Organisation for Economic Co-operation and Development

ENV/CBC/MONO(2026)25

Unclassified

English - Or. English

11 August 2026

ENVIRONMENT DIRECTORATE
CHEMICALS AND BIOTECHNOLOGY COMMITTEE

GUIDANCE DOCUMENT SUPPORTING OECD TEST GUIDELINE 443 ON THE EXTENDED ONE-GENERATION REPRODUCTIVE TOXICITY TEST, Second Edition

OECD Series on Testing and Assessment, No. 445

E-mail: ehscont@oecd.org.

JT03591374

This work was approved and declassified by the Chemicals and Biotechnology Committee on [22/06/2026].

Please cite this publication as:

OECD (2026), *GUIDANCE DOCUMENT SUPPORTING OECD TEST GUIDELINE 443 ON THE EXTENDED ONE-GENERATION REPRODUCTIVE TOXICITY TEST, Second Edition*, OECD Series on Testing and Assessment, No. 445, OECD Environment, Health and Safety, Paris, https://one.oecd.org/document/ENV/CBC/MONO(2026)25/en/pdf.

Contact us

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

E-mail: ehscont@oecd.org

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 38 countries in North and South America, Europe and the Asia and Pacific region, as well as the European Union, 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 Partner countries 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 twelve different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides**; **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; Safety of Manufactured Nanomaterials;** and **Adverse Outcome Pathways.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (https://www.oecd.org/en/topics/chemical-safety-and-biosafety.html).

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

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, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank, Basel, Rotterdam and Stockholm Conventions and OECD. 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.

FOREWORD

This Guidance Document (GD) provides guidance on the design, conduction and interpretation of results of the OECD Test Guideline (TG) 443 for an Extended One Generation Reproductive Toxicity Study. This test for reproductive endpoints covers the interaction of males with females, pregnant females, females with offspring and the development of the F1 offspring to full maturity, at approximately 14 weeks of age.

The project for developing a Guidance Document supporting TG 443 was proposed by the Secretariat and included in the workplan in 2010. The document was then developed by two consultants in close cooperation with the expert group on reproductive toxicity and the Secretariat and was sent three times to the WNT for comments between September 2011 and October 2012.

TG 443 was updated in 2024-2025 for some endpoints, as suggested in the context of the extended one-generation reproduction toxicity study review project performed between 2021 and 2023 (ECHA 2023a). As part of the same project, revisions to GD 151 were targeted in 2025/2026 in conjunction with the update of TG 443. Together with the GD 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, they constitute a set of two TG 443-specific Guidance Documents (GDs). Other GDs, also published under the Series on Testing and Assessment (i.e. GD 43, GD 106, GD 150), provide support to reproductive toxicity studies. They are therefore also highly relevant for TG 443 and can be consulted in addition to this current document, since relevant aspects are cross-referenced where applicable. Some targeted corrections have been made within the course of this project to GD 43 on Mammalian Reproductive Toxicity Testing and Assessment.

Table of contents

1 SECTION 1: Introduction

Background

1. This Guidance Document (GD) has been developed to support the use of the OECD Test Guideline (TG) 443 (OECD, 2025a) for an Extended One Generation Reproductive Toxicity Study (EOGRTS); a test for reproductive endpoints that covers the interaction of males with females, pregnant females, females with offspring and the development of the F1 offspring to full maturity, at approximately 14 weeks of age. The TG describes three cohorts of F1 animals:

- Cohort 1: assesses reproductive/developmental endpoints; this cohort may be extended to include an F_2 generation.
- Cohort 2: assesses the potential impact of chemical exposure on the developing nervous system.
- Cohort 3: assesses the potential impact of chemical exposure on the developing immune system.

2. TG 443 was adopted by the OECD Council in 2011 and provides details on how the EOGRTS should be conducted. During the development of the TG, specific needs for guidance were identified to support the TG, especially where several design alternatives are proposed or for some of the procedures and endpoints that require further explanation. This Guidance Document provides these further details and guidance on data interpretation. However, the GD does not provide guidance on further assessment of fertility and reproductive performance of the F1 offspring (OECD, 2011) or in situations in which it may be acceptable to omit assessment of developmental neurotoxicity (DNT) and/or developmental immunotoxicity (DIT), which will depend on existing knowledge for the chemical being evaluated, as well as the needs of various regulatory authorities (OECD, 2012).

3. It should be noted that the Mutual Acceptance of Data (Council Decision C(81)30) applies to the Test Guideline itself and not to this Guidance Document.

Objectives and organisation of this Guidance Document

4. The objective of this GD is to support study sponsors and laboratories planning to carry out an EOGRTS and scientists evaluating the results of an EOGRTS for scientific and/or regulatory purposes. TG 443 provides details on how an EOGRTS may be conducted, but the design of the study will depend on existing information, regulatory requirements and whether cohorts have been omitted. This document gives advice on study design, including the gathering of key data on the substance to be tested, endpoints and data interpretation issues not detailed in the TG.

5. The GD has been developed from information originally included in drafts of TG 443 during its development phase, such as footnotes, appendices and includes more details with tables and outlines designed to provide a better overview. Guidance notes are also provided on issues that were identified

by the expert group at the October 2009 (expert group) meeting as being relevant to the TG. Guidance is not provided on every aspect of the EOGRTS, only on those identified as needing it.

6. The GD has been organised so that it complements the structure of TG 443 and is intended to provide a logical flow for a reader considering conducting the assay. Pre-study considerations, including collation of data and study design for the substance of interest, are presented and followed by selective guidance on in-life and terminal observations. Finally, some advice on data interpretation is given.

Other relevant OECD Guidance Documents

7. OECD Guidance Document 43 on Mammalian Reproductive Toxicity Testing and Assessment (OECD, 2026) covers methodological aspects and interpretation of data in the testing of chemicals for potential human and other mammalian reproductive toxicity. OECD GD 43 refers to the procedures used in other OECD *in vivo* reproductive tests TGs 414, 416, 421, 422 and 426. Many of these procedures are also used in TG 443, and therefore GD 43 is highly relevant for the EOGRTS. This current GD (151) refers to GD 43 for areas where the advice in GD 43 is considered adequate. In some cases, advice in GD 43 has become outdated and new references have been identified; in these cases, the advice given here should supersede the advice given in GD 43. Guidance Document 20 for Neurotoxicity Testing (OECD, 2004) can also be consulted for additional guidance on methods for testing of chemicals for potential neurotoxicity.

8. OECD Guidance Document 117 on the Current Implementation of Internal Triggers in TG 443 for an Extended One Generation Reproductive Toxicity Study, in the United States and Canada (OECD 2011) provides guidance for situations in which the EOGRTS will be submitted to regulatory authorities requiring internal triggers for the assessment of the second generation.

9. OECD Guidance Document 106 on Histologic Evaluation of Endocrine and Reproductive Tests in Rodents (OECD, 2009) provides information on the preparation and evaluation of endocrine organs and vaginal smears that may be helpful for the EOGRTS.

10. OECD Guidance Document 150 on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption (OECD, 2018) provides advice on the use and interpretation of assays in the OECD Conceptual Framework for Testing and Assessment of Endocrine Disrupters. TG 443 is included in the Conceptual Framework (CF) (see OECD, 2018), and OECD Guidance Document 150 describes the use and interpretation of results from TG 443 within a number of scenarios. Endpoints in TG 443 affected by endocrine active substances, and the kind of effects that might be expected from substances interfering with oestrogen, androgen, thyroid and steroidogenesis disruption are also described.

11. The Guidance Documents listed above should be consulted in addition to this document, since relevant aspects are cross-referenced where applicable.

Study outline and endpoints

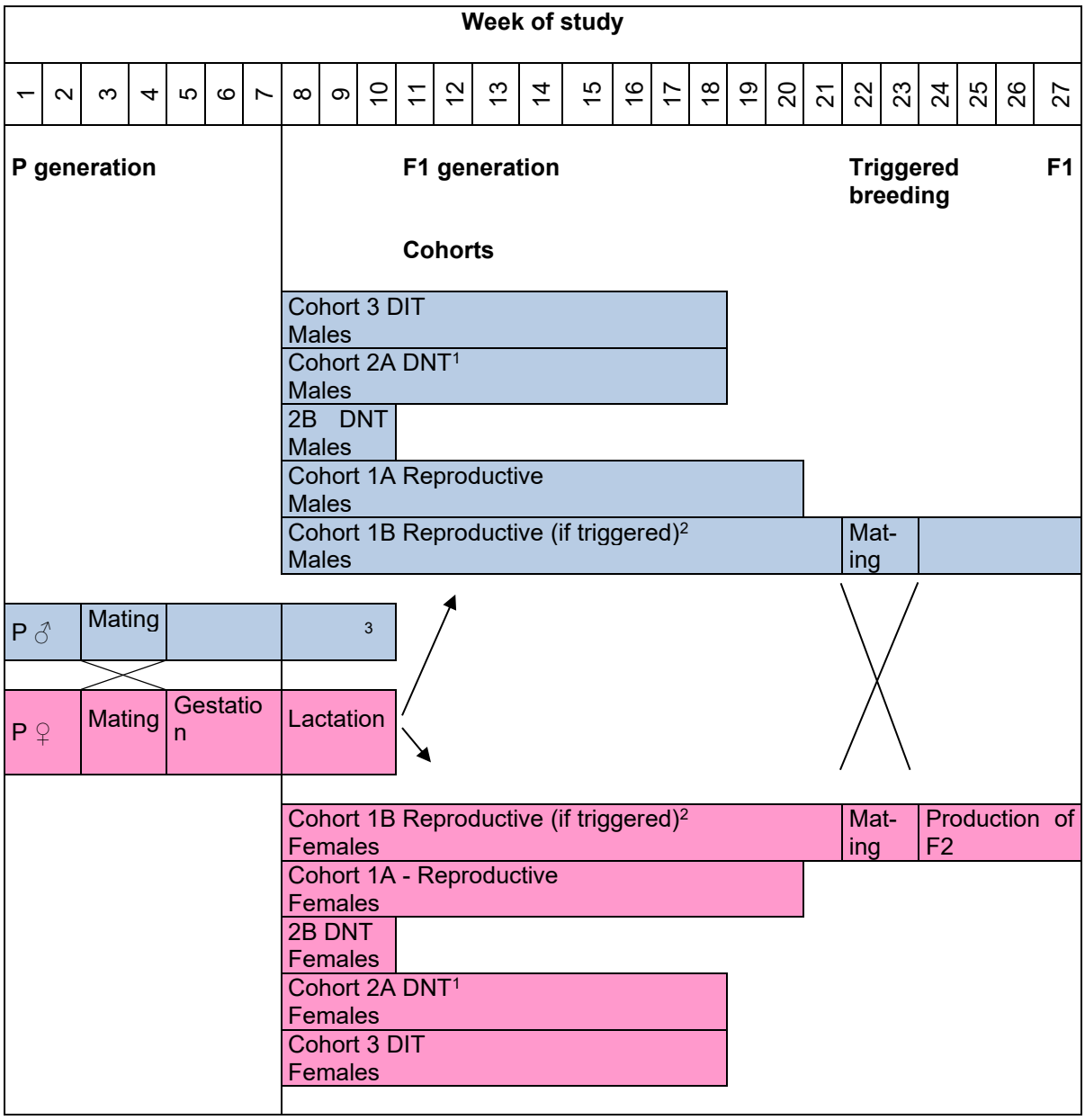
12. The design of the EOGRTS is provided in TG 443. A general summary is outlined in Figure 1 and Table 1, to give some context to the following sections of this GD. A detailed list of endpoints is given in Annex 1. In case the DNT (cohort 2) and/or DIT (cohort 3) cohorts are omitted or the F1 generation bred to produce an F2 generation (see paragraph 1), resulting changes should, however, maintain the required number of pups for reproductive assessment as detailed in this GD. Thus, whether the DNT and/or DIT assessments are performed or not, all animals in cohort 1A, 1B and cohort 3 (or 1C), should be maintained until sexual maturation to ensure that sufficient animals (2-3 pups/sex/litter/dose) are available for evaluation of critical endpoints such as sexual maturation parameters (vaginal opening and balano-preputial separation). For example, if there are 20 litters, 2

pups/sex from 10 litters and 3 pups/sex from the remaining 10 litters, should be chosen to yield 50/sex/dose for examination of these parameters. In studies where cohort 3 is omitted, the animals could be grouped into an additional cohort, labelled Cohort 1C to reach the required number of pups.

Animal welfare considerations

13. The number of animals used will be reduced, especially when EOGRTS is considered to fulfil information requirements on reproductive toxicity, developmental neurotoxicity and developmental immunotoxicity, as compared to the number of animals used in the three separate studies for these endpoints. As a one-generation study design, it also requires fewer animals than a two-generation study design because every generation increases the number of animals used.

