should be considered.

Existing Data to be Considered

Annex 2.6.8 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available (OECD, 2008a) and an OECD Detailed Review Paper in Jacobs *et al.* (2008). These methods, however, have not yet been validated.

Annex 2.6.9 Existing "Effects" data refer to *in vivo* effects that may come from level 3 or 4 tests in the CF *e.g.* UT or H assays. In these cases, it should be remembered that these assays are specifically designed to be sensitive to EDs. It is unlikely that the OECD TG 408 will be performed if higher tier data are already available as the OECD TG 408 offers no advantage over these assays. As mentioned above, the results of the study may be interpreted as part of a battery or group of tests carried out for regulatory purposes. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

Annex 2.6.10 When considering the results of the OECD TG 408 assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

Annex 2.6.11 A series of scenarios (A to R) are presented in table Annex 2.6 and represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be

indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study. Further considerations, specific to each scenario are given in the Table.

Annex 2.6.12 Scenarios A to I represent positive results in OECD TG 408 in the presence of positive and/or negative *in vitro* mechanistic data and positive and/or negative *in vivo* effects data.

Annex 2.6.13 The possibilities of equivocal or missing existing data are given in scenarios C, F, G, H and I. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to taken on a case-by-case basis; for example, some equivocal data may be considered positive whilst in other cases no conclusions may be possible and therefore the situation is effectively "data not available

Annex 2.6.14 Scenarios J to R represent negative results in the OECD TG 408 in the presence of positive and/or negative *in vitro* mechanistic data and positive and/or negative *in vivo* effects data.

Annex 2.6.15 The possibilities of equivocal or missing existing data are given in scenarios L, O, P, Q and R.

Annex 2.6.16 The next step to take to increase evidence will depend upon the existing information. The table is meant to provide a succinct guide and may not cover all circumstances or possibilities

Table Annex 2.6 . OECD TG 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents. Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results ** "Effects (*in vivo* effects of concern)" assumes various information, such as data from repeat oral toxicity studies, reproduction/developmental toxicity screen tests, read across from analogues, will be available.

***Note: a positive result is defined as a biologically significant change in any of the endocrine endpoints.

Scenarios	Result of OECD TG 408 (rodent 90 d) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+ ***	+	+	(Anti)-EATS activity. Increased evidence of (anti)-EATS activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from level 5 there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Consider route of exposures for effects data and possible

Scenarios	Result of OECD TG 408 (rodent 90 d) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						implications of ADME characteristics of the chemical.
B	+	+	-	(Anti)-EATS activity. Increased evidence of (anti)-EATS activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	If existing data are from an adequate level 5 assay then question why differences. If existing data are from a less sensitive assay then a higher level test may be required. Consider route of exposures and possible implications of ADME characteristics of the chemical.
C	+	+	Eq/0	(Anti)-EATS activity. Increased evidence of (anti)-EATS activity.	Perform assay from level 5 *e.g.* ext-1 or 2-gen assay	Check data on chemical analogues. Consider route of exposure for OECD TG 408 and follow-up assay. Possible implications of ADME characteristics of the chemical.
D	+	-	+	(Anti)-EATS activity. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical. Further mechanistic studies would

Scenarios	Result of OECD TG 408 (rodent 90 d) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						help determine MoA.
E	+	-	-	(Anti)-EATS activity. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences between OECD TG408 and existing data	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	If existing data are from an adequate level 5 assay then question why differences. If existing data are from a less sensitive assay then a higher level test may be required. Consider route of exposures and possible implications of ADME characteristics of the chemical.
F	+	-	Eq/0	(Anti)-EATS activity. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Check data on chemical analogues. Further mechanistic studies may help determine MoA. Upper level studies will provide hazard data.
G	+	Eq/0	+	(Anti)-EATS activity. May act via ER, AR,TR, S mechanism. (metabolic activation needed)	Perform *in vitro* ER, AR, TR, S assays. (for the "0" scenario,	If existing data are from level 5 then may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen

Scenarios	Result of OECD TG 408 (rodent 90 d) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					otherwise Eq result available) OR Perform *in vitro* ER, AR, TR, S assays with added metabolising system.	assay provides the most information). Check data on chemical analogues. Further mechanistic studies would help determine MoA. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
H	+	Eq/0	-	(Anti)-EATS activity. Acts via unknown mechanism or requires metabolic activation for activity. Route of exposure may account for the differences between OECD TG408 and existing data	For the "0" scenario, perform *in vitro* ER, AR, TR, S assays with added metabolising system. (otherwise Eq result available)	If existing data are from an adequate level 5 assay then question why differences. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	(Anti)-EATS activity. Acts via unknown mechanism. Unknown potential for adverse effects.	Perform *in vitro* ER, AR, TR, S assays. Perform *in vitro* ER, AR, TR, S	Check data on chemical analogues. Further mechanistic studies would help determine MoA.

Scenarios	Result of OECD TG 408 (rodent 90 d) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					assays with added metabolising system OR perform assay from level 5 *e.g.*ext-1 or 2-gen assay	
J	-	+	+	No evidence for (anti)-EATS activity in OECD TG408. Weak (anti)-EAS activity may not be detected by this assay. Metabolism or potency explains the difference from existing *in vitro*/ and *in vivo* data. Effects seen in existing studies are via non-EATS mechanism.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay then question why differences. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical. Further mechanistic studies would help determine MoA.
K	-	+	-	No evidence for (anti)-EATS activity in OECD TG408. Weak (anti)-EAS activity may not be detected by this assay. Metabolism or potency explains *in vitro*/*in vivo* differences.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less

Scenarios	Result of OECD TG 408 (rodent 90 d) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					from level 5 *e.g.*ext-1 or 2-gen assay.	sensitive assay then a higher level test may be required. Further mechanistic studies with metabolism may help determine MoA.
L	-	+	Eq/0	No evidence for (anti)-EATS activity in OECD TG408. Weak (anti)-EAS activity may not be detected by this assay. Metabolism or potency explains *in vitro*/*in vivo* differences. Unknown potential for adverse effects.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* activity is not realised. Consider possible routes of exposure implications of metabolism.
M	-	-	+	No evidence for (anti)-EATS activity in OECD TG408. Weak (anti)-EAS activity may not be detected by this assay. Effects seen in existing studies are via non-EATS mechanism.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay then question why differences. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical.
N	-	-	-	No evidence for (anti)-EATS activity in OECD TG408. Weak (anti)-EAS activity may not be detected by this assay.	Perform assay from level 5 *e.g.*ext-1 or 2-gen	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of

Scenarios	Result of OECD TG 408 (rodent 90 d) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				No evidence for (anti)-EATS activity *in vitro*. No evidence of adverse effects.	assay.	concern for endocrine disruption (the ext-1 gen assay provides the most information).
O	-	-	Eq/0	No evidence for (anti)-EATS activity in OECD TG408. Weak (anti)-EAS activity may not be detected by this assay. No evidence for (anti)-EATS activity *in vitro*. Unknown potential for adverse effects.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Consider route of exposures and possible implications for ADME characteristics of the chemical in follow-up assay.
P	-	Eq/0	+	No evidence for (anti)-EATS activity in OECD TG408. Weak (anti)-EAS activity may not be detected by this assay. Potential for adverse effects via unknown mechanism.	Perform *in vitro* ER, AR,TR, S assays	Consider route of exposure for OECD TG 408 assay and possible implications for differences from existing assay. Effects seen in existing studies may be in a more sensitive life stage. Further mechanistic studies would increase evidence.
Q	-	Eq/0	-	No evidence for (anti)-EATS activity in OECD TG408. Weak (anti)-EAS activity may not be detected by this assay. No evidence of adverse effects.	Perform *in vitro* ER, AR,TR, S assays.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Further mechanistic studies would increase evidence.
R	-	Eq/0	Eq/0	No evidence for (anti)-EATS activity in OECD TG408. Weak (anti)-EAS activity may	Perform *in vitro* ER, AR,TR, S	Further mechanistic studies would increase evidence.

Scenarios	Result of OECD TG 408 (rodent 90 d) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				not be detected by this assay.	assays, otherwise Eq result available.	Check data on chemical analogues.

Annex 2.7 OECD TG 451-3: Combined Chronic Toxicity/Carcinogenicity Studies

Annex 2.7.1 Modalities detected: (Anti)-Estrogen, (anti)-androgen, thyroid, steroidogenesis.

Endpoints: Weight of adrenals, epididymides, ovaries, testes, thyroid, uterus (chronic toxicity studies).

Histopathologic changes in adrenals, cervix, coagulating gland, epididymides, mammary glands, ovaries, pituitary, prostate, seminal vesicles, testes, thyroid gland, uterus.

Background to the Assay

Annex 2.7.2 These assays determine the general toxicity (OECD TG 452 and OECD TG 453) and carcinogenicity (OECD TG 451 and OECD TG 453) of chemicals in laboratory animals after exposure for a period lasting most of the lifespan. OECD TG 453 was revised in 2009 and replaced OECD TG 451 (older studies may have used OECD TG 451). General toxicity studies usually have a duration of 12 months whilst carcinogenicity studies usually have a duration of 18 or 24 months. They provide information on major toxic effects, target organ toxicity and carcinogenicity. Although they have not been validated for the detection of EDs and do not appear in the original CF, they contain many endpoints that are suitable for the determination of endocrine effects. Organ weights are not always included in the carcinogenicity phases of these studies as neoplastic changes may confound them but they are generally determined at 12 months. A comparison can be made with validation of the OECD TG 407 (28-day oral toxicity study) for endocrine endpoints (OECD, 2006b) where substances that were moderate and strong EDs for (anti)-estrogenicity and (anti)-androgenicity (*e.g.* ethinylestradiol and flutamide) and weak and strong modulators of thyroid hormone-related effects (*e.g.* propylthiouracil and methyl testosterone) were detected. Steroidogenesis inhibition was also detected although only one (potent) chemical was used in the validation study (CGS 18320B). OECD TG 453 and 452 are likely to be more sensitive than the OECD TG 407 because of the extended dosing period and the larger number of animals per group (20 or 50 rodents per sex per group for chronic or carcinogenicity studies respectively compared with 5 in OECD TG 407). OECD TG 453 and 452, however, do not contain some sensitive endpoints (*e.g.* thyroid hormones, estrous cyclicity) that may be included in OECD TG 407.

When/Why the Assay May be Used

Annex 2.7.3 These assays are likely to be used as part of a pesticide submission package and forms part of the standard information requirements in certain chemical legislations (*e.g.* REACH for chemicals which are manufactured or imported in quantities of 1000 tonnes or more). At least three dose levels are included so that an estimate of no-adverse-effect-levels or point of departure for benchmark doses can be determined. The assays are used for both hazard and risk assessment. It should be noted that, as these assays are not primarily designed to detect endocrine disruption, a higher degree of systemic toxicity is typically induced than is the case with the other level 3 and 4 assays. The possibly confounding effect of systemic toxicity on endocrine endpoints therefore needs to be considered.

Introduction to the Table of Scenarios

Annex 2.7.4 Table Annex 2.5 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

Annex 2.7.5 The results of OECD TG 451-3 are given in the second column. As they are not screening tests where a yes/no (qualitative) answer is obtained for the test as a whole, positive results would generally be assessed for individual endpoints. For the purposes of this guidance, however, a positive result is defined as a biologically significant change in any of the endocrine endpoints listed above *e.g.* statistically significant reductions in reproductive organ weights. Changes in related endpoints will increase their biological significance *e.g.* changes in the weights of testes and epididymides accompanied by histopathological/neoplastic changes. A negative result for OECD TG 452/ 453 is taken to be absence of biologically significant changes in both endocrine endpoints.

Annex 2.7.6 In the absence of other pertinent lines of evidence negative results in this test alone cannot be taken as evidence that the substance is not an ED. Further studies will be required as confirmation.

Annex 2.7.7 Equivocal results for the guideline are not considered in the table, partly for brevity but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative, whether the result must be put to one side or whether further testing should be carried out. Factors which may have interfered with the result *e.g.* composition of the diet used, environmental influences; should be considered.

Existing Data to be Considered

Annex 2.7.8 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs *et al.* (2008) and an OECD Detailed Review Paper (OECD, 2008a). These methods, however, have not yet been validated.

Annex 2.7.9 Existing "Effects" data refer to *in vivo* effects that may come from level 3 or 4 tests in the CF *e.g.* UT or H assays; or other sub-chronic repeat dosing studies *e.g.* OECD TG 407 (28-day) or OECD TG 408 (90-day). In these cases, it should be remembered that these assays are specifically designed to be sensitive to EDs. As mentioned above, the results of these studies may be interpreted as part of a battery or group of tests carried out for regulatory purposes. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

Annex 2.7.10 When considering the results of OECD TG 451-3 assays, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

Annex 2.7.11 A series of scenarios (A to R) are presented in table Annex 2.7 and represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study. Further considerations, specific to each scenario are given in the Table.

Annex 2.7.12 Scenarios A to I represent positive results in OECD TG 451-3 in the presence of positive and/or negative *in vitro* mechanistic data and positive and/or negative *in vivo* effects data.

Annex 2.7.13 The possibilities of equivocal or missing existing data are given in scenarios C, F, G, H and I. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to taken on a case-by-case basis; for example, some equivocal data may be considered positive whilst in other cases no conclusions may be possible and therefore the situation is effectively "data not available

Annex 2.7.14 Scenarios J to R represent negative results in OECD TG 451-3 in the presence of positive and/or negative *in vitro* mechanistic data and positive and/or negative *in vivo* effects data.

Annex 2.7.15 The possibilities of equivocal or missing existing data are given in scenarios L, O, P, Q and R.

Annex 2.7.16 The next step to take to increase evidence will depend upon the existing information. The table is meant to provide a succinct guide and may not cover all circumstances or possibilities

.

Table 2.7 Annex OECD TG 451-3: Combined Chronic Toxicity/Carcinogenicity Studies. Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results ** "Effects (*in vivo* effects of concern)" assumes various information, such as data from repeat oral toxicity studies, reproduction/developmental toxicity screen tests, read across from analogues, will be available.

***Note: a positive result is defined as a biologically significant change in any of the endocrine endpoints.

Scenarios	Result of OECD TG 451-3 (carc/1 year) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+ ***	+	+	(Anti)-EATS activity. Increased evidence of (anti)-EATS activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from level 5 there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.

Scenarios	Result of OECD TG 451-3 (carc/1 year) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
B	+	+	-	(Anti)-EATS activity. Increased evidence of (anti)-EATS activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	If existing data are from an adequate level 5 study then question why differences. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required.
C	+	+	Eq/0	(Anti)-EATS activity. Increased evidence of (anti)-EATS activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	Check data on chemical analogues. Consider route of exposure for OECD TG452/453 and follow-up assay. Possible implications of ADME characteristics of the chemical.
D	+	-	+	(Anti)-EATS activity. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system.	If existing data are from an adequate level 5 assay there may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.

Scenarios	Result of OECD TG 451-3 (carc/1 year) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
						Further mechanistic studies would help determine MoA.
E	+	-	-	(Anti)-EATS activity. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity. Route of exposure may account for the differences between OECD TG451-3and existing data	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay	If existing data are from an adequate level 5 assay then question why differences. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required.
F	+	-	Eq/0	(Anti)-EATS activity. Acts via non-ER,AR,TR, S mechanism or requires metabolic activation for activity.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Check data on chemical analogues. Further mechanistic studies may help determine MoA. Upper level studies will provide more hazard data.
G	+	Eq/0	+	(Anti)-EATS activity. May act via ER, AR,TR, S mechanism. (metabolic activation needed)	Perform *in vitro* ER, AR, TR, S assays. (for the "0" scenario,	If existing data are from level 5 then may be sufficient information to conclude evidence of concern for endocrine disruption (the ext-1 gen

Scenarios	Result of OECD TG 451-3 (carc/1 year) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					otherwise Eq result available) OR Perform *in vitro* ER, AR, TR, S assays with added metabolising system.	assay provides the most information). Check data on chemical analogues. Further mechanistic studies would help determine MoA. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
H	+	Eq/0	-	(Anti)-EATS activity. Acts via unknown mechanism or requires metabolic activation for activity. Route of exposure may account for the differences between OECD TG451-3and existing data	For the "0" scenario, perform *in vitro* ER, AR, TR, S assays with added metabolising system. (otherwise Eq result available)	If existing data are from an adequate level 5 assay then question why differences. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay then a higher level test may be required. Check data on chemical analogues. Further mechanistic studies would help determine MoA.
I	+	Eq/0	Eq/0	(Anti)-EATS activity. Acts via unknown mechanism. Unknown potential for adverse effects.	Perform *in vitro* ER, AR, TR, S assays. Perform *in vitro*	Check data on chemical analogues. Further mechanistic studies would help determine MoA.