Figure 1. Outline of study design. This figure is for illustrative purposes only. The details of the study are provided in TG 443 (2025), with further guidance in this GD.



Blocked colour indicates approximate treatment periods (males in blue, females in pink). The illustrated week of termination is approximate.

[1]Cohort 2A should be necropsied at approximately 11-12 weeks of age.

[2]Cohort 1B should be necropsied at approximately 14 weeks of age if a second generation is not produced or at approximately 20-25 weeks of age if a second generation is produced.

[3]Parental males require at least a 10-week treatment period and should be necropsied no sooner than indicated. If Cohort 3 is omitted, a sufficient number of pups need to be maintained to investigate sexual maturation in 3/sex/litter/dose (these animals could be grouped in a specific cohort, which could be labelled Cohort 1C, for example).

Table 1. General overview of the study design. This overview is also provided for illustrative purposes only. The details of the study are given in TG 443 with further guidance in the text of this GD.

- Parental (P) males and females are treated for a minimum two-week period followed by a two week mating period.
- Treatment is generally continuous, i.e. through pre-mating, mating, gestation and lactation stages of the parental generation and pre-weaning and post-weaning periods in offspring until termination
- The target is to achieve at least 20 litters per group from the P generation with sufficient numbers of F1 animals available for allocation to selected cohorts.
- Dams are allowed to litter and raise the pups. The litter size may be standardised on PND 4.
- After weaning, one male and one female F1 pup/litter are randomly assigned to cohorts 1A & 1B and one male or one female F1 pup/litter are randomly assigned to cohorts 2A, 2B and 3, as follows:
  - Cohort 1A: Assessment of effects on reproductive systems and toxicity (20M+20F/dose).
  - Cohort 1B: Assessment of reproductive performance (if required or triggered) and for obtaining additional histopathology data for reproductive or endocrine toxicity (20M+20F/dose).
  - Cohort 2A: Assessment of DNT post weaning (10M+10F/dose).
  - Cohort 2B: Assessment of DNT at weaning (10M+10F/dose).
  - Cohort 3: Assessment of DIT (10M+10F/dose).
- If there is an insufficient number of pups, then allocation to Cohort 1 should take precedence as the assessment of reproductive toxicity is the primary aim of the study.
- In-life measurements are determined as required by TG 443 (see Annex 1).
- The Cohorts are killed at approximately the following ages:
  - Cohort 1A: 13 weeks.
  - Cohort 1B: approximately 14 weeks if not mated, 20-25 weeks if mated.
  - Cohort 2A: 11-12 weeks.
  - Cohort 2B: 3 weeks (i.e. after weaning).
  - Cohort 3: 10-11 weeks.

2 SECTION 2: PRE-STUDY CONSIDERATIONS

14. Several factors will influence the design of the EOGRTS for a specific test substance. At the outset, all existing data should be reviewed, and all areas of the study considered to determine the most appropriate dose route, dosing schedule, number of animals, and other relevant factors. This section provides guidance and considerations on some important areas. TG 443 paragraph 80 specifies that the reliability and sensitivity of the test—such as positive historical control data and detection limits—must be provided. This applies especially to reproduction and endocrine disruption parameters (e.g., AGD in both sexes, retained areolas in males), as well as any parameter with statistically significant or biologically relevant changes or dose-response relationships.

Collation of existing data

15. Information from existing studies should be thoroughly reviewed when designing an EOGRTS. Suitable *in vivo* studies to be reviewed are those using repeated doses and reproductive studies. Of particular use are studies where effects on fertility, reproduction, development, reproductive organs or endocrine axes have been investigated. *In vivo* studies of the types and with the purposes of those contained in Levels 3-5 of the revised OECD Conceptual Framework (CF) for the Testing and Assessment of Endocrine Disrupters in GD 150 (OECD, 2018) have the most useful endpoints. The revised CF includes the most current OECD repeat dosing studies and reproductive studies. Non-standard studies addressing these endpoints are also very useful. Data on structural analogues of the substance tested in these assays may also be relevant.

16. A number of *in vitro* assays can provide data which may help setting dose levels or interpreting findings in the EOGRTS. These include, for example, the embryonic stem cell test (EST), bovine *in vitro* fertilisation assay (bIVF), bovine *in vitro* maturation assay (bIVM) and bovine *in vitro* pre-implantation assay (bIVP) (see summary in AXLR8, 2010 and Adler et al, 2011). *In vitro* assays providing data on endocrine mechanisms are listed in the revised OECD CF and in OECD Guidance Document 150 (OECD, 2018). Data on structural analogues, the use of QSAR tools (e.g. OECD, 2024) as well as a check of (potential) interaction with endocrine systems may help planning and interpretation of EOGRTS results.

17. In many cases only limited data on the chemical of interest may be available prior to the design and conduct of an EOGRTS. In the absence of any data or previous information on possible effects on reproduction, it is recommended that a preliminary reproductive study be conducted (see below in “Selection of route and dose”).

Consideration of toxicokinetic data

18. The use of toxicokinetic (TK) data (ADME – Absorption, Distribution, Metabolism, Excretion) during the planning of the EOGRTS is strongly recommended. This information will aid informed decisions on selection of the route of administration, choice of vehicle and selection of doses and in relation to considering whether extension of the duration of the pre-mating period is relevant (see paragraph 34). These data can also provide information regarding potential exposure of the offspring (*in utero* or via breast milk). Toxicokinetic information is also very valuable for interpretation of data obtained during the conduct of the EOGRTS.

19. Knowledge of the absorption, distribution, metabolism and excretion characteristics of a substance in the test species may help dose selection. For example, absorption of a substance may be saturated at a certain dose level. If toxicokinetic data are available beforehand, then the highest dose level could be set with the intention of avoiding saturation of toxicokinetic processes, as any higher dose level may not increase systemic exposure unless other factors, such as loss of the integrity of intestinal lining or microbial activity in the intestine, contribute in case of oral absorption. Saturation of TK processes may be included in the rationale for dose setting (see paragraph 26 provided that human exposures are expected to be well below the point of saturation. (Creton et al 2012; Saghir et al. 2012; McCoy et al. 2012; McFadden et al. 2012). If available, certain types of ADME data may also be used in relation to be considering whether the default pre-mating period of 2 weeks is sufficient. In case data on clearance suggest that an internal steady state concentration will not be achieved within the 2 weeks, it should be considered to extend the period long enough to allow a steady state to be reached.

20. Information from several sources can be utilised in order to design (and/or interpret) the EOGRTS. These include structural and physico-chemical parameters (e.g. chemical structure, molecular weight, water solubility, log P, physical state, vapour pressure and particle size: see ECHA (2023b) for a detailed discussion) which may allow a qualitative evaluation of TK behaviour. Information from *in vitro* testing may provide data on partition coefficients, permeation through membranes and the metabolic profile of a chemical. The latter may be addressed by the use of *in vitro* metabolising systems (Jacobs et al, 2008) and may also be useful to address potential gender and life-stage differences in xenobiotic metabolising enzymes relevant to the substance under consideration (Hines, 2008; Myllynen et al, 2009). *In vivo* information can be obtained from specific toxicokinetic studies (e.g. OECD TGs 417, 427, 428) where quantitative estimates for all aspects of ADME can be derived. Information on the differences in the toxicokinetics (e.g., biliary excretion or various membrane transport proteins) between human and the test species may assist in evaluating potential differences in the occurrence and/or potency of the observed effects. Limited toxicokinetic data are sometimes available from repeated dose toxicity studies or acute studies where additional endpoints have been included. Finally, Physiologically Based Toxicokinetics (PBTK) models may allow the estimation of the internal disposition towards a chemical during pregnancy, in the mother and in the embryo and fetus (Corley et al., 2003; Lee et al., 2002). Lactational transfer from the mother to the infant may also be assessed by measuring the compound or biologically active metabolites in the milk or pup tissues (Byczkowski and Lipscomb, 2001; Yoon and Barton, 2008).

21. TG 443 states that TK data at the following time points from late pregnancy, mid-lactation and early post-weaning in dams and offspring would be very useful in the planning of the EOGRTS:

- Gestation Day 20 (late pregnancy) - maternal blood and foetal blood
- Postnatal day 10 (mid-lactation) - maternal blood, pup blood and/or milk
- Postnatal day 28 (early post-weaning) - weanling blood samples

22. These data would provide information on the passage of the substance across the placenta, and/or lactational transfer, and thus reveal information regarding exposure of both dams and pups. Pups start to eat for themselves around late in the second postnatal week/early third postnatal week.

Before this, if pups do not receive the substance in milk or by direct dosing, there is a gap in exposure during a potentially critical window of development, from birth until the pup starts to eat for itself (in dietary studies) or when direct dosing commences (gavage studies) typically at weaning.

23. It is therefore suggested that evaluation of pup exposure be incorporated into the preliminary work designed to aid dose selection. Concentrations of test substance in pup blood and milk samples can be compared to maternal plasma levels at the same time point. Milk samples can be obtained from the stomach or directly by physical manipulation of the mammary glands, following an oxytocin injection, at about Day 10 of lactation and levels of test substance compared to maternal plasma and offspring plasma levels.

24. The results of these evaluations should be used to estimate whether exposure of the offspring is considered to be satisfactory for toxicity testing/safety evaluation. This may be discussed with some regulatory authorities. Where exposure levels in the offspring are not adequate, direct dosing of the offspring should be considered (see paragraph 28) during some stages of the pre-weaning period. The period of direct dosing the pups should not overlap other sources of exposure (e.g., during the third week of lactation if the test substance is included in the diet).

25. Where there is clear toxicity to the offspring, first signs of which appear during the lactation phase, it may be related to the transfer of the test substance to the offspring via the milk, in which case a detailed evaluation of their exposure level in milk or blood may not be required. However, reduced offspring growth, relative to controls, may also be a consequence of reduced milk production or quality or other maternal toxicity, and therefore, such results must be interpreted with caution.

Selection of route and dose

26. TG 443 (paragraphs 20-24) provides advice on dose selection. If there are no other relevant data, a preliminary reproductive study is recommended to evaluate endpoints critical to the EOGRTS prior to the main study. It will also aid in selecting the appropriate dose levels. Endpoints in the preliminary study should include mating success, fertility, litter production, and pup survival. It is also suggested that evaluation of pup exposure during gestation and lactation (as outlined above) is included in this preliminary study. Based on these measures, justification for the selected dose levels should be included in the report of the EOGRTS. Observations of decreased fertility (either from range finding study or from EOGRTS) should be clearly reported in order to allow derivations of effect levels (LOAELs) in case no further signs of toxicity are seen in the F1 generation.

27. The EOGRTS is designed to assess fertility and to evaluate the pre- and postnatal effects of chemicals on development. Based on the weight of evidence and/or specific regulatory authority's requirements, evidence of systemic toxicity or reproductive toxicity may be required at the highest dose level in order to ensure that the test system is optimised to be able to investigate any reproductive toxic property of a substance measured in the test system. As noted in paragraph 18, toxicokinetics may also be considered in dose selection. Consulting with regulatory authorities might be appropriate. It is recognised that some dose levels of the test substance may affect fertility, such that an insufficient number of pups may be produced for assessment of the F1 generation. In situations where fertility is affected, the lower dose levels should therefore be carefully selected to ensure the objectives of the study can be met.

28. The route of administration should consider the most relevant route for human exposure. Relevant guidance on routes of exposure for parents and offspring is also provided in GD 43 (74-77) (OECD, 2026).

29. When the route of administration is oral (by gavage, via the food or via the drinking water), exposure of the mothers can be continued through the period of birth and in the neonatal period.

However, it is uncertain if the offspring will have been exposed to the test substance during lactation (via maternal milk) unless there is evidence to demonstrate this (see paragraphs 22-23). It may therefore be necessary to consider the direct dosing of pups during some stages of the pre-weaning period. The potential technical, logistical and dosage problems involved in directly dosing young pups should not be underestimated.

30. Although there is a risk of injury to delicate tissues or accidental death, if done correctly, direct dosing of pups does not produce excessive stress in the pups. Careful consideration should be given to the impact of such procedures on toxic response and data interpretation (ILSI, 2003; Moser et al., 2005).

31. For whole body inhalation studies, the parent animals and the pups should be exposed simultaneously as separation of the dam from the litter is not favoured (OECD, 2008). Exposure routes in this situation will be dermal and oral (via grooming) in addition to inhalation. However, the benefits of not separating the pups from their parents are considered to outweigh the disadvantages of the exposures via the unintended routes.

32. The dermal route of exposure is not recommended for the EOGRTS. Although the dermal route may be the major exposure pathway in humans, the technical difficulties associated with reproductive testing by the dermal route outweigh the advantages of replicating the human exposure scenario. Some of the technical difficulties typically encountered from dermal exposure include: (i) disruption of nursing due to occlusion of application sites in maternal animals, (ii) techniques to ensure dermal exposure of the neonates, and (iii) incidental oral ingestion by the offspring during nursing. In addition, maternal care behaviour (*e.g.,* nesting, licking, grooming) may be affected due to the methods used to occlude the application site (not the compound) leading to stress and/or behavioural changes in the offspring. These factors could, in turn, confound interpretation of effects in the offspring (*e.g.,* decreased pup weight or changes in immune response may be affected by stress due to poor maternal care, or motor activity in the offspring may be affected due to occlusion of the application site). Other studies, such as ADME studies should be undertaken to facilitate extrapolation from the oral to the dermal route, if this is required.

Benchmark dose

33. When designing an EOGRTS, if a benchmark dose (BMD) approach is considered rather than using the No Observed Adverse Effect Level (NOAEL) as the point of departure (POD) for risk assessment, then dose levels should be selected that will enable its use. Guidance on the BMD approach can be found in the United States Environmental Protection Agency (US EPA)'s benchmark dose technical guidance document (US EPA 2000). The lower confidence level of the benchmark dose (BMDL) may be used as the POD. Default values for the magnitude of the response for which the BMDL is derived (i.e. the benchmark response – BMR) as well as the recommended dose-response models can be found in the European Food Safety Authority (EFSA) guidance on the use of the benchmark dose approach in risk assessment (EFSA, 2022). The EFSA document also provides guidance on reporting the results of a BMD analysis.

Pre-mating dosing schedules

i) Pre-mating exposure duration

34. TG 443 states that the "parental (P) generation should be dosed for a defined pre-mating period (selected based on the available information for the test substance; but for a minimum of two weeks)". This differs from TG 416 which requires a 10-week pre-treatment period. The basis for the minimum

requirement of a two-week pre-mating period is that a two week pre-pairing treatment is normally sufficient to establish effects upon male mating behaviour (Sakai, 2000) and effects on epididymal sperm maturation. In addition, histopathology and sperm analysis, which are included in TG 443, are considered more sensitive than male fertility assessment by mating and can detect the actions of a testicular toxicant at the end of the overall treatment period, which will be at least 10 weeks. For females, the two-week exposure period covers approximately 3 complete oestrous cycles. Thus, two weeks is also considered adequate for treatment-induced disruption of oestrous cyclicity to become established (Sanbuissho, 2009). The mating of the P (F0) females allows for the assessment of corpora lutea function during pregnancy. It should be noted however, that as the duration of the full folliculogenesis in female rats is at least 61 days (Latini, 2008), exposure covers the full folliculogenesis only in F1 animals. And if F1 animals are mated, it also includes the function of the corpora lutea and fertilisation of the ova exposed from primordial stage until maturation.

35. The adequacy of a two-week pre-mating period is justified below:

- For most substances, two weeks of treatment is sufficiently long to achieve steady state exposure conditions (Gibaldi and Perrier, 2009).
- Spermatozoa acquire their motility during the post-testicular phase when they transit the epididymis. This process takes 2 weeks in the rat, thus epididymal toxicity resulting in impaired sperm maturation, morphology and function will be detected by a 2 week pre-mating treatment period.
- Unless the test substance affects the mating ability of the males or females or the reproductive function of females, a litter will be produced even from males that have a clear testicular effect (on the most sensitive meiotic germ cell population) after 2 weeks of treatment. Therefore, maternal function and pre-and postnatal developmental toxicity of the substance may be evaluated, without the addition of animals or lowering of dose levels.
- When the P (F0) animals are evaluated for reproductive organ toxicity and sperm parameters following the overall 10 weeks of treatment any testicular toxicity will have had sufficient time to propagate throughout the downstream germ cell stages in the testis and the epididymis and have become detectable by testicular histopathology as well as by counting testicular and epididymal sperm.

36. Collaborative studies and review papers confirm that for rodents, a direct evaluation of testicular changes reliably detects effects on spermatogenesis and is more sensitive than a mating test to reveal the affected spermatogenesis (Ulbrich & Palmer, 1995; Mangelsdorf et al., 2003: Sakai et al., 2000 and Creasy, 2003). This view is also upheld for the detection of toxicity to reproduction for medicinal products and toxicity to male fertility (ICH Harmonised Tripartite Guideline, 2005). The ICH Guideline describes the justification for a two-week pre-mating period and the supporting collaborative studies (Sakai et al., 2000; Takayama et al., 1995).

37. Further guidance is included in GD 43 (OECD, 2026) under the section Methodological Issues for examination of male reproductive organs (paragraph 167) and for sperm parameters (paragraphs 169-174).

ii) Adaptation of the pre-mating exposure scenario

38. The pre-mating exposure schedule and duration of the EOGRTS may be extended when adequately justified. Consideration should be given to the duration of the dosing schedule based on available information on the test substance, including existing toxicity data, induction of metabolism or bioaccumulation. TG 443 (paragraph 27) states that the pre-mating exposure scenarios for males may be adapted if testicular toxicity (impairment of spermatogenesis) or effects on sperm integrity and function have been clearly identified in previous studies or, for females, if there are known effects of the

test substance on the oestrous cycle and possibly sexual receptivity. In addition, although for most substances; 2 weeks is sufficiently long to establish steady state exposure conditions, pre-mating exposure may be extended for a substance that requires a longer exposure period to reach steady state (cf. paragraphs 18 and 19 in this GD 151).

Effect of animal numbers on statistical power

39. The EOGRTS examines a total of 70 F1 animals per sex per dose, in various cohorts:

Cohort 1A – 1/sex/litter for a total of 20/sex/dose

Cohort 1B – 1/sex/litter for a total of 20/sex/dose

Cohort 2A – 1 male or 1 female/litter for a total of 10/sex/dose

Cohort 2B - 1 male or 1 female/litter for a total of 10/sex/dose

Cohort 3 - 1 male or 1 female/litter for a total of 10/sex/dose

40. As noted in paragraph 61 of TG 443, all F1 animals are examined macroscopically for any structural abnormality or pathological change at the time of termination or premature death (including those removed from the litter on PND 4 and at weaning). Special attention should be paid to the organs of the reproductive system (appropriate to the stage of development). Detailed examination of all offspring prior to necropsy will increase the likelihood of detection of rare or low incidence malformations such as hypospadias, or other effects on the reproductive axis. This is also relevant for the animals investigated for timing of sexual maturation (vaginal opening or preputial separation). Sexual maturation should be evaluated for required number of animals (aiming at 50 pups/sex/group) in cohort 1A, 1B and cohort 3 (or 1C), and results should be analysed statistically as described in paragraph 47 of TG 443. In studies where cohort 3 is omitted, the animals could be grouped in an extra cohort labelled Cohort 1C to reach the required number of pups (paragraph 35 of TG 443). Assessment of Cohort 2A animals for sexual maturation should not be performed, because the daily handling of the animals during the pre-pubertal and pubertal period could potentially affect the behavioural results of the DNT testing.

41. Twenty rats per sex per dose (i.e., 1/sex/litter) in Cohort 1A should be examined on PND 90 (gross necropsy with organ weights and histopathology) as defined in TG 443. An additional 20 rats per sex per dose, Cohort 1B, is included for termination at approximately 14 weeks (if not mated) or 20-25 weeks (if mated) of age and should be subject to gross necropsy with organ weights and tissues processed to block for future analysis, if required. In cases where results are ambiguous or equivocal within Cohort 1A, (e.g. atypical dose-response curves, lack of statistical significance, occurrence of rare or serious effects), or in cases of suspected reproductive or endocrine toxicants, the tissues from Cohort 1B should then be examined histologically to further characterise the nature of the effects.

Rationale for animal allotment in reproduction cohort

42. The concept that intra-litter variability exists and that there is value in assessing more than 1 animal/sex/litter has been highlighted in several published papers by different laboratories (Elswick et al, 2000; Gray and Gray 2006; Gray et al, 2004; George et al., 2003; Hotchkiss et al, 2008; Willoughby et al, 2000; Blystone et al, 2010). In fact, OECD Guidance Document 43 (OECD, 2026) states:

> *"The power of a study, that is, the probability that a study will demonstrate a true effect, is important in the evaluation of prenatal toxicity data. Factors that may influence the statistical power include the sample size used in the study (with the assumption that the litter is the basic unit of analysis), the background incidence of the finding, the variability in the incidence of the endpoint, the robustness of the data, and the method of analysis."*

> *"For multigeneration studies, the detection of structural abnormalities in the F1 and F2 pups*

> *has been shown to be dependent not only on the number of litters assessed, but also on the number of pups from each litter that are examined for each endpoint, and on the degree of relatedness of the effects in one pup in the litter to another. Since the pups are not identical, there is statistical value (improved power) gained from examining all of the pups in a litter for a postnatal malformation, as is done in the developmental toxicity study. Examining many pups/litter in the F1 generation greatly enhances the ability to detect low incidence[1] effects. Even when litter mean values are analyzed, examining more than one pup per sex per litter can improve the statistical precision of the analysis (reducing the error mean square used to calculate the F statistic). In general, the size of "litter effect" is not the same for all endpoints in a multigeneration study, the size of the litter effect varies across dose (being larger at high, more effective dose levels), and the litter effect for an organ varies from one chemical mode of action to another."*

43. In their publication Hotchkiss et al., calculated intra-litter correlation coefficients (the degree to which pups within a litter differ from one another) and examined how using different numbers of male pups in a litter affected the power calculations for a number of reproductive endpoints from several of their studies (Hotchkiss et al. (2008)). The higher the variation among pups within the litter, the more power was enhanced, and hence standard error of the mean was reduced by examining an increased number of pups from the same litter. The variability among pups within a litter appeared partly dependent on the mode of action of the test chemical and the endpoints evaluated. They also reported that:

> *"If 20 animals per dose group are examined for malformations then lesions occurring at an incidence of 25% or greater can be detected whereas an incidence of 10% can be detected if all the pups are examined from 20 litters. If only ten males per group are examined, as recommended for histopathological analyses in some regulatory agency test guidelines, then effects are only detected statistically if about 50% or more of the tissues/organs are affected; a level of statistical power that many would consider inadequate."*

Similar observations have been made by Blystone et al (2010) who found that evaluation of three males per litter, retained to adulthood, provided a substantial increase in the detection rate of male reproductive tract abnormalities compared to situations in which only 1 male per litter was evaluated.

44. These studies therefore demonstrate the importance of maximising the use of the animals on test to improve the ability to detect rare or low incidence effects of test chemicals. As is current practice, the litter mean values should still be considered in the analysis of the data and a statistical method based on data from all investigated pups should be used. The biological relevance of the findings should also be considered separately from statistical significance.

Housing and feeding - phytoestrogen content of the diet

45. TG 443 (paragraphs 12-13) recommends the use of standardised open-formula laboratory animal diets in which oestrogenic substances have been reduced. There are many literature reports of diets high in phytoestrogen content reducing the sensitivity of endocrine assays, particularly the immature rat uterotrophic screening assay (Boettger-Tong et al, 1998; Thigpen et al, 2007; Owens et al, 2003). Common sources of dietary phytoestrogens are soy and alfalfa that contain isoflavones (genistein and daidzein) and coumestrol respectively. An analysis carried out during the OECD validation of the uterotrophic assay showed that whilst high levels of phytoestrogens (350 µg/g total genistein equivalents in rats, extrapolated to 175 µg phytoestrogens/g diet for mice[2]) could diminish the responsiveness of the assay, levels below this had little effect (Owens et al, 2003). There are no reports

[1] The paragraph cited in GD 43 reads "low dose" but the expert group drafting GD 151 agreed that this should read "low incidence" as this is the intended meaning.

[2] as the food consumption of mice on a body weight basis is higher than that of rats.

of similar effects in apical tests such as the EOGRTS and there is much value in laboratories using the same diet used in previous studies. Furthermore, alteration of the constituents of standardised open-formula laboratory animal diets is not encouraged as the loss of key constituents and trace elements are known to adversely affect parturition and survival of the neonates. However, phytoestrogen levels are not always predictable and may vary from batch to batch of dietary constituents (Thigpen et al, 2004; Kato et al, 2004). Unless available from the supplier (e.g. closed formula diets supplied with analytical certificates), the concentrations of phytoestrogens in the diet and cage bedding - phytoestrogens have also been reported in laboratory animal bedding (Markaverich et al, 2002) - should be examined before the start of the study. Diets and beddings with high concentrations of phytoestrogens should not be used

Choice of animals

46. TG 443 (paragraph 10) states that the rat is the preferred species for the EOGRTS. In selecting the strain of rat for use in this study, it is necessary to consider the mean litter size and the probability of obtaining a sufficient number of pups to meet the objectives of the study and to provide adequate litter representation for each intended cohort. The strain of rats chosen should be one that has a reliable reproductive performance. Usually, the strain in the laboratory for which there are historical data is used for both the reproduction and repeated dose toxicity studies. Wistar or Sprague Dawley rats are most commonly used. There is some evidence of a positive relationship between the body weight of the dam and the number of oocytes produced (Harper, 1964). It may therefore be appropriate to delay the age at mating for those strains where an increase in litter size can be obtained by mating the females slightly later than the TG 443 recommendation of 90 days of age. It may be recommended to start mating when the females reach a body weight range of 200 – 230 g, with an upper limit of 250 g. However, it should be kept in mind that the factor most likely to affect pregnancy rate is not so much body weight but rather body fat (for which body weight is only an indirect indicator) (Palmer and Ulbrich, 1997). Strain differences in sensitivity to endocrine active substances have been reported in rats but this varies according to the substance tested and the endpoint determined (Putz et al, 2001; Steinmetz et al, 1998; Long et al, 2000; US EPA 2007). This may be relevant when comparing results from the EOGRTS with studies conducted in different strains.