Scenarios	Result of OECD TG 451-3 (carc/1 year) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					ER, AR, TR, S assays with added metabolising system OR perform assay from level 5 *e.g.*ext-1 or 2-gen assay	
J	-	+	+	No evidence for (anti)-EATS activity in OECD TG451-3. Effects on reproduction will not be detected by these assays. Metabolism or potency explains the difference from existing *in vitro*/ and *in vivo* data. Effects seen in existing studies are via non-EATS mechanism.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay then question why differences. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical. Further mechanistic studies would help determine MoA.
K	-	+	-	No evidence for (anti)-EATS activity in OECD TG451-3. Effects on reproduction will not be detected by these assays. Metabolism or potency explains *in vitro*/*in vivo* differences.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). If existing data are from a less

Scenarios	Result of OECD TG 451-3 (carc/1 year) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					from level 5 *e.g.*ext-1 or 2-gen assay.	sensitive assay then a higher level test may be required. Further mechanistic studies with metabolism may help determine MoA.
L	-	+	Eq/0	No evidence for (anti)-EATS activity in OECD TG451-3. Effects on reproduction will not be detected by these assays. Metabolism or potency explains *in vitro*/*in vivo* differences. Unknown potential for adverse effects.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* activity is not realised. Consider possible routes of exposure implications of metabolism.
M	-	-	+	No evidence for (anti)-EATS activity in OECD TG451-3. Effects on reproduction will not be detected by these assays.. Effects seen in existing studies are via non-EATS mechanism.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay then question why differences. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical.
N	-	-	-	No evidence for (anti)-EATS activity in	Perform assay	If existing data are from an adequate

Scenarios	Result of OECD TG 451-3 (carc/1 year) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
				OECD TG451-3. Weak (anti)-EAS activity may not be detected by this assay. No evidence for (anti)-EATS activity *in vitro*. No evidence of adverse effects.	from level 5 *e.g.*ext-1 or 2-gen assay.	level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information).
O	-	-	Eq/0	No evidence for (anti)-EATS activity in OECD TG451-3. Effects on reproduction will not be detected by these assays. No evidence for (anti)-EATS activity *in vitro*. Unknown potential for adverse effects.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Consider route of exposures and possible implications for ADME characteristics of the chemical in follow-up assay.
P	-	Eq/0	+	No evidence for (anti)-EATS activity in OECD TG451-3. Effects on reproduction will not be detected by these assays. Potential for adverse effects via unknown mechanism.	Perform *in vitro* ER, AR,TR, S assays	Consider route of exposure for OECD TG452/453 and possible implications for differences from existing assay. Effects seen in existing studies may be in a more sensitive life stage.
Q	-	Eq/0	-	No evidence for (anti)-EATS activity in OECD TG451-3. Effects on reproduction will not be detected by these assays.. No evidence of adverse effects.	Perform *in vitro* ER, AR,TR, S assays	If existing data are from an adequate level 5 assay there may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Further mechanistic studies would increase evidence.
R	-	Eq/0	Eq/0	No evidence for (anti)-EATS activity in OECD TG451-3. Effects on reproduction will not be detected by these assays.	Perform *in vitro* ER, AR,TR, S assays, otherwise	Further mechanistic studies would increase evidence. Check data on chemical analogues.

Scenarios	Result of OECD TG 451-3 (carc/1 year) assay	Existing Results		Possible conclusions	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
					Eq result available.	

Annex 2.8 OECD TG 421 Reproduction/Developmental Toxicity Screening Test and OECD TG 422 Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test

Annex 2.8.1 Modalities detected: (Anti)-Estrogen, (anti)-androgen, steroidogenesis (OECD TG 421 & 422) thyroid (OECD TG 422).

Endpoints: (for OECD TG 422, note that endpoints for OECD TG 421 are not as extensive)

Time to mating, male fertility, female fertility, dystocia, gestation length, number of implantations & corpora lutea, number of live births and post implantation loss, litter size, sex ratio, litter/pup weight, pup survival index

Weights of: (Parents only) testes, epididymides (OECD TG 421 & 422), plus adrenals (OECD TG 422 only)

Histopathologic changes in: (Parents only) testis, epididymides, accessory sex organs, ovaries (OECD TG 421 & 422), plus uterus, adrenals, thyroid (OECD TG 422 only).

Background to the Assay

Annex 2.8.2 These assays are designed to provide limited information about the effects of a chemical on the male and female reproductive systems including gonadal function, mating, conception, gestation, development of the conceptus and parturition. Although the titles of the TGs imply that they are screening tests, they are not screens as given in the definition in Section A but are apical assays. The TGs have similar experimental schedules but OECD TG 422 includes a more detailed assessment of repeated dose toxicity and thus more endpoints. The studies are not designed to detect EDs but they have some endpoints relevant for the assessment of possible endocrine disruption and provide data on adverse effects related to reproduction and development. The developing fetus is a life stage that has been shown to be particularly sensitive to EDs. The scope of these assays is much smaller than OECD TG 415 and OECD TG 416 (one and two generation assays) *e.g.* duration of premating exposure is much shorter, group sizes are generally half and post-natal development is not included.

Annex 2.8.3 Although they have not been validated for the detection of EDs, OECD TG 421 and 422 contain endpoints that are suitable for the determination of endocrine effects. In addition to reproduction/development, OECD TG 421 and 422 may both provide information about endocrine effects on male reproductive organs. OECD TG 422 will also include information about effects on the thyroid. Female reproductive organs are also examined but detection of endocrine effects in these organs may be obscured because of pregnancy. Male animals are dosed for a total period of 28 days. A comparison can be made with OECD TG 407 (28-day oral toxicity study) where validation studies (OECD, 2006b) demonstrated that substances that were moderate and strong EDs for (anti)-estrogenicity and (anti)-androgenicity (*e.g.* ethinylestradiol and flutamide) and weak and strong modulators of thyroid hormone-related effects (*e.g.* propylthiouracil and methyl testosterone) were detected. Steroidogenesis inhibition was also detected although only one (potent) chemical was used in the validation study (CGS 18320B). As all the endpoints are apical, it is difficult to discern mechanism of action from these tests alone. Information on mechanism of action needs to be obtained from *in vitro* EATS assays or *in vivo* lower tier tests such as UT and H assays.

When/Why the Assay May be Used

Annex 2.8.4 These assays are frequently used for initial hazard assessments for chemicals, as part of the Screening Information Data Set (SIDS) for the assessment of chemicals for which there is little information or for dose setting for more extensive reproduction/developmental assays.

Introduction to the Table of Scenarios

Annex 2.8.5 Table Annex 2.8 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "Mechanism" and "Effects" data (3rd and 4th columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

Annex 2.8.6 The results of OECD TG 421/422 are given in the second column. As these are not tests where a yes/no (qualitative) answer is obtained for the test as a whole, positive results would generally be assessed for individual endpoints and as a whole. It is not possible to provide guidance on all endpoints individually and therefore the endpoints have been pragmatically divided into "parental endocrine organ" endpoints *i.e*. effects on the endocrine organs of the parental animals determined during the repeated dose toxicity phase of the studies and "reproductive/developmental" endpoints determined during the reproduction/ conceptus development phase. OECD TG422 has a more extensive repeated dose toxicity phase and therefore more endocrine endpoints available for assessment, but both TGs are still useful.

Annex 2.8.7 "Parental endocrine organ" endpoints are weights of testes, epididymides, and adrenals; histopathologic changes in testis, epididymides, accessory sex organs, ovaries, uterus, adrenals, thyroid.

"Reproductive/developmental" endpoints are: Time to mating, male fertility, female fertility, gestation length, number of implantations & corpora lutea, number of live births and post implantation loss, litter size, sex ratio, litter/pup weight, pup survival index

Annex 2.8.8 Three possible outcomes for a positive result are therefore envisaged in Table Annex 2.8:

1) Reproductive/developmental endpoints positive and parental endocrine organ endpoints positive.

2) Reproductive/developmental endpoints only positive.

3) Parental endocrine organ endpoints only positive.

Annex 2.8.9 A positive endpoint is defined as a biologically significant change, *e.g.* statistically significant reductions in reproductive organ weights. Changes in related endpoints will increase their biological significance *e.g.* changes in the weights of testes and epididymides accompanied by reduced male fertility.

Annex 2.8.10 A negative result for the OECD TG 421/422 is taken to be absence of biologically significant changes in all endocrine endpoints.

In the absence of other pertinent lines of evidence negative results in this test alone cannot be taken as firm evidence that the substance is not an ED. Further studies may be required as confirmation but will depend on the usage of the chemicals and potential for exposure.

Annex 2.8.11 Equivocal results for the guideline are not considered in the table, partly for brevity but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative or whether the result must be put to one side and the test repeated (using the same or a different test guideline). Factors which may have interfered with the result *e.g.* composition of the diet used, environmental influences; should be considered.

Existing Data to be Considered

Annex 2.8.12 Existing "Mechanism" *in vitro* data are assumed to be available from ER, AR and steroidogenesis based assays (level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available and so therefore judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs *et al.* (2008) and an OECD Detailed Review Paper (OECD, 2008a). These methods, however, have not yet been validated.

Annex 2.8.13 Existing "Effects" data refer to *in vivo* effects that may come from level 3 or 4 tests in the CF *e.g.* UT or H assays. In these cases, it should be remembered that these assays are specifically designed to be sensitive to EDs. Given the usage of these assays for general chemical testing, it is possible that a OECD TG 407 (28-day test) is available. It is unlikely that OECD TG 421/422 will be performed if higher tier reproduction/developmental toxicity data are already available as they offer no advantage over these assays. The results of the study may also be interpreted as part of a battery or group of tests carried out for regulatory purposes. Data may also be available on effects in wildlife species although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in environmental species may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

Annex 2.8.14 When considering the results of the OECD TG 421/422, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include read-across data from structural analogues and QSAR.

Scenarios: Positive and Negative Results Combined with Existing Data

Annex 2.8.15 A series of scenarios (A to R) are presented in table Annex 2.8 and represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend upon the regulatory environment but the considerations given here are generally science based. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain and exposure route should always be considered. In general lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage unless it is apparent that a Level 5 test will be required anyway. At Level 5, the extended one-generation study (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study. Further considerations, specific to each scenario are given in the Table.

Annex 2.8.16 Scenarios A to I represent positive results in OECD TG 421/422 in the presence of positive and/or negative *in vitro* mechanistic data and positive and/or negative *in vivo* effects data. Each positive OECD TG 421/422 result scenario is divided into the three possible outcomes given above.