Litter standardisation

47. TG 443 (paragraph 33) refers to the optional procedure for standardising the litter size to five males and five females by elimination of the extra pups in the litter on day 4 after birth. If an adjustment is made, the selection of pups must be done on the specified day, be random, and the procedure fully documented.

48. Guidance on litter standardisation is given in GD 43 (paragraphs 70-73) (OECD, 2026). The target adjusted litter size should be based on the normal litter size of the strain used. Ten pups per sex is generally appropriate for Sprague Dawley rats with a natural litter size of about 14, but 8 per litter is more achievable for Wistar rats, which normally have smaller litter sizes. Selection of 4 males + 4 females per litter provides sufficient animals for allocation to all cohorts. The decision to standardise to 5 males + 5 females or to 4 males + 4 females should be taken on a study basis and in relation to the chosen strain of animal and not on an individual litter basis during the study.

49. For those litters where there are sufficient pups but an unequal number of males and females such that selection of 5 males + 5 females (or 4 males + 4 females) cannot be achieved, it is acceptable to standardise the litter to 10 (or 8) such that each sex is represented as far as possible e.g. 2 males + 8 females or 6 males + 2 females. This variation in procedure is to avoid the unnecessary waste of animals when these could be used to generate additional data. However, consideration of the possible

effect of litters with an imbalance in sex ratio should be included in the statistical analysis of the data where the litter is the unit of analysis.

50. For those litters where there is an insufficient number of pups to allow standardisation of the litter to 10 (or 8), a decision to remove the affected litters on study should be made on a case-by-case basis. Any removal of any litter should be justified. However, when there is any indication that the reduced litter size is treatment-related it would be appropriate to retain the litters (and treatment group) on study. All decisions should consider the fulfilment of the study objectives without compromising the study integrity as well as the impact on data interpretation and statistical analysis.

Selection of pups to cohorts

51. TG 443 (paragraph 34) states that on PND 21, selected F1 pups are required to be randomly assigned to cohorts. A detailed description of cohort allocation is also given paragraph 39 of this GD. Allocation of the pups to the cohorts should follow this regime as far as possible. It is recommended that the total numbers are not exceeded to ensure a consistent group size at the outset and to maintain the balanced distribution of litter representatives. Where the lack of available pups/litters necessitates an increase in litter representation, careful consideration should be given to the allocation of pups to each cohort, to optimise litter distribution. For example, cohorts 2A, 2B and 3 have a smaller group size in comparison with cohorts 1A and 1B, with total of 10/sex/dose taken ideally, from 20 different litters. When a shortfall in available pups/litters is encountered, it may be preferable to maintain this litter distribution for cohorts 2A, 2B and 3 (or 1C) as far as possible and to allocate additional pups to the larger cohorts 1A and 1B. At all times, the representation of litters should be optimised within each cohort.

52. As litters are generally born over a number of days, the temporal spread needs to be considered when allocating animals to cohorts, particularly to cohort 2A and 3. This consideration is necessary to ensure that equal representation of groups is maintained as far as possible to ensure no bias or temporal differences with the subsequent evaluations e.g. motor activity. GD 43 (OECD, 2026) provides guidance on the General Methodological Considerations for Conducting Neurobehavioural Measures.

Achieving the correct balance at necropsy

53. At necropsy, consideration should be given to ensure equal representation of animals by sex and treatment group as far as possible on any one day, to prevent bias and temporal differences.

Conducting the study in blocks

54. For some laboratories that do not have the capacity to perform all examinations at the same time, the EOGRTS may be conducted in blocks (e.g., parental animals divided into two or three groups with a staggered study start – e.g. 1 week between blocks), to allow for easier animal management and data collection. If the EOGRTS is conducted in blocks, each dose group must be equally represented in each block, and each block should start on study as soon as possible to prevent variance that may be introduced by temporal differences in data collection. If a block design is used, the decision whether to breed the Cohort 1B animals may be delayed pending the complete collection of all data considered as potential triggers for the assessment of the second generation. In this case, necropsy of the Cohort 1B animals will be later than week 14 even if the second generation is not produced. With a block design, necropsy of Cohort 1B animals may occur between approximately 14 - 17 weeks of age if the second generation is not produced and between approximately 20 - 25 weeks of age if the second

generation is produced. The block design EOGRTS can adhere to the specified ages for data collection for other endpoints described in the test guideline. Consideration should be given to including "block" as a factor in statistical analyses.

Additional endpoints

55. The exposure of animals to test substances during critical developmental windows may provide an opportunity to collect data on endpoints not included in TG 443. There are many possible endpoints that could be included (see paragraph 56 in TG 443), but care should be taken when considering these so that they do not compromise the standard endpoints described in TG 443 (see Annex 1 of this document). The primary purpose of the study and the relevance of any additional endpoints to human health should be considered. The use of excess blood samples to measure other endpoints, for example, is unlikely to affect standard in-life endpoints; whilst addition of new in-life endpoints may provide such a conflict by increasing the complexity of the study or by adding other variables. Consultation with regulatory authorities would be advisable before additional endpoints are included, to ensure that the compliance of the study is maintained.

56. With scientific progress, increased knowledge of biological mechanisms and new emerging technologies, it is recognised that the EOGRTS provides an opportunity to measure additional investigative endpoints. These endpoints could relate to metabolic disturbances and include additional biochemical parameters, and body fat measurements (Heindel and Vom Saal, 2009; Ronti et al., 2006, Plagemann et al., 2009; Mathieu et al., 2009; Yki –Jarvinen, 2010), as well as to changes in the neuroendocrine system (Gore, 2008) and the neuroimmune system (Spencer et al., 2011; Merlot et al., 2008). As an example, part of the neural circuitry involved in control of gonadal function and sexually dimorphic behaviours are known, so that specific molecular targets can be studied in relation to functional changes. If validated, molecular markers could yield additional endpoints and could improve the sensitivity and mechanistic specificity of the protocol, assisting in comparisons with in vitro data and extrapolation to humans.

57. The 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" No. 178 (OECD, 2012) reviews the hypothalamus:pituitary:adrenocortical (HPA) axis, the hypothalamus:pituitary:gonad (HPG) axis, the somatotropic axis, the retinoid signalling pathway, the hypothalamus:pituitary:thyroid (HPT) axis, the vitamin D signalling pathway and the peroxisome proliferator-activated receptor (PPAR) signalling pathway. Many OECD TGs, including TG 443, could thus be modified to include new endpoints for the assessment of endocrine active chemicals disrupting these axes. This would expand the scope of the existing EOGRTS.

58. New investigative techniques may also help to improve the assessment of endpoints already included in TG 443. As an example, some reports have suggested that a useful extension of mammary gland histopathology is the analysis of whole mounts (You L et al., 2002; Munoz-de-Toro M et al., 2005), especially for weanlings (PND22), where histopathologic sections do not give as much information about the mammary glands. More work is necessary before this technique can be used in TG 443.

3 SECTION 3: IN-LIFE OBSERVATIONS WHERE FURTHER GUIDANCE IS PROVIDED

59. The in-life observations required are described in TG 443 and also detailed in Annex 1 of this guidance document. GD 43 (OECD 2026) provides guidance on these endpoints; however, some more recent literature references are provided below for anogenital distance, areola/nipple retention, and vaginal opening / balano-preputial separation. These endpoints are under hormonal regulation and therefore warrant specific attention in view of potential endocrine disruption. However, some of them (vaginal opening / balano-preputial separation) may also be sensitive to non-endocrine mediated changes. In general, blinding assessors to the exposure group is a methodological safeguard that strengthens reliability by minimising bias in outcome assessment (OECD, 2025b). Whenever technically possible, blinding of the observers performing in-life observations of the animals is therefore recommended. If blinding is not feasible, it is recommended to randomise animals across treatment groups as an alternative, as recommended in GD 116 (OECD, 2014, paragraphs 253-255). Another important issue to consider is that differences in scoring criteria among assessors within a study can introduce substantial variability into the results. Therefore, training sessions should be held to harmonise assessment techniques and minimise technical variability. Whenever possible, it is advisable that the same observer perform the assessment of a given endpoint. If this is not technically feasible, some demonstration of inter-observed reliability is required. To ensure an accurate evaluation, it is also important that assessors allocate sufficient time to assess endpoints in accordance with the established standardised procedures.

Anogenital distance

60. TG 443 (2025a) requires that anogenital distance (AGD) in all pups (i.e., all dose groups including controls) be measured at least once, on the same day of age, between PND 0 and PND 4, before any culling is performed. It is essential that all pups are measured on the same day of age, as rapid growth can influence AGD. For all pups, individual AGDs and body weights should be reported, and the assessment day should be noted. AGD measurements should be performed with sufficient precision to allow reporting of the results in mm, with at least two significant digits (as specified in TG 443 para 46). The following considerations apply to the standardised procedure for AGD measurements:

- A calibrated measuring instrument (e.g., slide gauge /calliper or a stereomicroscope with an integrated scale) should be used for all assessments.

- Clear and consistent anatomical criteria must be applied for every pup. It is essential to establish precise measurement landmarks—for example, measuring from the centre of the anus to the centre of the genital bud, or from the anterior margin of the anus to the posterior margin of the genital bud. A detailed description of this issue can be found in Wise (2024).
- Careful and consistent handling of the animals during measurement is crucial to minimise variability. This includes applying a standardised degree of stretching and ensuring uniform positioning of the perineal region, as differential stretching or positioning can introduce unwanted variation between animals.
- As already described in detail in paragraph 59, attention should be given to reducing variability in the assessments by using a few, well-trained assessors, who are given sufficient time to assess AGD according to the established standardised procedures. In-house training sessions should be performed to secure as much consistency among the technical staff as possible. Additionally, whenever technically possible, blinding observers to the exposure group is recommended.
- Because AGD correlates with offspring body weight, data analyses should use a standardised approach that accounts for this relationship. This issue is described in detail in paragraph 96 (on data interpretation).

Nipple/Areola retention

61. Because hair growth starting around postnatal day (PND) 15 in rat offspring makes it difficult or impossible to see areolae on the skin without shaving the animals, it is important to establish the correct time for the assessment of areola retention. As specified in TG 443 (OECD, 2025), the presence of areolae in male pups should be counted on the same PND between PND 12 and 14, when the nipples are visible in female littermates, for all pups (i.e., all dose groups, including controls). All pups should be evaluated on the same postnatal day, as marked differences can develop as maturation progresses. To ensure a high level of accuracy in the evaluation of this endpoint, adequate lighting is important, as areolae in male offspring are only visible as small dark spots on the skin, at locations corresponding to the clearly visible nipples in female littermates. Additionally, sufficient time allocation and thorough technical training of assessors are advised. Where practicable, the number of assessors and inter-observer variability should be minimised. Some guidance on the assessment of nipple/areola retention is provided in GD 43 (OECD 2026, paragraph 91). In line with the recommendations in GD 116 (OECD 2014, paragraphs 256–258) and OECD 2025b, it is advised to ensure that assessors are blinded to exposure groups. It is, however, recognised that in large reproductive toxicity studies like EOGRTS, implementing a single-blinded assessor for all measurements throughout the study may not always be operationally feasible. Accordingly, while the principle of blinding should be maintained wherever practicable, flexibility in the number of assessors may be acceptable, provided that:

- All assessors are demonstrably trained and supported with standardised procedures to ensure proficiency in data collection.
- Standardised and well-documented scoring criteria are applied uniformly, and
- Quality assurance measures are in place to verify the reliability and comparability of the data generated.

Vaginal opening, first oestrus and balano-preputial separation

62. F1 animals in Cohorts 1A, 1B and 3 (or 1C) should be examined daily for vaginal opening (VO, females) or for balano-preputial separation (BPS, males), as detailed in paragraphs 12 and 40 of this GD. Further guidance is provided in GD 43 (OECD 2026, paragraph 92). In accordance with TG 443 (OECD, 2025), examinations should begin on PND 35 for BPS in males and on PND 24 for VO in females and should continue until BPS and VO have occurred. Sexual maturation can be influenced by e.g. hormonal balance and physical development of the animal. Therefore, body weight should be recorded for each animal on the day VO or BPS is first observed. In addition, VO and BPS observations should be reported together with body weight measurements for all pups examined on the same PND (e.g. PND 21, 28 or 35). Interpretation of VO and BPS data in relation to body weight changes is discussed in paragraph 107 of this GD. Additionally, the first oestrus should be identified by conducting daily vaginal smear examinations in all F1 females assigned to Cohort 1A. Smears should begin on the day VO is observed and continue daily until the first cornified smear is detected (first oestrus). This allows calculation of the interval between VO and the onset of the first oestrus. To confirm that animals began cycling normally, it is recommended that Cohort 1A females continue to be smeared once daily until progression from oestrus to metestrus, or for up to three consecutive days if they remain in oestrus. Thus, whereas VO is assessed in female offspring from Cohorts 1A, 1B and 3 (or 1C), the determination of first oestrus is performed only in Cohort 1A.

63. One issue arises in taking vaginal smears in the rare instance where the developing female exhibits VO accompanied by the presence of vaginal threads (mesenchyme surrounded by keratinised epithelial cells). In some cases, VO may occur with a vaginal thread present. If a thread is observed, it should be recorded daily for one week after VO. If the thread persists one week after VO, it should be considered persistent. A female may well begin cycling during a thread's presence. While vaginal insertion of an eye dropper may break the thread (although some can be quite resilient), this is something that should be avoided, particularly if its persistence is associated with a toxicant exposure (e.g. TCDD, Gray and Ostby, 1995). When VO is first observed, those animals with such threads should still be smeared using a fire-polished narrow tip Pasteur pipette (instead of a dropper), inserting it adjacent to the thread in order to be able to record the appearance of first oestrus and the onset of cyclicity. Also, pinhole openings may be noted, but smears should not begin until full VO is present. In all cases, VO, pinholes and threads should be analysed separately.

Developmental neurotoxicity

64. Functional observation battery, motor activity and acoustic startle habituation are performed on cohort 2A. In addition to the information available in TG 443, guidance on these neurotoxicity examinations is covered in a number of published documents: Test Guideline 426 on Developmental Neurotoxicity (OECD, 2007), Guidance Document 20 for Neurotoxicity testing (OECD, 2004), Guidance Document 43 on Mammalian Reproductive Toxicity Testing and Assessment (OECD, 2026), Crofton et al., 2008, Makris et al., 2008. Thus, further guidance has not been detailed in this Guidance Document. Particular attention may need to be paid to sexually dimorphic behaviour (Tyl et al 2008).

4 SECTION 4: TERMINAL OBSERVATIONS WHERE FURTHER GUIDANCE IS PROVIDED

65. The terminal observations required are described in TG 443 and also detailed in Annex 1 of this document. GD 43 (OECD 2026) provides guidance on some of these endpoints, but some guidance related to more recent methods and other endpoints is provided below. It should be noted that for all terminal procedures, consideration should be given to ensure equal representation of animals by sex and treatment group as far as possible on any one day, to prevent bias and temporal differences.

Clinical biochemistry / Haematology

66. TG 443 requires the collection of fasted blood samples from a defined site. Care should be taken to ensure animals are in a comparable state (i.e. exsanguinated, fasted) at termination for organ weights and histological examination. For females with litters, fasting should start after removal of the pups on PND day 21 prior to termination on PND 22.

67. Full scale haematology, clinical biochemistry and assay of thyroid hormones are recommended. Thyroid hormone levels may be affected by fasting. If this is problematic, for example if comparison with non-fasting hormone levels is required, then non-fasting blood samples could be taken at the same stage (1 day before necropsy) in non-food deprived animals. In addition to the parameters required by TG 443, the examination of others may be indicated by the known effect profile of the test substance on a case-by-case basis. Serum markers of acute tissue damage may be considered for chemicals in certain classes or if a specific effect of the test substance has been observed in repeated-dose studies using special techniques, these could also be incorporated into this study. Examples might be acetyl cholinesterase activity for compounds known to inhibit this and blood methaemoglobin concentration for compounds known to increase methaemoglobin formation or specific hormone measurements for endocrine modulators.

Thyroid hormones

i. Thyroid Hormone Measurements

68. The importance of maintaining appropriate systemic concentrations of thyroid hormones for normal development, especially maturation and function of the central nervous system, has been well established (Zoeller & Rovet, 2004; Yang et al., 2009). The EOGRTS specifies the measurement of

thyroxine (T4) and/or thyroid stimulating hormone (TSH) in parental and F1 offspring at various life-stages to assess direct effects on thyroid function or indirect effects via the hypothalamic-pituitary-thyroid axis. Since some thyroid toxicants have been reported to induce changes in total T4 without concomitant changes in TSH (Zoeller et al., 2005; Chang et al., 2007; Zorrilla et al., 2009), serum concentrations of both hormones are assayed when possible, to provide information on the mode of action of the test chemical and its potential effect. Apart from the limited quantity of blood in very young pups on PND 4 that necessitates the assay of only T4, both hormones are expected to be assayed at all other life-stages. If there's a reason to measure both T4 and TSH, then it may be necessary to pool samples from 2 pups of the same sex and litter, or by litter. As thyroid hormones have diurnal fluctuations and TSH is pulsatile in nature (Zoeller et al., 2007), care should be taken to collect blood samples in a consistent manner, within a reasonable amount of time and at a similar time of day across treatment groups. Statistically or biologically significant changes in either one or both hormones between treated and control groups can be evaluated in conjunction with any changes in thyroid gland weight and histopathology, as well as neurological or other developmental adversities used for risk assessment.

ii. Thyroid hormone assay validation (T4) and quality control (T4 and TSH)

69. There are numerous immunoassay kits commercially available to measure total T4 and rodent TSH. Commercial assay kits for rodent TSH are recommended because of the species specificity of TSH (Davies, 1993; Christian & Trenton, 2003). Species specificity does not generally apply to T4; therefore, commercial T4 assay kits that have been developed for use with human serum can be validated and adopted for use with rodent serum (Veterinary Application Documents for Siemens Medical Solutions Diagnostics, 1993). For PND 4 pups, serum T4 might be below detection level of kits. The detection level of the commercial assay kits should be verified.

70. The performance of the T4 and rodent TSH assays should be examined prior to accepting and examining the thyroid hormone results between treated and control groups for any study. The standard reference curves, level of hormone sensitivity, quality control samples and within- and between-assay coefficients of variation should be within acceptable limits according to the manufacturer's specifications and laboratory SOPs. This information should be reported along with the results of the assay.

Ovary examination

71. Ovarian toxicants may cause loss of oogonia, oocytes, or supportive somatic cells resulting in adverse effects on reproduction. In reproductive toxicity study, identifying ovarian toxicants relies on several endpoints including evaluation of oestrous cyclicity, determination of fertility, organ weights and histopathology. Histopathological evaluation of the ovary should be informed by the normal ovarian morphology and the normal changes associated with the oestrous cycle and ageing. GD 106 (OECD, 2009) Part 3 (sections 1-5): Female Reproductive System, describes the normal female reproductive tract histology including a section on the ovary with description of the follicles.

72. In the EOGRT study, ovarian histopathology and enumeration of ovarian follicles and corpora lutea in F1 adults may be the only measures of fertility in females exposed in utero. F1 animals are actually exposed through the full folliculogenesis (follicle and ovum development and maturation phases). If they are mated, this also includes the functional phase of the corpora lutea and fertilisation of the ova. TG 443 (paragraph 73) specifies qualitative histopathological evaluation of the ovaries from the P and F1 females. The qualitative assessment of ovaries should include evaluations of any abnormalities/lesions such as ovarian atrophy as well as presence or absence of follicle maturational stages. Evaluation should be carried out by a trained veterinary pathologist using standardised nomenclature (Dixon et al. 2014). In addition, it specifies quantitative evaluation of primordial and small

growing follicles, and corpora lutea in the ovaries of the F1 females. Qualitative (histopathology) and quantitative analysis (follicles and corpora lutea counts) can be carried out on the same histological samples, provided they are prepared to allow for both types of evaluation (e.g. haematoxylin and eosin staining). To minimise operator bias and variation in quantitative morphological assessments (e.g., follicular counts, corpora lutea counts), all examinations within a study should ideally be carried out by a single observer trained in the specific technique. Efforts should be made to reduce the variability and generate reliable counts. As such, it is recommended that laboratories strive to minimise the % coefficient of variations (%CVs) when enumerating ovarian follicles and corpora lutea. Historical control values for %CVs should be provided by the testing facility to support interpretation of study-specific variability, particularly when study-specific %CVs exceed 25% for follicles or 20% for corpora lutea. The use of new technologies (e.g., digital pathology or use of machine learning) must be clearly disclosed in the study report, with supporting validation data for the tool and methodology utilised to ensure adequate reliability and accuracy for the intended purpose.

Follicle counts

73. TG 443 requires an enumeration of primordial and small growing follicles. To support a consistent approach across laboratories, it is recommended that the primordial follicles are defined as an individual oocyte surrounded by a partial layer of granulosa cells, which typically are flattened and squamous-like (corresponding to Pedersen Type 1; Pedersen and Peters, 1968), and the small growing follicles are defined as an individual oocyte surrounded by up to one complete single layer of cuboidal granulosa cells (corresponding to Pedersen Types 2, 3a and 3b; Pedersen and Peters, 1968). Including follicles with a single layer of granulosa cells will enhance consistency and reproducibility in follicle counts and support a more accurate functional interpretation of follicle numbers. Laboratories should provide a description (e.g. as above) of the follicles counted. The average number of follicles per ovary (expressed as mean or median) in the treatment group(s) should be compared to the average number of follicles per ovary in the concurrent control group.

74. For the enumeration of primordial and small growing ovarian follicles, the number of animals examined, the approach used for section selection, and the overall sampling strategy should be statistically appropriate for the evaluation method employed. As noted in TG 443, at least 1% of the ovary, or a minimum of 10 single sections per ovary, should be included in the quantitative assessment of follicles. Sufficient methodological detail should be provided to ensure clarity and reproducibility of the evaluation. This includes:

- number of ovaries evaluated;
- the location of the sections within the ovary (e.g., bisecting the ovary along its long axis through the central suspending ligament, with both halves embedded in the wax block);
- section thickness (e.g., 5 µm);
- the interval between levels for step-sectioning (e.g., 100 µm);
- the total number of levels examined; and
- the type of histological staining used (e.g., haematoxylin and eosin).

These details should be documented in the study report to allow for appropriate interpretation of the results.

75. The following provides some general guidance with respect to sampling procedures for follicle counts.

- At least one ovary from each of the females in a given group (i.e. the concurrent control and the treatment group) should be examined to optimise the statistical unit.

- To reduce variability in follicle counts, sectioning should consistently begin from the same anatomical region, selected sections should come from a defined area, or serial sectioning of the entire ovary should be performed. As an example, sections are taken from the middle third of the ovary. This could be achieved by cutting one ovary in half across its long axis and through the central suspending ligament. Both halves could then be mounted in the same wax block to provide serial sections from the central plane of both halves of the ovary at the same time.
- For section sampling, beginning from a smooth cut surface, the block could be cut at sections 5 microns thick, and every 20th section (or the section from every 100-micron interval) retained from each half of the ovary, for a total of 5 double sections (or 10 single sections) per ovary. This method is estimated to provide sections from a 1% sample of the ovary and judged to give a statistically appropriate sample size.
- Currently available methods for the enumeration of ovarian follicles include standard haematoxylin and eosin staining as well as methods involving immunohistochemistry (e.g. Yoshida et al., 2009; Picut et al., 2008; Muskhelishvili et al., 2005; Muskhelishvili et al., 2002). While immunohistochemistry may enhance contrast between follicles and ovarian stroma, the establishment of clear criteria used for follicle identification is necessary to ensure consistent and reliable quantitation, supported by a proper validation for the specificity and performance of the antibodies.
- To avoid double counting of follicles, only follicles with an oocyte with a visible nucleus should be counted.

76. As an aid to the evaluation of change in follicle counts, the following is suggested: where high dose group counts are not less than 85% of control and are not statistically significantly different, no further evaluation is considered necessary. When a statistically significant difference in count is obtained, this would trigger an evaluation of the ovaries of females from the mid and/or low dose groups. Where high dose group counts are less than 85% of control but are not statistically significantly different, it is recommended that the second ovary be processed, with an evaluation of an increased number of sections (e.g. 50 sections or approximately 5% of the ovary) to establish if the intergroup difference is statistically significant. Animals in Cohort 1B may be used if it is considered that evaluating additional ovaries could aid in clarifying the results. However, if Cohort 1B is mated to generate a 2nd generation, the females in Cohorts 1A (nulliparous) and 1B (post-lactational) would be in different physiological states which may add further variability in the results. If relevant, study reports should clearly identify this potential confounder in their assessment, and data from different cohorts should not be combined.