Annex 2.8.17 The possibilities of equivocal or missing existing data are given in scenarios C, F, G, H and I. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases but the nature of equivocal data means that decisions need to taken on a case-by-case basis; for example, some equivocal data may be considered positive whilst in other cases no conclusions may be possible and therefore the situation is effectively "data not available.

Annex 2.8.18 Scenarios J to R represent negative results in OECD TG 421/422 in the presence of positive and/or negative *in vitro* mechanistic data and positive and/or negative *in vivo* effects data. As a negative result for the OECD TG 421/422 is taken to be negative findings for both "parental endocrine organ" endpoints and "reproductive/developmental" endpoints then (unlike the situation with positive outcomes) there is only one possible negative outcome.

Annex 2.8.19 The possibilities of equivocal or missing existing data are given in scenarios L, O, P, Q and R.

Annex 2.8.20 The next step to take to increase evidence will depend upon the existing information. The table is meant to provide a succinct guide and may not cover all circumstances or possibilities.

Table Annex 2.8 OECD TG 421 Reproduction/Developmental Toxicity Screening Test and OECD TG 422 Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test. Guidance for scenarios of combinations of results with existing data.

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data, and existing *in vivo* data. The symbol '+' indicates that the data in question represent a positive result, '-' indicates a negative result, and 'Eq/0' indicates that the data are either equivocal or are not available.

Existing Results * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from ER, AR and steroidogenesis based assays (level 2). TR and other assays concerning mechanisms of thyroid disruption may be available but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step".

Existing Results ** "Effects (*in vivo* effects of concern)" assumes effects have been observed in other *in vivo* screens/tests which give rise to concern that the test chemical may be an endocrine disrupter. These may be other repeated dose toxicity tests, UT and H assays or read across from chemical analogues..

***Note: three possible outcomes for a positive result are given:

1 Reproductive/developmental endpoints and parental endocrine organs positive
2) Reproductive/developmental endpoints positive
3) Parental endocrine organs positive

Scenarios	Result of OECD TG 421/422	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Reproductive/developmental endpoints and parental endocrine organs positive** **2) Reproductive/developmental endpoints positive** **3) Parental endocrine organs positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern) ****			
A	+ *****	+	+	1) Evidence of adverse effects on endocrine/ reproductive/developmental endpoints via EATS mechanism. 2) Evidence of adverse effects on reproductive/developmental endpoints via EATS mechanism / 3) Evidence of adverse effects on endocrine organs via EATS mechanism but without reproductive/developmental effects.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	May be sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Consider potency of effects for existing results and whether EATS mechanism is credible for reproductive/developmental effects or whether there may be non-endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
B	+	+	-	1) Evidence of adverse effects on endocrine/ reproductive/developmental endpoints via EATS mechanism. 2) Evidence of adverse effects on reproductive/developmental endpoints via EATS mechanism.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	May be sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine

Scenarios	Result of OECD TG 421/422	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Reproductive/developmental endpoints and parental endocrine organs positive 2) Reproductive/developmental endpoints positive 3) Parental endocrine organs positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern) **			
				3) Evidence of adverse effects on endocrine organs via EATS mechanism but without reproductive/developmental effects.		disruption. Consider potency of effects for existing results and whether EATS mechanism is credible for reproductive/developmental effects or whether there may be non- endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
C	+	+	Eq/0	1) Evidence of adverse effects on endocrine/ reproductive/developmental endpoints via EATS mechanism. 2) Evidence of adverse effects on reproductive/developmental endpoints via EATS mechanism. 3) Evidence of adverse effects on endocrine organs via EATS mechanism but without reproductive/developmental effects .	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Consider potency of effects for existing results and whether EATS mechanism is credible for reproductive/developmental effects or whether there may be non- endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
D	+	-	+	1) Evidence of adverse effects on endocrine/ reproductive/developmental endpoints but not via EATS mechanism or requires metabolic activation for activity. 2) Evidence of adverse effects on	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	May be sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Note that the ext-1 gen assay provides

Scenarios	Result of OECD TG 421/422	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Reproductive/developmental endpoints and parental endocrine organs positive** **2) Reproductive/developmental endpoints positive** **3) Parental endocrine organs positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern) ****			
				reproductive/developmental endpoints but not via EATS mechanism or requires metabolic activation for activity. 3) Evidence of adverse effects on endocrine organs but without reproductive/developmental effects and not via EATS mechanism or requires metabolic activation for activity.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	the most information on endocrine disruption. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non- endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
E	+	-	-	1) Evidence of adverse effects on endocrine/ reproductive/developmental endpoints but not via EATS mechanism or requires metabolic activation for activity. 2) Evidence of adverse effects on reproductive/developmental endpoints but not via EATS mechanism or requires metabolic activation for activity. 3) Evidence of adverse effects on endocrine organs but without reproductive/developmental effects and not via EATS mechanism or requires metabolic	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added	May be sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non- endocrine

Scenarios	Result of OECD TG 421/422	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Reproductive/developmental endpoints and parental endocrine organs positive 2) Reproductive/developmental endpoints positive 3) Parental endocrine organs positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern) **			
				activation for activity.	metabolising system.	mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
F	+	-	Eq/0	1) Evidence of adverse effects on endocrine/ reproductive/developmental endpoints via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. 2) Evidence of adverse effects on reproductive/developmental endpoints via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. Apical endpoints may be less sensitive or unaffected. 3) Evidence of adverse effects on endocrine organs but without reproductive/developmental effects, via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	May be sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non- endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
G	+	Eq/0	+	1) Evidence of adverse effects on endocrine/ reproductive/developmental endpoints, may act via EATS mechanism and may require metabolic activation for activity.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	May be sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism.

Scenarios	Result of OECD TG 421/422	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Reproductive/developmental endpoints and parental endocrine organs positive** **2) Reproductive/developmental endpoints positive** **3) Parental endocrine organs positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern) ****			
				2) Evidence of adverse effects on reproductive/developmental endpoints, may act via EATS mechanism and may require metabolic activation for activity. 3) Evidence of adverse effects on endocrine organs but without reproductive/developmental effects may act via EATS mechanism and may require metabolic activation for activity.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Note that the ext-1 gen assay provides the most information on endocrine disruption. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non- endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
H	+	Eq/0	-	1) Evidence of adverse effects on endocrine/ reproductive/developmental endpoints via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. 2) Evidence of adverse effects on reproductive/developmental endpoints via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity. 3) Evidence of adverse effects on endocrine	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added	May be sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Consider potency of effects for existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or

Scenarios	Result of OECD TG 421/422	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Reproductive/developmental endpoints and parental endocrine organs positive** **2) Reproductive/developmental endpoints positive** **3) Parental endocrine organs positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern) ****			
				organs but without reproductive/developmental effects via non-EATS/non endocrine disruption mechanism or requires metabolic activation for activity.	metabolising system.	whether there may be non- endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
I	+	Eq/0	Eq/0	1) Evidence of adverse effects on endocrine/ reproductive/developmental endpoints via unknown mechanism. 2) Evidence of adverse effects on reproductive/developmental endpoints via unknown mechanism. 3) Evidence of adverse effects on endocrine organs but without reproductive/developmental effects via unknown mechanism.	Perform assay from level 5 *e.g.*ext-1 or 2-gen assay. To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	May be sufficient information to conclude evidence of concern for reproductive toxicity via endocrine disruption mechanism. Note that the ext-1 gen assay provides the most information on endocrine disruption. Consider existing results and whether endocrine disruption mechanism is credible for reproductive/developmental effects or whether there may be non- endocrine mechanisms. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical.
J	-	+	+	No evidence of adverse effects in OECD TG 421/422. Effects seen in existing (lower level) studies do not lead to adverse outcome in level	Perform *in vitro* ER, AR, TR, S assays with	Consider route of exposures and possible implications for ADME characteristics of the chemical.

Scenarios	Result of OECD TG 421/422	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Reproductive/developmental endpoints and parental endocrine organs positive** **2) Reproductive/developmental endpoints positive** **3) Parental endocrine organs positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern) ****			
				5 assay. Metabolism or potency explains the difference from existing *in vitro*/ and *in vivo* data.	added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	Effects seen in existing studies may be in a more sensitive life stage. Further mechanistic studies with metabolism may help determine MoA.
K	-	+	-	No evidence of adverse effects in OECD TG 421/422. Metabolism or potency explains *in vitro*/*in vivo* differences.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	There may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* activity is not realised. Further mechanistic studies with metabolism may help determine MoA.
L	-	+	Eq/0	No evidence of adverse effects in OECD TG 421/422. Metabolism or potency explains *in vitro*/*in vivo* differences. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR	Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* activity is not realised. Further mechanistic studies with metabolism may help determine MoA. Consider route of exposures and possible implications for ADME

Scenarios	Result of OECD TG 421/422	Existing Results		Possible conclusions **Note: three possible outcomes for a positive result are given:** **1) Reproductive/developmental endpoints and parental endocrine organs positive** **2) Reproductive/developmental endpoints positive** **3) Parental endocrine organs positive**	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	**Effects (*in vivo* effects of concern) ****			
					Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	characteristics of the chemical
M	-	-	+	No evidence of adverse effects in OECD TG 421/422. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system. OR Perform assay from level 5 *e.g.*ext-1 or 2-gen assay.	If existing data are from an adequate level 5 assay then question why differences. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical.
N	-	-	-	No evidence of adverse effects in OECD TG 421/422.	Consider existing data, there may be no need for further testing.	There may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information).
O	-	-	Eq/0	No evidence of adverse effects in OECD TG 421/422. No evidence for (anti)-EATS activity *in vitro*.	Consider existing data, there may be no need for further testing.	There may be sufficient information to conclude absence of concern for endocrine disruption. Further mechanistic studies would increase evidence. Consider route of exposures and

Scenarios	Result of OECD TG 421/422	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Reproductive/developmental endpoints and parental endocrine organs positive 2) Reproductive/developmental endpoints positive 3) Parental endocrine organs positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern) **			
					To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	possible implications for ADME characteristics of the chemical. Check data on chemical analogues.
P	-	Eq/0	+	No evidence of adverse effects in OECD TG 421/422. Effects seen in existing (lower level) studies do not lead to adverse outcome in level 5 assay. Effects seen in existing studies are via unknown mechanism.	Perform *in vitro* ER, AR,TR, S assays	Further mechanistic studies would increase evidence. Consider route of exposures and possible implications for ADME characteristics of the chemical. Effects seen in existing studies may be in a more sensitive life stage. Check data on chemical analogues.
Q	-	Eq/0	-	No evidence of adverse effects in OECD TG 421/422.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays	There may be sufficient information to conclude absence of concern for endocrine disruption (the ext-1 gen assay provides the most information). Further mechanistic studies would increase evidence.

Scenarios	Result of OECD TG 421/422	Existing Results		Possible conclusions Note: three possible outcomes for a positive result are given: 1) Reproductive/developmental endpoints and parental endocrine organs positive 2) Reproductive/developmental endpoints positive 3) Parental endocrine organs positive	Next step which could be taken to increase evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern) **			
					with added metabolising system.	Check data on chemical analogues.
R	-	Eq/0	Eq/0	No evidence of adverse effects in OECD TG 421/422.	To further discern mechanism could perform *in vitro* ER, AR, TR, S assays with added metabolising system.	Further mechanistic studies would increase evidence. Check data on chemical analogues

Annex 3. Information on Endocrine Assay Costs

It has not been possible to obtain fully accurate costs for the performance of all assays for endocrine-disrupters because many of these procedures are relatively new and are not yet widely required by regulatory authorities. Testing companies have therefore only been able to provide estimated costs in many cases. The information reported in this annex should therefore be treated with caution.

The most comprehensive information to date has been compiled by the United States Environmental Protection Agency (USEPA) in order to estimate the costs of Tier 1 screening assays required by the Endocrine Disruptor Screening Program (EDSP). These data are available in an unpublished OECD document (ENV/JM/TG/EDTA(2010)1) which reproduces a 2009 USEPA report entitled '*Laboratory testing of chemicals for endocrine disrupting potential – analysis of market factors'*. A summary of the main screening test costs from this report is reproduced below in Table Annex 3.1.

The USEPA also gathered information about the various ancillary costs associated with setting up the Tier 1 endocrine screens, and these are reproduced in Table Annex 3.2. A total of 15 laboratories contributed costs data to the USEPA. Most were based in the United States, but a proportion were based in Canada, South America, Europe and Asia.

It is apparent from Table Annex 3.1 that estimated costs of these screens varied widely (by more than a factor of 10 in some cases), but it is clear (unsurprisingly) that the median costs of the *in vivo* assays are several times greater than those of the *in vitro* assays.

As part of the preparation of the present GD during 2010, more up-to-date information covering a wider range of assays responsive to endocrine disrupter was sought. A total of 12 responses were obtained from approximately 20 laboratories which were approached, although one 'response' was actually a mean of 2-3 unidentified Japanese laboratories. A few of the laboratories which responded had already contributed to the USEPA exercise described above. The responding group were based in the USA, Europe and Japan. Most of the laboratories were commercial testing houses, but one described itself as non-profit-making public sector. Set-up costs were not generally obtained, and the costs did not include chemical analytical costs except in one case.

A summary of the estimated assay costs obtained in 2010 is shown in Table Annex 3.3. Again, costs varied widely, even when only a few responses were obtained, suggesting that most laboratories have not yet geared up for the full range of endocrine assays which are likely to be employed in the future, and are essentially making educated guesses about likely assay prices.

The median costs of the *in vitro* assays lay between US$ 9,000 and US$36,000, which is a larger range than the USEPA data, but in the same order of magnitude. The wildlife assays (covering screening and higher tier tests) had median costs of between US$ 55,000 and US$ 355,000. However, note that only one laboratory responded with information on avian lifecycle test costs, and only two responded on the amphibian growth and development test. The mammalian *in vivo* assays (also covering screening and higher tiers) had a similar range of median costs to the wildlife assays (US$ 34,000 – US$ 318,000), but note that only 2 laboratories responded with information about the long-term rat studies.

Conclusion

In summary, these data cannot be regarded as definitive, for the reasons stated above. However, they clearly show that the *in vitro* assays tend to be less expensive than *in vivo* assays, and that *in vivo* screening assays tend to be less expensive than higher tier *in vivo* tests. This latter point, while not unexpected, reinforces the importance of avoiding, as far as possible, the use of ethically less desirable higher-tier tests with their intensive use of test animals.

Table Annex 3.1. Summary of costs (US $) for Tier 1 screening assays employed in the Endocrine Disruptor Screening Program. Obtained from commercial laboratories in 2009 by the US Environmental Protection Agency (ref.: OECD ENV/JM/TG/EDTA(2010)1).

Assay	Costs			# of Responses
	Median	Minimum	Maximum	
In vitro				
Rat ER binding	$17,250	$7,000	$55,000	7
Rat AR binding	$17,250	$7,000	$55,000	7
Human aromatase	$25,000	$4,000	$60,000	6
Human steroidogenesis	$13,750	$4,500	$31,400	4
Human ER transcriptional	$15,000	$3,000	$55,000	6
In vivo rat tests				
Uterotrophic	$43,050	$9,500	$65,500	6
Hershberger	$47,400	$9,200	$72,000	6
Pubertal female	$87,100	$47,200	$152,500	6
Pubertal male	$93,500	$47,200	$160,000	6
In vivo tests of aquatic species				
Amphibian metamorphosis	$80,597	$20,000	$97,700.	6
Fish reproduction	$92,500	$85,000	$137,650	7
Total				
All 11 assays	$532,397			

Table Annex 3.2. Additional costs of Tier 1 screens (ref.: OECD ENV/JM/TG/EDTA(2010)1).

Assay	Additional Costs				
	Method Validation		Preliminary Analytical Work		Average Total Additional Costs
	Range	Average	Range	Average	
In vitro					
Rat ER binding	$7,000 to $9,000	$8,000	Included in the cost of the assay (Table 6).		$8,000
Rat AR binding					
Human aromatase					
Human steroidogenesis					
Human ER transcriptional					
In vivo rat tests					
Uterotrophic	$8,000 to $15,000	$11,500	$1,000 to $2,500[a]	$2,000[a]	$13,500
Hershberger					
Pubertal female					
Pubertal male					
In vivo tests of aquatic species					
Amphibian	$5,000 to $10,000	$7,333	$5,000 to $20,000[b]	$10,000[c]	$17,333
Fish					

[a] For verification of dosing concentration
[b] For preliminary solubility and delivery trial setup
[c] The average for soluble compounds was $5,000 while less soluble compounds was $15,000. The mean of these two averages is $10,000.

Table Annex 3.3. Summary of estimated endocrine assay costs (US$) obtained in 2010 during preparation of this GD.

Endocrine assay cost estimates	Number of responding labs	Median cost	Minimum cost	Maximum cost
Wildlife assays:				
Amphibian metamorphosis assay (USEPA 890.1100) (OECD TG 231)	7	$75,000	$50,000	$96,000
Fish short-term reproduction assay (USEPA 890.1350) (OECD TG 229)	9	$90,450	$40,000	$130,000
Fish 21 d assay (OECD TG 230)	5	$55,000	$30,000	$83,000
Fish sexual development test	4	$139,500	$110,000	$160,000
Fish lifecycle toxicity test (USEPA 850.1500)	4	$251,250	$97,400	$386,000
Medaka multi-generation test	4	$355,000	$255,000	$500,000
Amphibian growth and development test	2	$216,000	$50,000	$382,000
Avian reproduction test (OECD TG 206)	1	$100,000	$100,000	$100,000
Avian 2-generation test	1	$275,000	$275,000	$275,000
Mammalian *in vitro* assays:				
Androgen Receptor Binding (Rat Prostate) (USEPA 890.1150)	5	$36,300	$13,450	$57,000
Aromatase (Human Recombinant) (USEPA 890.1200)	5	$19,000	$8,450	$40,000
Estrogen Receptor Binding (USEPA 890.1250)	5	$32,000	$13,450	$52,000
Estrogen Transactivation (USEPA 890.1300) (OECD TG 455)	4	$8,750	$6,650	$40,000
Steroidogenesis (Human Cell Line H295R) (USEPA 890.1550)	4	$22,750	$9,975	$40,000