Corpora lutea counts

77. Quantitative evaluation of corpora lutea should be conducted histologically using a single section per ovary (Sato et al., 2014; Kucheryavenko et al., 2025, accepted). The section should be taken from the middle of the ovary, providing the maximum cut surface to ensure a representative selection of structures (OECD, 2009). Use of a single section level minimises the risk of counting the same corpora lutea more than once, thereby avoiding biologically implausible total numbers. Corpora lutea from the current cycle and those undergoing regression, should be included. Corpora lutea counts should be evaluated in conjunction with ovarian histopathology to identify morphological or cellular changes indicative of altered luteal development and regression (Taketa, 2021). Histopathological examination may reveal signs of accelerated regression such as increased apoptosis, fibrosis, disrupted vascularisation, or premature structural collapse that cannot be identified from counts alone. This integrated approach supports accurate interpretation of ovulatory activity and luteal function and is essential for detecting subtle toxicant-induced effects.

78. The average number of corpora lutea per ovary in the treatment group(s) should be compared directly with the average number of corpora lutea in the concurrent control group to determine whether

the test substance affects ovulation. The study report should clearly describe the methodology used for the identification and enumeration of corpora lutea, including the criteria applied to distinguish the stages of luteal development. As described in paragraph 72, provision of historical control data is recommended to support the interpretation of corpora lutea counts.

Testicular and epididymal histopathology; and sperm parameters

79. In the EOGRTS, detailed testicular histopathology examinations are conducted on the F1 males (cohort 1A and if needed, cohort 1B) in order to identify treatment-related effects on testis differentiation and development and on spermatogenesis. Guidance is provided in GD 106 (OECD, 2009) Part 2: Male Reproductive System. This document describes the normal physiology and structure of the reproductive system, the normal background variation of structure, morphologic patterns of hormone disruption, the recommended terminology and severity grading for histopathological findings and the critical aspects of histopathological evaluation. GD 43 (paragraphs 181-182) (OECD, 2026) provides advice on data interpretation for male reproductive organs with reference to organ weights and histopathology.

80. In P males, in addition to examining gross lesions such as atrophy or tumours, detailed testicular histopathology examinations should be conducted in order to identify treatment-related effects such as retained spermatids, missing germ cell layers or types, multinucleated giant cells or sloughing of spermatogenic cells into the lumen (Russell et al., 1990).

81. Examination of the intact epididymis should include the caput, corpus, and cauda, which can be accomplished by evaluation of a longitudinal section (Creasy, 2003). The epididymis should be evaluated for leukocyte infiltration, change in prevalence of cell types, aberrant cell types, phagocytosis of sperm, and the absence of clear cells in the caudal epithelium (Lanning et al., 2002).

Mammary gland histopathology

82. Histopathological analysis of the mammary gland is specified in TG 443 and may provide sensitive endpoints for detection of endocrine-related effects (Lucas et al, 2007; Fenton et al, 2002). It is recommended that longitudinal sections (parallel to the skin) be taken for examination (RITA, 2003; Hvid et al 2010). In the OECD validation of the TG 407 assay, the male mammary gland in particular, was pointed out as sensitive to oestrogenic substances and some histopathological guidance is included in GD 106 Part 4 (OECD, 2009). In TG 443, evaluation of the mammary gland may be carried out on both adults and weanlings. The following footnote is included in the TG (paragraph 69): “Research has shown the mammary gland, especially in early life mammary gland development, to be a sensitive endpoint for oestrogen action. It is recommended that endpoints involving pup mammary glands of both sexes be included in this Test Guideline, when validated”. Histopathological guidance on adult mammary glands is described in GD106 part 4. As described in GD106 part 4, it is important in both cases for the pathologist to be aware of the typical sexual dimorphism of the rat mammary gland (Cardy, 1991) as well as the pattern of morphologic changes associated with xenobiotic-induced hormonal perturbations.

83. Development of the terminal end buds into differentiated structures is of particular interest. A positive association has been reported between mammary tumour incidence or multiplicity and increased numbers of terminal end buds relative to controls in post-pubertal control and chemically exposed rodents that have been challenged with a carcinogen (Brown et al., 1988; Russo and Russo, 1996) or allowed to potentially form spontaneous mammary lesions (Padilla-Banks et al., 2006; Vandenberg et al., 2008).

Neurology: Assessment of potential developmental neurotoxicity in Cohort 2

84. Neurohistopathology and brain measurements are performed at weaning (cohort 2B) and post weaning (cohort 2A). Guidance on how to perform these examinations can be found in OECD, 2004; OECD, 2007; OECD, 2008; Tyl et al., 2008. Particular attention may need to be paid to sexually dimorphic structures.

Immunology: Measurement of IgM and IgG responses to assess potential immunotoxicity in Cohort 3

85. Inclusion of Cohort 3 in TG 443 requires determination of the primary and secondary Immunoglobulin M (IgM) and Immunoglobulin G (IgG) antibody response to a T-cell dependent antigen. The preferred antigen to be used is Keyhole Limpet Hemocyanin (KLH), consistent with current immunotoxicity testing procedures (Wang et al., 2018; Gore et al., 2004). The immunization protocol is specified in the TG 443 (2025).

86. Based on a review of TG 443 studies submitted within the context of the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) Regulation (Vandebriel et al., 2024), and as confirmed by members of the OECD ad hoc immunotoxicity expert group in 2024, KLH is the conventional and preferred antigen used at present. Whereas other routes of immunization, other antigens, and other doses of antigens could in principle lead to proper immunization of the animals, a harmonized protocol facilitates comparison across studies. The knowledge and guidance on the dose to be used for other antigens besides KLH (and the appropriate routes of immunization) are less developed and specified (Wang et al., 2018), which may result in a potential risk of animals not being immunized correctly.

87. Both SRBC and KLH are T-dependent antigens that can be used to stimulate IgM and IgG antibody production in the Cohort 3 offspring. The TG 443 (2025) specifies that if sheep red blood cells (SRBC) are required as the antigen by a regulatory authority, the detailed immunization protocol should be agreed upon with that authority.

88. KLH has the advantage of being stable for extended periods when stored frozen as dried powder. KLH should have certified and consistent batch-to-batch water solubility. Following immunization, KLH-specific IgM and IgG responses are to be measured in serum samples for example by ELISA. Primary responses generally peak at approximately five days for IgM and at 14 days for IgG following intravenous immunization (Wang et al., 2018). Consequently, IgM should be measured five days after both primary and secondary immunization(s), and IgG should be measured 14 days after the primary immunization and five days after the secondary immunization (see figure 2 in TG 443). If the peak responses for IgM and/or IgG differ from the time points indicated above (e.g. due to strain-related differences), samples should be collected at the appropriate peak-response day; however, supporting evidence for the alternative sampling time must be provided.

89. The PFC approach has not been applied within Cohort 3 of the TG 443. Consequently, references to the PFC approach were removed in the updated version of the TG 443. Furthermore, the PFC method is not compatible with the revised testing scheme as specified in the 2025 version of the TG 443.

90. The antigenic potency of SRBC from different donors may vary, and therefore, a reliable source of immunogenic SRBC should be established. SRBC also have a relatively short shelf life and cannot be frozen for later use. Due to these issues with SRBC, laboratories now also use KLH rather than SRBC to stimulate a humoral response.

91. Data generated with either antigen has been generally accepted by regulatory authorities worldwide. Current US EPA testing guidelines for assessment of immunotoxicity in adult animals (OPPTS 870.7800) specifies the use of SRBC, although this policy is subject to revision as experience with KLH increases.

92. The developmental immunotoxicity literature clearly supports evaluation of the antibody response to a T-dependent antigen as a means to detect adverse immune system effects. The first edition of this guidance document, published in 2013, indicated that both the SRBC and KLH antigens are suitable for developmental immunotoxicity testing in Cohort 3 animals, subject to each laboratory optimizing the chosen assay prior to use in testing. Optimisation of the immunisation protocol should include determining the peak day of antibody production or serum titres, as well as the optimum dose that induces a significant IgM and IgG response while retaining the ability to detect mild to moderate suppression of the antibody response (e.g., White et al., 2007). Any deviations from the standard practices specified in TG 443 should be justified. In view of the advantages of KLH, as outlined above, and this preference is reflected in the updated of the TG 443 (2025).

93. Laboratories testing the DIT Cohort of the EOGRTS should have sufficient experience with immune function assays to ensure accurate and consistent results. A positive control group exposed to a known immunosuppressant (i.e., cyclophosphamide, cyclosporin, dioctyltinoxide, di-n-octyltindichloride) is useful in the interpretation of the results or verification of the assay sensitivity (Loveless et al, 2007). Positive control experiments, including exposure to two doses of e.g. cyclophosphamide, adjusted to 90-100% as well as about 50% inhibition and a vehicle control, are performed every 6 months, which is well accepted by authorities. Such positive control studies could be done in the context of the EOGRTS, but may also be performed in another context, however such positive control study needs to be provided by the same facility performing the T-cell dependent antibody (TDAR) assay. Information on the performance of the assay and how this was assessed is included in the study report.

5 SECTION 5: INTERPRETATION OF DATA -SPECIFIC ISSUES WHERE FURTHER GUIDANCE IS PROVIDED

94. In this section, some guidance is provided on the interpretation of specific endpoints. It is stressed, however, that interpretation should be based on the study as a whole, as many endpoints are interconnected. Guidance on data interpretation of reproductive endpoints is also provided in GD 43 (OECD, 2026). Guidance on assessment of endocrine disruption in the EOGRTS is also provided in GD 150 (OECD, 2018).

Reproductive performance

95. Reproductive performance is the ability of male and female animals to mate successfully and produce viable offspring. The major indices usually determined are: male and female mating indices, male and female fertility indices, gestation length, gestation index and survival index. These should be reported in TG 443. Calculation of these indices and discussion on the interpretation of reproductive performance can be found in GD 43 (OECD, 2026, paragraph 180).

Anogenital distance

96. AGD is influenced by the body weight of the animal, and therefore, this factor should be considered in the evaluation of the data. The available data on pup size is typically body weight, and it has been shown that the cubic root of body weight is a better measure for normalising AGD data than body weight itself (Gallavan et al., 1999). Therefore, AGD normalisation using the cube root of body weight is recommended in GD 43 (OECD, 2026, paragraph 165). Normalisation can be performed either by dividing the individual AGD values by the corresponding cube root of body weight or by including body weight as a covariate in the statistical analysis (OECD, 2026). There are pros and cons to both methods. Dividing AGD by the cube root of body weight is a simple and transparent method for normalising AGD data, but it may yield skewed results because the relationship between AGD and body weight is not necessarily linear. Using the cube root of body weight as a covariate is statistically more appropriate for accounting for body weight's influence on AGD, but it can also yield less transparent results. In most cases, the two methods yield outcomes that are quite similar; therefore, both methods of presenting data from OECD TG 443 are acceptable. Decreased AGD in male rats is a hallmark of anti-androgenic substances (Gray et al., 1999; Hass et al., 2007; Schwartz et al 2019). A review by Wise (2024) highlights that AGD measurements can differ substantially depending on methodological factors, such as whether microscopy or callipers are used and how the measurement point on the

genital tubercle is defined (e.g., top, middle, or base). Harmonisation of measurement methods is therefore essential to ensure accuracy, consistency, and reproducibility across studies. A statistically significant change in AGD that cannot be explained by the size of the animal indicates an adverse effect of exposure and should be considered when setting the NOAEL (OECD, 2026).

Nipple/areola retention

97. Increased nipple/areola retention in male rats is one of the effects associated with fetal exposure to anti-androgenic substances (Christiansen et al, 2010; Schwartz et al. 2021). A statistically significant change in nipple/areola retention should be evaluated similarly to an effect in AGD, as both endpoints indicate an adverse effect of exposure and should be considered in setting the NOAEL. In contrast to anogenital distance, there appears to be no need to consider body weight in analysing the data. While fetal androgen signalling typically causes the areola to regress in control males, some biological variability exists, and it is expected that at least one male rat out of the hundreds examined across a TG 443 study will show areola retention (Schwartz et al., 2021). Therefore, unless the performing test facility can document its ability to detect this endpoint in male offspring, studies showing zero areolae in all examined male offspring should be interpreted with caution. To ensure accurate and unbiased assessments, the recommendations presented in paragraphs 59-61 should be followed.

Follicle count

98. The information on follicle counts should be examined in conjunction with the other endpoints evaluated in the F1 females. A change in follicle number may be apparent prior to a change in organ weight or histopathology (Muhammad et al., 2009) and may reflect an adverse effect on reproduction (Sanbuissho et al., 2009). A change in follicle count could indicate either direct oocyte toxicity, or an effect on the granulosa or thecal cells that alters the paracrine control of folliculogenesis.