Endocrine assay cost estimates	Number of responding labs	Median cost	Minimum cost	Maximum cost
Mammalian *in vivo* assays:				
Hershberger (Rat) (USEPA 890.1400) (OECD TG 441)	5	$57,000	$35,000	$129,000
Female Pubertal (Rat) (USEPA 890.1450)	5	$102,000	$61,000	$314,000
Male Pubertal (Rat) (USEPA 890.1500)	5	$104,000	$66,000	$305,000
Uterotrophic (Rat) (USEPA 890.1600) (OECD TG 440)	5	$34,000	$24,000	$122,000
2-generation reproduction (Rat) (OECD TG 416)	2	$318,300	$311,600	$325,000
Extended 1-generation reproduction (Rat)	2	$675,000	$137,883	$900,000
Repeated dose 28-day oral (Rat) (OECD TG 407)	2	$62,126	$52,972	$71,280

Annex 4. Glossary of Acronyms and Technical Terms

ADME	Absorption, distribution, metabolism and excretion
AFSS	Androgenised Female Stickleback Screen (GD 140)
AGD	Anogenital distance
AhR	Aryl hydrocarbon receptor
AMA	Amphibian Metamorphosis Assay (OECD TG 231)
AOP	Adverse Outcome Pathway
Apical endpoints	Results of an *in vivo* assay which describe a response by the organism as a whole, (*e.g*. fecundity or growth) which have possible implications for its biological fitness, rather than a response of the endocrine system alone (including physiological changes dependent on the endocrine system, such as VTG induction). Apical responses may or may not result from endocrine changes (*e.g*. fecundity may be affected both by some EDs and by some non-EDs)
AR	Androgen receptor
AR STTA	The Stably Transfected Human Andogen Receptor Transactivation Assay for Detection of Androgenic (ant)Agonist-Activity of Chemicals
ART	Avian Reproduction Test (OECD TG 206)
Assay	An experimental system that can be used to obtain a range of information from chemical properties through the adverse effects of a substance. The terms 'assay' and 'test method' may be used interchangeably for wildlife as well as for mammalian studies (OECD, 2005).
ATGT	Avian Two Generation Test
BCF	Bioconcentration factor
CF	Conceptual Framework
EATS	Estrogen/androgen/thyroid/steroidogenesis
ECx	x% effect concentration
ED	Endocrine disrupter
EDSP	USEPA Endocrine Disruptor Screening Program
EDTA AG	Endocrine Disruption Testing and Assessment Advisory Group
Endocrine active substance	A substance that affects endocrine endpoints – not necessarily an ED because the effects may not be adverse.
Endocrine disruption	"An ED 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." "A potential ED 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. (WHO, 2002).
EOGRTS	Extended One Generation Reproductive Toxicity Study) (OECD TG 443)
Epigenesis	Inherited changes in phenotype or gene expression caused by mechanisms other than alteration in gene sequences *e.g*. DNA methylation.
ER	Estrogen Receptor
ER STTA	The Stably Transfected Human Estrogen Receptor-alpha Transactivation Assay for Detection of Estrogenic Agonist-Activity of Chemicals (OECD TG 455)
FLCTT	Fish Lifecycle Toxicity Test
FSDT	Fish Sexual Development Test (draft OECD TG 234)
FSH	Follicle stimulating hormone
FSTRA	Fish Short Term Reproduction Assay (OECD TG 229)
GD	Guidance Document
GnRH	Gonadotropin Releasing Hormone
GR	Glucocorticoid Receptor
H assay	Hershberger Bioassay
HPG axis	Hypothalamic/pituitary/gonadal axis
HPT axis	Hypothalamic/pituitary/thyroid axis
ICCVAM,	Interagency Coordinating Committee on the Validation of Alternative Methods
Indicators of hormonal activity	These are endpoints in an *in vivo* assay which show whether or not the endocrine system has been stimulated, and often provide information of mechanistic value. In other words, they are not apical endpoints (see definition above). It is possible in some cases for indicators of hormonal activity to respond to a test chemical while apical endpoints do not respond, while in other cases, both types of endpoint give a response or only apical endpoints respond.
In vivo assay	Assay where a whole live animal is treated. This may be a mammalian assay where individual animals are treated or a wildlife assay where a population of animals is treated.
In vitro assay	Assay where whole live animals are not used. Systems used may include cell lines or subcellular preparations from untreated animals.
LABC	Levator ani plus bulbocavernosus muscle complex

LAGDA	Larval Amphibian Growth and Development Assay
LH	Luteinizing hormone
LOEC	Lowest-Observed-Effect-Concentration
MMGT	Medaka Multi-Generation Test
MOA	Mode/Mechanism of action
NOEC	No-Observed-Effect-Concentration
OECD	Organisation for Economic Cooperation and Development
OECD TG	OECD Test Guideline
OVX	Ovariectomised/ovariectomy
PEC/PNEC	Predicted environmental concentration / Predicted no-effect concentration
Possible ED	A possible endocrine disrupter is a chemical that is able to alter the function of the endocrine system but for which information about possible adverse consequences of that alteration in an intact organism is uncertain
PP assay	Peripubertal assay (male or female)
PPS	Preputial separation
(Q)SAR	(Quantitative) Structure Activity Relationship
REACH	Registration, Evaluation, Authorisation & Restriction of CHemicals
Screen	*In vitro or in vivo* assays which provide information on an endocrine disruption mechanism, but not generally information on adverse effects, for use in hazard or risk assessment.
STTA	Stably Transfected Transactivation Assay
T3	Tri-iodothyronine (thyroid hormone)
T4	Thyroxine (thyroid hormone)
Test	*In vivo* assays which can provide evidence to support a conclusion that a chemical is an ED that can cause adverse effects in an intact organism.
TG	Test Guideline
TSH	Thyroid Stimulating Hormone
TR	Thyroid hormone receptor
USEPA	United States Environmental Protection Agency
UT assay	Uterotrophic bioassay
Validation	The process by which the reliability and relevance of a particular approach, method, process or assessment is established for a defined purpose

	(OECD, 2005).
Validated assay (also equivalent to validated test method)	A test method for which validation studies have been completed to determine the relevance (including accuracy) and reliability for a specific purpose. It is important to note that a validated test method may not have sufficient performance in terms of accuracy and reliability to be found acceptable for the proposed purpose (OECD, 2005).
VO	Vaginal opening (or patency)
VTG	Vitellogenin
Weight of evidence	In the context of this document, this phrase implies that all relevant data from the test being evaluated, and from other tests on the chemical in question, should be considered before making decisions about interpretation of the new data and the possible need for additional testing. Each datum is not necessarily given equal weight – such weighting will depend on the type and reliability of the datum. Although it is possible to provide guidelines for weight of evidence assessment, its effective use will always depend to some extent on the application of expert judgement.
WHO	World Health Organisation
Wildlife	In this context (wildlife screens and tests), the test species are fish, amphibia and birds.
YAS	Yeast androgen screen
YES	Yeast estrogen screen

Annex 5. References

Andersen H R, Vinggaard AM, Rasmussen TH, Gjermandsen IM, and Bonefeld-Jorgensen EC (2002). Effects of currently used pesticides in assays for estrogenicity, androgenicity, and aromatase activity in vitro. Toxicol. Appl. Pharmacol. 179, 1-12

Ankley GA, Bennett RE, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR,Nichols JW, Russom CL, Schmieder PK, Serrrano JA, Tietge JE, and Villeneuve DI (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. Environ. Toxicol. Chem. 29, 730-741

Anway MD and Skinner MK (2006). Epigenetic transgenerational actions of endocrine disruptors. Endocrinology 147 (6) (Suppl.), S43-S49.

Ashby J and Lefevre PA (2000). The peripubertal male rat assay as an alternative to the Hershberger castrated male rat assay for the detection of antiandrogens, oestrogens and metabolic modulators. J.Appl.Toxicol. 20, 35-47.

Axelstad M, Hansen PR, Boberg J, Bonnichsen M, Nellemann C, Lund SP, Hougaard KS, and Hass U (2008). Developmental neurotoxicity of propylthiouracil (PTU) in rats: relationship between transient hypothyroxinemia during development and long-lasting behavioural and functional changes. Toxicol Appl Pharmacol. 232, 1-13. Battelle (2005). Draft final report on avian dosing study. Report no. 072205 to US Environmental Protection Agency – US EPA contract number 68-W-01-023, Battelle, Columbus, Ohio. 337 pp.

Battelle (2005). Draft final report on avian dosing study. Report no. 072205 to US Environmental Protection Agency – US EPA contract number 68-W-01-023, Battelle, Columbus, Ohio. 337 pp.

Benoit DA (1981). User's guide for conducting life-cycle chronic toxicity tests with fathead minnows (*Pimephales promelas*). United States Environmental Protection Agency, Duluth. EPA-600/8-81-011, 25 pp.

Birkhøj M, Nellemann C, Jarfelt K, Jacobsen H, Andersen HR, Dalgaard M, and Vinggaard AM (2004).The combined antiandrogenic effects of five commonly used pesticides.Toxicol Appl Pharmacol. 201,10-20.

Borgert CJ, Mihaich EM, Ortego LS, Bentley KS, Holmes CM, Levine SL and Becker RA (2011). Hypothesis-Driven Weight of Evidence Framework for Evaluating Data Within the U.S. EPA's Endocrine Disruptor Screening Program. Regulatory Toxicology and Pharmacology. Accepted for publication. Brown KH, Schultz IR, and Nagler JJ (2009). Lack of a heritable reproductive defect in the offspring of male rainbow trout exposed to the environmental estrogen 17-alpha-ethynylestradiol. Aquat. Toxicol. 91, 71-74.

Combes RD, (2000). Endocrine disruptors: a critical review of in vitro and in vivo testing strategies for assessing their toxic hazard to humans. ATLA 28, 81-118.

Cooper RL, Laws SC, Das PC, Narotsky MG, Goldman JM, Lee Tyrey E, and Stoker TE. (2007). Atrazine and reproductive function: mode and mechanism of action studies. Birth Defects Res. B Dev. Reprod. Toxicol. 80, 98-112

Cooper RL, Lamb JC, Barlow SM, Bentley K, Brady AM, Doerrer NG, Eisenbrandt DL, Fenner-Crisp PA, Hines RN, Irvine LF, Kimmel CA, Koeter H, Li AA, Makris SL, Sheets LP,

Speijers G, and Whitby KE (2006). A tiered approach to life stages testing for agricultural chemical safety assessment. Crit. Rev. Toxicol. 36, 69-98.

Crews D and McLachlan JA (2006). Epigenetics, evolution, endocrine disruption, health, and disease. Endocrinology 147 (6) (Suppl.) S4-S10.

Du G, Shen O, Sun H, Fei J, Lu C, Song L, Xia Y, Wang S, and Wang X (2010). Assessing hormone receptor activities of pyrethroid insecticides and their metabolites in reporter gene assays. Toxicol. Sci. 116, 58-66

Escande A, Pillon A, Servant N, Cravedi JP, Larrea F, Muhn P, Nicolas JC, Cavaillès V, and Balaguer P (2006). Evaluation of ligand selectivity using reporter cell lines stably expressing estrogen receptor alpha or beta Biochem. Pharmacol. 71, 1459-1469.

Fang H, Tong W, Perkins R, Soto AM, Prechti NV and Sheehan DM (2000). Quantitative comparisons of in vitro assays for estrogenic activities. Environ. Health Perspect. 108, 723-729.

Goldman JM, Laws SC, Balchak SK, Cooper RL. and Kavlock RJ (2000). Endocrine-disrupting chemicals: prepubertal exposures and effects on sexual maturation and thyroid activity in the female rat. A focus on the EDSTAC recommendations. Crit. Rev.Toxicol. 30, 135-196.

Hartig PC, Bobseine KL, Britt BH, Cardon MC, Lambright CR, Wilson VS, and Gray LE Jr (2002). Development of two androgen receptor assays using adenoviral transduction of MMTV-luc reporter and/or hAR for endocrine screening.Toxicol. Sci. 66,82-90.

ICCVAM (2003). ICCVAM evaluation of in vitro test methods for detecting potential endocrine disruptors: estrogen receptor and androgen receptor binding and transcriptional activation assays. NIH Publication No: 03-4503.

Jacobs MN, Janssens W, Bernauer U, Brandon E, Coecke S, Combes R, Edwards P, Freidig A, Freyberger A, Kolanczyk R, Mc Ardle C, Mekenyan O, Schmieder P, Schrader T, Takeyoshi M, and van der Burg B (2008).The use of metabolising systems for in vitro testing of endocrine disruptors. Curr. Drug Metab. 9, 796-826.

Jones RC and Edgren RA (1973). The effects of various sterols on the vaginal histology in the rat. Fertil. Steril. 24, 284-291

Klimisch H-J, Andreae M, and Tillmann U (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul. Toxicol. Pharmacol. 25, 1–5.

Knacker T, Boettcher M, Frische T, Rufli H, Stolzenberg H-C, Teigeler M, Zok S, Braunbeck T, and Schäfers C (2010). Environmental effect assessment for sexual endocrine-disrupting chemicals: fish testing strategy. Integ. Environ. Assess. Manag. 6, 653-662.

Kuiper GG, Lemmen JG, Carlsson B, Corton JC, Safe SH, van der Saag PT, van der Burg B, and Gustafsson JA (1998). Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. Endocrinology 139, 4252-3263.

Laws SC, Carey SA, Ferrell JM, Bodman GJ, and Cooper RL (2000). Estrogenic activity of octylphenol, nonylphenol, bisphenol A and methoxychlor in rats. Toxicol. Sci. 54,154-167.

Lo S, King I, Alléra A, and Klingmüller D (2007). Effects of various pesticides on human 5alpha-reductase activity in prostate and LNCaP cells. Toxicol. In Vitro. 21, 502-508

Marty MS, Crissman JW, and Carney EW. (2001). Evaluation of the male pubertal onset assay to detect testosterone and steroid biosynthesis inhibitors in CD rats. Toxicol. Sci. 60, 285-295.

Matsuura I, Saitoh T, Ashina M, Wako Y, Iwata H, Toyota N, Ishizuka Y, Namiki M, Hoshino N, and Tsuchitani M (2005). Evaluation of a two-generation reproduction toxicity study adding endpoints to detect endocrine disrupting activity using vinclozolin. J. Toxicol. Sci. 30 Spec No., 163-188.

NTP (2010). Multigenerational Reproductive Toxicology Study of Ethinyl Estradiol (CAS No. 57-63-6) in Sprague-Dawley Rats TR-547.

O'Connor JC, Frame SR, and Ladics GS (2002a). Evaluation of a 15-day screening assay using intact male rats for identifying anti-androgens. Toxicol. Sci. 69, 92-108.

O'Connor JC, Frame SR, and Ladics GS (2002b). Evaluation of a 15-day screening assay using intact male rats for identifying steroid biosynthesis inhibitors and thyroid modulators. Toxicol. Sci. 69, 79-91.

Odum J, Lefevre PA, Tittensor S, Paton D, Routledge EJ, Beresford NA, Sumpter JP, and Ashby J. (1997). The rodent uterotrophic assay: critical protocol features, studies with nonyl phenols, and comparison with a yeast estrogenicity assay. Regul. Toxicol. Pharmacol. 25,176-188.

Odum J, Tinwell H, Tobin G, and Ashby J (2004). Cumulative dietary energy intake determines the onset of puberty in female rats. Environ. Health Perspect. 112, 1472-1480.

OECD (2005). Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. No. 34.

OECD (2006a). Detailed Review Paper on thyroid hormone disruption assays. OECD Series on Testing and Assessment, OECD, Paris. No. 57.

OECD (2006b). Report of the Validation of the Updated Test Guideline 407: Repeat Dose 28-Day Oral Toxicity Study in Laboratory Rats. No.59.

OECD (2006c). OECD Report of the Validation of the Rodent Uterotrophic Bioassay: Phase 2 - Testing of Potent and Weak Oestrogen Agonists by Multiple Laboratories No. 66OECD (2007a). Guidance Document on the Uterotrophic Bioassay - Procedure to Test for Antioestrogenicity. OECD Series on Testing and Assessment. No. 71.

OECD (2007b). Detailed review paper for avian two-generation toxicity test. OECD Series on Testing and Assessment no. 74. Organisation for Economic Cooperation and Development, Paris. 167 pp.

OECD (2007c). Guidance document on amphibian thyroid histology. OECD Series on Testing and Assessment no. 82. Organisation for Economic Cooperation and Development, Paris. 47 pp.

OECD (2008). Report of a Workshop on Integrated Approaches to Testing and Assessment (IATA) No. 88

OECD (2008a). Detailed Review Paper on the Use of Metabolising Systems for In Vitro Testing of Endocrine Disruptors No. 97

OECD (2008b). Report of the OECD Validation of the Rodent Hershberger Bioassay: Phase 2: Testing of Androgen Agonists, Androgen Antagonists and a 5 α-Reductase Inhibitor in Dose Response Studies by Multiple Laboratories No. 86

OECD (2008c). Guidance document on mammalian reproductive toxicity testing and assessment. OECD Series on Testing and Assessment no. 43. Organisation for Economic Cooperation and Development, Paris. 88 pp.

OECD (2009a). Guidance document for histologic evaluation of endocrine and reproductive tests in rodents. OECD Series on Testing and Assessment no. 106. Organisation for Economic Cooperation and Development, Paris. 26 pp.

OECD (2009b). Report of the expert consultation to evaluate an estrogen receptor binding affinity model for hazard identification. OECD Series on Testing and Assessment no. 111. Organisation for Economic Cooperation and Development, Paris. 119 pp.

OECD (2009c) Guidance Document on the Weanling Hershberger Bioassay in Rats: A Short-term Screening Assay for (Anti) Androgenic Properties No. 115

OECD (2010a). Guidance document on the diagnosis of endocrine-related histopathology in fish gonads. OECD Series on Testing and Assessment no. 123. Organisation for Economic Cooperation and Development, Paris. 114 pp.