The influence of body weight changes on endocrine organ weights

99. Organ weights are key endpoints in TG 443, but reductions in body weight may confound the interpretation of organ weight changes. Terminal body weight is generally used as a covariate during statistical analysis of organ weights and weights adjusted for covariance with terminal body weight may be shown in addition to absolute weights. Alternatively, “relative weight” (organ weight g/100 g body weight) may also be used. The disadvantage of reporting relative organ weight is that a simple relationship between organ weight and body weight is assumed, and this may not be correct. Organ to brain weight ratio may also be used, as brain weight tends to be less sensitive than other organs to body weight changes. The relationship of organ weight with body weight varies according to organ, age and sex of the animal, and pregnancy status. Some guidance on the interpretation of organ weight data is provided in Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies (OECD 2002a) and Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies (OECD 2002b).

100. Several publications have examined the relationship between body weight changes on organ weight data in the rat (Scharer, 1977; Chapin et al, 1993; Levin et al, 1993; Keenan et al, 1994; Seki et al, 1997; Odum et al, 2001; Marty et al, 2003; Carney et al, 2004; Terry et al, 2005; Laws et al, 2007). In these publications, body weight and organ weight reductions were achieved by the use of feed restriction or different proprietary diets. Data analyses of the influence of body weight on organ weights provide additional guidance on the interpretation of organ weight data relevant to TG 443. A summary of the findings from these studies, as well as detailed numeric results from these analyses, are

presented and tabulated in Annex 2. The data demonstrate that the relationship between organ weight and terminal body weight varies according to organ. For example, whilst a simple correlation appears to exist for organs such as liver and kidney, others such as testes and epididymides are less affected by body weight reduction unless it becomes substantial (e.g. 40%).

Measurements of vaginal opening and first vaginal oestrus

101. Onset of vaginal opening (VO) and age at first vaginal oestrus (defined as the predominance of cornified cells in a vaginal smear) are determined in the EOGRTS as measures of the onset of puberty. It is recommended that all females in Cohort 1 A of each group continue to be smeared once daily until the smear has progressed to oestrus and then metestrus or a maximum of three days of oestrus. This procedure is to confirm that the animals are starting to cycle normally.

102. The information obtained from the measurement of VO and first vaginal oestrus can be useful for determining how a test chemical influences the pubertal process. By evaluating at minimum, VO and the first day of oestrus, one can identify a disruption of this process and obtain information about the potential adverse reproductive effects of the test substance. For example, if the test chemical possesses estrogenic properties, exposure of the developing female will likely cause a significant advancement of the age of VO, but not necessarily advance first ovulation (e.g., Rodriguez et al., 1993). A similar dissociation between these two parameters has been reported following prepubertal androgen exposure and has been linked to the presence of aromatase activity in the vaginal epithelium of immature rats (Mathews et al., 1987; Lephart et al., 1989) suggesting that the androgen, after local conversion to oestradiol (Gupta et al., 1986; Rangaraj and Gupta, 1997), is acting directly to induce vaginal canalisation.

103. In most cases, environmental oestrogens will cause early VO and a pattern of persistent vaginal oestrus, (i.e., pseudo-precocious puberty) which may or may not continue as the animal matures. Thus, evaluating the first vaginal oestrus following VO will provide information as to whether there are group/dose differences in the timing of these two events that would signal an abnormal progression through puberty. As indicated above, first oestrus may be affected in time proportional to the appearance of VO, or the two may be disconnected, indicating independent alterations in response to a test chemical within the vagina and the hypothalamic-pituitary control of first ovulation at puberty (Firlit and Schwartz, 1977). Caution should, however, be exercised as vaginal smears represent a single point measurement within a dynamic oestrous cycle. To enhance consistency and support interpretation, smears should be collected at the same time of day (e.g. early morning hours) to minimise variability associated with circadian-driven changes. This data should also be interpreted alongside histopathological analysis of the uterus, cervix and vagina. GD 106 (OECD, 2009a) Part 5 describes the preparation, reading and reporting of vaginal smears and Part 3 describes the morphological changes during the oestrous cycle.

104. Both VO and first oestrus can be influenced by body weight and condition (potential consequences of toxicity to the dam during gestation and/or lactation). Mothers exposed to high doses of a test material often have offspring with lower body weights. This will affect offspring maturation and hence the pubertal process (Goldman et al., 2000). Thus, any treatment related differences in body weight should be considered in the interpretation and statistical methods used to analyse the chemical's effect on VO (i.e. covariate analysis). These considerations are important for interpreting VO and BPS results. TG 443 (paragraph 47) states that 'Sexual maturity of F1 animals is compared to physical development by determining age and body weight at BPS or VO for males/females respectively'. This means that body weight at the time of acquisition of VO and BPS should be reported and used to interpret data on sexual maturation. TG443 (paragraph 80) further states that 'VO and BPS data should also be reported together with body weight measurements on all pups on the same PND day (e.g. PND

21, 28 or 35)'. Sexual maturation data should therefore be interpreted in relation to body weight data on the day of VO/BPS acquisition and to body weight changes prior to puberty. Reporting and evaluating these parameters together enable determination of whether systemic toxicity or test substance–induced developmental delays before puberty underlie any changes in sexual maturation, or whether the observed effects are more consistent with an endocrine mode of action.

105. In order to increase the sensitivity of the test, examination of data on VO and BPS should be performed on the required number of animals (aiming at 50 pups/sex/group) in Cohort 1A, Cohort 1B and Cohort 3 (or 1C), as described in paragraph 12 of this guidance. Statistical analysis should be conducted on the combined dataset (i.e. pooling results from all relevant cohorts) and should use the litter as the statistical unit.

Influence of the oestrous cycle on female reproductive organ weights

106. The weight of the uterus (and vagina, although this is not specified in TG 443) in adult rats may vary 3 to 4-fold depending upon the stage of the oestrous cycle (GD 43, OECD 2026, paragraph 186). Uterine weights may therefore have a high variance unless they are related to the stage of the oestrous cycle at the time of evaluation. Compounds that cause loss of cyclicity (e.g. oestrogen antagonists, steroidogenesis inhibitors) may cause the uterus to become atrophic and the reduction in weight along with reduced variance of mean uterine weight may be indicative of this effect.

Evaluation of developmental neurotoxicity

107. Guidance on interpretation of developmental neurotoxicity data can be found in several published documents (OECD, 2004; OECD, 2007; OECD, 2008; Crofton et al., 2008; Makris et al., 2008; Tyl et al., 2008) and is thus not detailed in this Guidance Document. The neurotoxicity testing in TG 443 aims to provide an initial assessment of neurotoxicity potential but does not include all facets of a complete DNT study. Thus, it is not intended as a replacement for a DNT study, nor is it appropriate to interpret the results from the TG 443 DNT assessment as a replacement for conducting a DNT study where one may be required. Interpretation of TG 443 DNT test results should take into account available information on mechanisms of action, toxicokinetics, maternal toxicity and potential indirect effects on offspring, as well as any available data on neurotoxic effects of the specific test chemical. Such results may indicate additional targeted DNT testing may be required.

Evaluation of developmental immunotoxicity

108. The current immunotoxicity endpoints included in TG 443 can detect modulation (suppression and stimulation) of the immune response. Lymphoid organ weights and histopathology, white blood cell parameters, and lymphocyte subset analysis (phenotyping) derived from Cohort 1A animals contributes to the weight of evidence for immunotoxicity (TG 443 paragraph 84) but are generally not regarded as having the predictive power of functional tests. Dose-related changes in these endpoints, in the absence of effects in antibody production, may be an indication that exposure affected cell mediated immunity; certain developmental immunotoxicants have been reported to suppress cellular immunity without affecting antibody production. Under these circumstances, if immune parameters of Cohort 1A represent the LOAEL for all TG 443 endpoints, it is recommended that at least the immunotoxicology portion of the study be investigated further. In addition, functional assessment of cell mediated immunity

(e.g., the delayed type of hypersensitivity response) is recommended to determine if cellular immune function is compromised.

109. Effects on immune function may be sex specific. Therefore, absence of an effect in one sex should not be interpreted as evidence against effects observed in the other sex. Similarly, cross generational comparisons of immune assessments linked to organ weights, clinical chemistry, haematology and histopathology, observed only in offspring (F1) and not seen in parental animals (F0) should not be disregarded, as developing animals are generally considered more sensitive than adults to immunotoxic effect (IPCS, 2012).

110. In some cases, changes in the functional test such as the TDAR may be the only findings observed. As the TDAR assay is considered to be the most predictive test for detecting immunotoxicity, lack of evidence in other immune parameters should not be used to disregard the positive findings in the TDAR (Luster e*t al,* 1992).

111. It is essential to ensure the proper functioning of the antigen used. For immunization to be considered valid, a robust antibody response i.e., an increase in the KLH-specific antibody production, must be observed in the negative control group. As noted by Gore et al., (2004) “similarly, all rats immunized with KLH (300 µg/kg) by the i.v. route tested positive for anti-KLH IgM and IgG antibodies […]”. If the negative control group does not exhibit a robust increase in the antibody production, the immunization is considered unsuccessful, and the results obtained from the test and positive control groups cannot be used. Sporadic low responses or non-responders may occur due to interindividual variation; however, the number of animals not responding to immunization within a dose group should not exceed one animal per dose group.

112. Historical control data is critical for the interpretation and evaluation of TDAR responses, including KLH-specific IgM and IgG findings. In particular, TDAR data obtained from a positive control substance tested according to the same immunotoxicity protocol at the original facility can be used to assess the sensitivity and performance of the procedures employed (White, et al., 2007; Fukuyama et al., 2013).

113. When assessing observations related to immunotoxicity, it is important to recognise that inherent interindividual variation occurs in both humans and rodents (rats). This variability may be further influenced by environmental factors (Zimmerman and Curtis, 2019; Nakamura, et al., 2024). Functional immunotoxicity studies in rats have demonstrated considerable interindividual variability in antibody production following TDAR assays. For example, a study evaluating primary and secondary responses to KLH in different rat strains (Sprague-Dawley, Wistar, Fischer, and Lewis) reported considerable variability in KLH-specific IgM and IgG production (Kawai et al., 2013, Ai et al., 2014).

114. Due to this inherent biological variability of the immune system, it is important to evaluate the data beyond mean values and statistical significance. Individual data should be inspected to identify potential outliers that may skew mean values. In such cases, exclusion of such outliers may be appropriate supported by an appropriate statistical test. Various methodologies exist to identifying outliers, and it is essential to ensure that the statistical method is applicable to the dataset. Given the considerable variability of the data (CVs), in addition to the consideration of potential outliers, biologically relevant suppression or stimulation of antibody production with clear dose response may be observed even in the absence of statistical significance. Comparison of individual baseline (pre-immunization) values with post-immunization values and reporting of individual fold changes can provide important information on biologically relevant changes. Therefore, effects lacking statistical significance should not be disregarded and could be considered relevant for the determination of hazardous properties.

115. Although not required under the TG 443, if Cohort 1 or 3 show evidence ofimmunotoxicity , additional investigation into the underlying mechanism of immunotoxicity could be conducted. Selection

of endpoints will depend on the outcomes seen in the Cohorts 1 and 3 and may include appropriate in-depth studies. In certain cases, it may be helpful to use defined animal disease models (resistance to infection, allergic hypersensitivity, autoimmune disease) or new approach methodologies to understand the relationship between altered immune function and development of actual disease.

116. Detailed advice on interpretation of immunotoxicity data and the use of immunotoxicity data for risk assessment, including suppression and unintended stimulation of immune function, is available in the document Guidance for Immunotoxicity Risk Assessment for Chemicals (IPCS, 2012).

Weight of evidence evaluation

117. The results from all endpoints of the EOGRTS should be considered together in a weight of evidence (WoE) evaluation. It is not advisable to view the endpoints in isolation as they may be inter-related. Advice on evaluation of results for reproductive toxicity and neurotoxicity is also provided in GD 43 (OECD 2026), GD 20 (OECD 2004) and Tyl et al (2008). Elements that should be given special consideration include:

- Relationship between the dose and the presence, incidence, and severity of effects (see TG 443 paragraph 78)
- Deviations from the protocol and their potential impact on the study,
- Quality of the data, including variability of results,
- Relevant TK data such as placental transfer and milk excretion,
- Sex-specific effects,
- Possible consequences of systemic toxicity on reproductive, neurological, immunological and endocrine endpoints (see paragraph below).