OECD (2010b). Workshop report on OECD countries activities and management of endocrine disrupters. OECD Series on Testing and Assessment no. 118. Organisation for Economic Cooperation and Development, Paris. 83 pp. + appendices.

OECD (2010c). Detailed review paper (DRP) on molluscs life-cycle toxicity testing. OECD Series on Testing and Assessment no. 121. Organisation for Economic Cooperation and Development, Paris. 182 pp.

OECD (2010d). Draft detailed review paper on environmental endocrine disruptor screening: the use of estrogen and androgen receptor binding and transactivation assays in fish. OECD Series on Testing and Assessment no. 135. Organisation for Economic Cooperation and Development, Paris..

OECD (2010e). Peer Review Report for the H295R Cell-Based Assay for Steroidogenesis No. 133.

OECD (2010f). Draft guidance document in support of the test guideline on the extended one generation reproductive toxicity study No. 151. Organisation for Economic Cooperation and Development, Paris.

OECD (2011a), Draft Detailed Review Paper State of the Science on Novel In Vitro and In Vivo Screening and Testing Methods and Endpoints for Evaluating Endocrine Disruptors. Organisation for Economic Cooperation and Development, Paris.

OECD (2011b). Validation report (Phase 1) for the fish sexual development test for the detection of endocrine active substances. OECD Series on Testing and Assessment no. 141, ENV/JM/MONO(2011)22, Organisation for Economic Cooperation and Development, Paris, 91 pp.

OECD (2011c), Guidance Document on the Current Implementation of Internal Triggers in Test Guideline 443 for an Extended One Generation Reproductive Toxicity Study, in the United States and Canada, Series on Testing and Assessment No. 117. Organisation for Economic Cooperation and Development, Paris.

Pickford DB (2010). Screening chemicals for thyroid-disrupting activity: a critical comparison of mammalian and amphibian models. Crit. Rev. Toxicol. early online 1-48.

Rockett JC, Narotsky MG, Thompson KE, Thillainadarajah I, Blystone CR, Goetz AK, Ren H, Best DS, Murrell RN, Nichols HP, Schmid JE, Wolf DC, and Dix DJ (2006). Effect of conazole fungicides on reproductive development in the female rat. Reprod. Toxicol. 22, 647-658.

Routledge EJ and Sumpter JP (1996). Estrogenic activity of surfactants and some of their degradation products assessed using a recombinant yeast screen Environ. Toxicol. Chem. 15, 241-248.

Sheahan DA, Brighty GC, Daniel M, Jobling S, Harries JE, Hurst MR, Kennedy J, Kirby SJ, Morris S, Routledge EJ, Sumpter JP, and Waldock MJ (2002). Reduction in the estrogenic activity of a treated sewage effluent discharge to an English river as a result of a decrease in the concentration of industrially derived surfactants. Environ. Toxicol. Chem. 21, 515-419.

Sohoni P and Sumpter JP (1998). Several environmental oestrogens are also anti-androgens. J Endocrinol. 158, 327-339.

Schultz T (2010). Adverse Outcome Pathways: A way of linking chemical structure to in vivo toxicological hazards. IN: Issues in Toxicology No. 7; In Silico Toxicology: Principles and Applications, (Eds.) M.T.D. Cronin and J.C. Madden., Royal Society of Chemistry.

Sohoni P and Sumpter JP (1998). Several environmental oestrogens are also anti-androgens. J. Endocrinol. 158, 327–339.

Soto AM and Sonnenschein C (2001). The two faces of Janus: sex steroids as mediators of both cell proliferation and cell death. J. Natl. Cancer Institute 93, 1673-1675.

Stoker TE, Parks LG, Gray LE, and Cooper RL (2000). Endocrine disrupting chemicals: prepubertal exposures and effects on sexual maturation and thyroid function in the male rat. A focus on the EDSTAC recommendations. Endocrine Disrupter Screening and Testing Advisory Committee. Crit. Rev.Toxicol. 30, 197-252.

Tan BL, Kassim NM, and Mohd M A (2003). Assessment of pubertal development in juvenile male rats after sub-acute exposure to bisphenol A and nonylphenol. Toxicol. Lett. 143, 261-270.

Takeyoshi M, Yamasaki K, Sawaki M, Nakai M, Noda S, and Takatsuki M (2002).The efficacy of endocrine disruptor screening tests in detecting anti-estrogenic effects downstream of receptor-ligand interactions. Toxicol. Lett. 126, 91-98

Trenzado CE, Carrick TR and Pottinger TG (2003). Divergence of endocrine and metabolic responses to stress in two rainbow trout lines selected for differing cortisol responsiveness to stress. Gen. Comp. Endocrinol. 133, 332-340.

Tyl RW, Myers CB, Marr MC, Brine DR, Fail PA, Seely JC and Van Miller JP (1999). Two-generation reproduction study with para-tert-octylphenol in rats. Regul. Toxicol. Pharmacol. 30, 81-95.

Tyl RW, Myers CB, Marr MC, Castillo NP, Seely JC, Sloan CS, Veselica MM, Joiner RL, Van Miller JP, and Simon GS (2006).Three-generation evaluation of dietary para-nonylphenol in CD (Sprague-Dawley) rats. Toxicol. Sci. 92, 295-310.

USEPA (1996). Fish Life Cycle Toxicity., USEPA Ecological Effects Test Guidelines, OPPTS 850.1500, United States Environmental Protection Agency, Washington. 4 pp.

USEPA (2007). Integrated summary report for the validation of an androgen receptor binding assay as a potential screen in the Endocrine Disruptor Screening Program http://www.epa.gov/scipoly/oscpendo/pubs/assayvalidation/status.htm

USEPA (2007a). Integrated summary report on aromatase http://www.epa.gov/scipoly/oscpendo/pubs/assayvalidation/status.htm

USEPA (2007b). Integrated summary report for Validation of a Test Method for Assessment of Pubertal Development and Thyroid Function in Juvenile Male Rats as a Potential Screen in the Endocrine Disruptor Screening Program Tier-1 Battery http://www.epa.gov/scipoly/oscpendo/pubs/assayvalidation/status.htm

USEPA (2007c). Integrated summary report for Validation of a Test Method for Assessment of Pubertal Development and Thyroid Function in Juvenile Female Rats as a Potential Screen in the Endocrine Disruptor Screening Program Tier-1 Battery http://www.epa.gov/scipoly/oscpendo/pubs/assayvalidation/status.htm

USEPA (2007d). Integrated summary report for validation of 15-day intact adult male rat assay as a potential screen in the endocrine disruptor screening program tier-1 battery http://www.epa.gov/scipoly/oscpendo/pubs/assayvalidation/status.htm

USEPA (2009). Endocrine Disruptor Screening Program Test Guidelines OPPTS 890.1300: Estrogen Receptor Transcriptional Activation (Human Cell Line (HeLa-9903) http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series890.htm

USEPA (2009a). Integrated summary report for Validation of an Estrogen Receptor Binding Assay using Rat Uterine Cytosol as Source of Receptor as a Potential Screen in the Endocrine Disruptor Screening Program Tier 1 Battery http://www.epa.gov/scipoly/oscpendo/pubs/assayvalidation/status.htm

USEPA (2009b). Endocrine Disruptor Screening Program Test Guidelines OPPTS 890.1150: Androgen Receptor Binding (Rat Prostate Cytosol). http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series890.htm

USEPA (2009c). Endocrine Disruptor Screening Program Test Guidelines OPPTS 890.1200: Aromatase (Human Recombinant). http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series890.htm

USEPA (2009d). Endocrine Disruptor Screening Program Test Guidelines OPPTS 890.1500: Pubertal Development and Thyroid Function in Intact Juvenile/ Peripubertal Male Rats http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series890.htm

USEPA (2009e). Endocrine Disruptor Screening Program Test Guidelines OPPTS 890.1450: Pubertal Development and Thyroid Function in Intact Juvenile/ Peripubertal Female Rats http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series890.htm

USEPA (2011). Endocrine Disruptor Screening Program. Weight-of-Evidence: Evaluating Results of EDSP Tier 1. Screening to Identify the Need for Tier 2 Testing. http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2010-0877-0021

Weybridge Conference (1996). European workshop on the impact of endocrine disrupters on human health and wildlife. Report of proceedings, EUR 17549, DG XII of the European Commission.

WHO (2002). Global assessment of the state-of-the-science of endocrine disruptors. Eds: Damstra, T., Barlow, S., Bergman, A., Kavlock, R. and Van der Kraak, G., WHO/PCS/EDC/02.2, World Health Organisation, Geneva. 180 pp.

WHO (2007). Harmonization Project Document No. 4, Part 1: IPCS Framework for Analyzing the Relevance of a Cancer Mode of Action for Humans and Case-Studies, Part 2: IPCS Framework for Analyzing the Relevance of a Non-Cancer Mode of Action for Humans

Witters H, Freyberger A, Smits K, Vangenechten C, Lofink W, Weimer M, Bremer S, Ahr PH, and Berckmans P (2010). The assessment of estrogenic or anti-estrogenic activity of chemicals by the human stably transfected estrogen sensitive MELN cell line: results of test performance and transferability. Reprod. Toxicol. 30, 60-72.