118. The results of any dose-ranging or other studies conducted at the same or at higher dose levels (see paragraph 26) may also be relevant to the assessment of any effects of systemic toxicity. The effect of reduced body weight (from food restriction) on reproductive function, pubertal development, fertility, organogenesis and immunological endpoints have been described in (Chapin et al, 1993; Seki et al, 1997; Marty et al, 2003; Carney et al, 2004; Terry et al, 2005; Fleeman et al, 2005; Laws et al, 2007). The influence of body weight on organ weight is discussed in paragraphs 93-94 and Annex 2. If possible, treatment-related changes resulting from body weight loss should be distinguished from specific toxic effects. When results from the study are evaluated, there should be consistency of response between interrelated endpoints and, ideally, evidence of a dose-response relationship. Results that are clearly positive or negative should be distinguished from those that are marginal. Irreversible effects, such as developmental abnormalities, may be distinguished from potentially reversible effects. The overall conclusions drawn from the study should be based on the results in their entirety.

119. The interpretation of the results of the EOGRTS in the context of other studies on the substance should be considered in a wider weight of evidence evaluation. The US EPA (2011) has defined WoE as “the process for characterizing the extent to which the available data support a hypothesis that an agent causes a particular effect”. This process is described as involving a number of steps starting with assembling the relevant data, evaluating that data for quality and relevance followed by an integration of the different lines of evidence to support conclusions concerning a property of the substance. WoE is not a simple tallying of the number of positive and negative studies; it relies on professional judgment. Thus, transparency is important to any WoE analysis. A WoE assessment explains the kinds of data available, how they were selected and evaluated, and how the different lines of evidence fit together in drawing conclusions. The significant issues, strengths, and limitations of the data and the uncertainties that deserve serious consideration are presented, and the major points of interpretation highlighted (US

EPA 2011). TG 443 paragraphs 80-84 also provide some advice on interpretation of the EOGRTS and stress the importance of considering all available data, including physico-chemical data, TK data, results from QSAR models and result from structural analogues of the substance tested. A number of approaches to WoE evaluations have been published that are useful because they provide a structured framework for the analysis (Boobis et al, 2008; ECHA, 2010; Bogert et al 2011; US EPA 2006, 2011). The IPCS mode of action framework for noncancer endpoints (Boobis et al, 2008) provides a series of steps to determine whether the WoE, based on experimental observations, is sufficient to establish mode of action (MoA). The approach is based on the Bradford Hill criteria and includes considerations of biological plausibility and coherence, strength, and consistency of the body of evidence. This framework then allows analysis of the human relevance of the findings. This MoA approach has been used in a series of case studies including reproductive toxicants, developmental toxicants and developmental neurotoxicants (Foster, 2005; Wiltse, 2005; Zoeller and Crofton, 2005). Whichever WoE approach is used, transparency and scientific rigour are key to ensure adequate evaluation of the data.

Overview of the literature during the initial drafting of this GD

120. At the time of writing this GD initially, only one full EOGRTS has been published where all endpoints included in TG 443 were evaluated. This was conducted on the fungicide vinclozolin (Schneider et al, 2011). A “cut down” version of the EOGRTS has also been performed on the plastic stabiliser di-n-octyltin dichloride (DOTC). In this study, animal numbers and treatment regime were very similar to TG 443 but endpoints were focussed on DIT and did not include all parameters for reproductive toxicity or any parameters for DNT (Tonk et al, 2011). The objective of both studies was the testing of the protocol and endpoints of the EOGRTS, rather than assessment of the substances. Both studies were carried out before publication of TG 443 and therefore there are inevitable differences compared to the current study design. Administration of both vinclozolin and DOTC was via the diet.

121. DOTC had previously been shown to cause thymus atrophy and suppression of thymus-dependent immune responses in the rat. In the EOGRTS, there were no treatment-related effects on fertility or reproductive performance in the P generation. In male and female F1 offspring, there was reduced pup viability at PND 4, some isolated effects on body weight but no effects on sexual maturation. In this study, effects on immune parameters in F1 offspring were assessed at 3, 6 and 10 weeks of age. No effects were seen at 3 weeks but at 6 and 10 weeks, effects were seen on lymphocyte populations in the spleen at the highest dose level (30 mg/kg bw/day). Thymus weight was reduced at 6 weeks of age. In TG 443, analysis of splenic lymphocyte subpopulations is conducted at 8 weeks of age (Cohort 3) and 13 weeks of age (Cohort 1A), indicating that at these time points, the endpoints are sensitive to immunotoxicants. Other parameters of immunotoxicity were included in Tonk et al (2011) but have not been included in TG 443.

122. Vinclozolin is a well-studied reproductive/developmental toxicant where the mode of action is via androgen receptor antagonism. Endpoints associated with development of the male reproductive system in the F1 generation (such as genital abnormalities, reduced AGD, increased nipple retention, delayed preputial separation, reduced reproductive organ weight) were all altered. Some effects on the female reproductive system in the F1 generation were observed, such as reduced age at vaginal opening and oestrous cycle disturbance. Reproductive performance in the P generation was unaffected. There was an original intention to breed a second generation, but this was not carried out as all the high dose animals had hypospadias and were judged to be unable to breed. There were also no effects on endpoints of DNT or DIT. The sensitivity of the study was similar to that of previous two-generation reproduction toxicity studies as a very similar NOAEL (4 mg/kg bw/day) was obtained. The most sensitive endpoints were male AGD, incidence of nipple/areola incidence in males and preputial separation where clear effects were seen at the mid dose level (20 mg/ kg bw/day).

123. The authors commented that the protocol was a significant challenge, that the complexity should not be underestimated and that flexibility was essential. Each post-weaning cohort represented a full study in terms of logistics and resource. PND 21 was the study day with the highest workload, which included:

- Clinical observations, food consumption and body weight in 200 parents.
- Clinical observations, body weight, AGD, nipple retention in about 1200 offspring.
- Final selection of cohorts.
- Necropsy of 200 parents including macroscopic pathology, implants, sperm analysis, organ weights, tissue preservation.
- Necropsy including macroscopic pathology of about 520 surplus weanlings.
- Perfusion fixation, macroscopic pathology, organ weights, preparation of central and peripheral nervous system tissue in 80 weanlings.

124. It was concluded, however, that the study is technically feasible for laboratories with sufficient experience. Conducting the study in blocks may be envisioned in some circumstances (see paragraph 54).

ABBREVIATIONS

ADME	Absorption, Distribution, Metabolism, Excretion
AGD	Anogenital distance
bIVF	Bovine *in vitro* fertilisation assay
bIVM	Bovine *in vitro* maturation assay
bIVP	Bovine *in vitro* pre-implantation assay
BMD	Benchmark dose
BMDL	Lower confidence level of the benchmark dose
BMR	Benchmark response
BPS	Balano-preputial separation
BW	Body weight
CF	Conceptual framework
CL	Corpora lutea
CV	Coefficient of variation
DIT	Developmental immunotoxicity
DNT	Developmental neurotoxicity
DOTC	2,3,7,8-tetrachlorodibenzo-p-dioxin
ECHA	European Chemical Agency
EFSA	European Food Safety Authority
ELISA	Enzyme-linked immunosorbent assay
EOGRTS	Extended One Generation Reproductive Toxicity Study (OECD TG 443)
EPA	United States Environmental Protection Agency
EST	Embryonic stem cell test
F	Female
F0	Parental generation
F1	First filial generation
F2	Second filial generation
FC	Food consumption
FR	Feed restricted
GD	Guidance document
GMP	Good Manufacturing Practice
HPA	Hypothalamus:pituitary:adrenocortical axis
HPG	Hypothalamus:pituitary:gonad axis
HPT	Hypothalamus:pituitary:thyroid axis
ICH	International Conference on Harmonisation
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IPCS	International Programme on Chemical Safety
KLH	Keyhole limpet haemocyanin

LOAEL	Lowest observed adverse effect level
M	Male
MoA	Mode of action
NA	Not applicable
NOAEL	No Observed Adverse Effect Level
NR	Nipple/Areola retention
OECD	Organisation for Economic Cooperation and Development
P	Parental
Para/s	Paragraph/s
PBTK	Physiologically Based Toxicokinetics
PFC	Plaque forming cell
PND	Postnatal day
POD	Point of departure
PPAR	Peroxisome proliferator-activated receptor
QSAR	Quantitative Structure Activity Relationship
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals
SOP	Standard Operating Procedure
SRBC	Sheep red blood cells
T4	Thyroxine (thyroid hormone)
TDAR	T-cell dependent antibody response
TG	Test Guideline
TK	Toxicokinetics
TSH	Thyroid Stimulating Hormone
US EPA	United States Environmental Protection Agency
VO	Vaginal opening
WoE	Weight of Evidence

REFERENCES

Adler S, Basketter D, Creton S, Pelkonen O, van Benthem J, Zuang V, Andersen KE, Angers-Loustau A, Aptula A, Bal-Price A, Benfenati E, Bernauer U, Bessems J, Bois FY, Boobis A, Brandon E, Bremer S, Broschard T, Casati S, Coecke S, Corvi R, Cronin M, Daston G, Dekant W, Felter S, Grignard E, Gundert-Remy U, Heinonen T, Kimber I, Kleinjans J, Komulainen H, Kreiling R, Kreysa J, Leite SB, Loizou G, Maxwell G, Mazzatorta P, Munn S, Pfuhler S, Phrakonkham P, Piersma A, Poth A, Prieto P, Repetto G, Rogiers V, Schoeters G, Schwarz M, Serafimova R, Tähti H, Testai E, van Delft J, van Loveren H, Vinken M, Worth A, Zaldivar JM (2011). Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. Arch Toxicol. 2011 May; 85(5):367-485. doi: 10.1007/s00204-011-0693-2.

Ai W, Huo Y, Liu X, Liu F, Zhou X, Miao Y, Jiang H, Zhang L, Shen L, Piao J, Bo L (2014). Relative sensitivities of TDAR, cytokine production, and immunophenotyping assays in immunotoxicity assessment. Toxicology Research, Volume 3, Issue 6, November 2014, Pages 465–473

AXLR8 (2010). AXLR8 Alternatives Testing Strategies 2010 Progress Report. www.http://www.axlr8.eu/axlr8-2010-progress-report.pdf

Blystone, C. R., Kissling, G. E., Bishop, J. B., Chapin, R. E., Wolfe, G. W., and Foster, P. M. (2010). Determination of the di-(2-ethylhexyl) phthalate NOAEL for reproductive development in the rat: importance of the retention of extra animals to adulthood. Toxicol Sci 116(2), 640-646.

Boettger-Tong, H., Murthy, L., Chiappetta, C., Kirkland, J. L., Goodwin, B., Adlercreutz, H., Stancel, G. M., and Makela, S. (1998). A case of a laboratory animal feed with high estrogenic activity and its impact on *in vivo* responses to exogenously administered estrogens. Environ Health Perspect. 106(7), 369-373.

Bogert, C.J., Mihaich, E.M., Quill, T.F., Marty, M.S., Levine, S.L., Becker, R.A. (2011) Evaluation of EPA‘s Tier 1 endocrine screening battery and recommendations for improving the interpretation of screening results. Regulatory Toxicology and Pharmacology 59:397-411.

Boobis, A.R., Doe, J.E., Heinrich-Hirsch, B., Meek, M.E., Munn, S., Ruchirawat, M,. Schlatter, J., Seed, J., Vickers, C. (2008). IPCS framework for analyzing the relevance of a noncancer mode of action for humans. Critical Reviews in Toxicology. 38(2):87-96.

Brown NM, Manzolillo PA, Zhang JX, Wang J, Lamartiniere CA.(1998). Prenatal TCDD and predisposition to mammary cancer in the rat. Carcinogenesis. Sep;19(9):1623-9.

Byczkowski JZ and Lipscomb JC (2001): Physiologically based pharmacokinetic modeling of the lactational transfer of methylmercury. Risk Anal., 21, 869-882.

Cardy RH (1991). Sexual dimorphism of the normal rat mammary gland. Vet Pathol 28: 139-145

Carney, E. W., Zablotny, C. L., Marty, M. S., Crissman, J. W., Anderson, P., Woolhiser, M., and Holsapple, M. (2004). The effects of feed restriction during in utero and postnatal development in rats. Toxicol Sci 82(1), 237-249.

Chang SC, Thibodeaux JR, Eastvold ML, Ehresman DJ, Bjork JA, Froehlich JW, Lau CS, Singh RJ, Wallace KB and Butenhoff JL. 2007. Negative bias from anolog methods used in the analysis of free thyroxin in rat serum containing perfluorooctanesulfonate (PFOS). Toxicology 234:21-33.

Chapin, R. E., Gulati, D. K., Barnes, L. H., and Teague, J. L. (1993). The effects of feed restriction on reproductive function in Sprague-Dawley rats. Fundam. Appl. Toxicol 20(1), 23-29.

Christian MS, Trenton NA. 2003. Evaluation of thyroid function in neonatal and adult rats: The neglected endocrine mode of action. Pure and Applied Chemistry 75:2055-2068.

Christiansen, S., Boberg, J., Axelstad, M., Dalgaard, M., Vinggaard, A. M., Metzdorff, S. B., and Hass, U. (2010). Low-dose perinatal exposure to di(2-ethylhexyl) phthalate induces anti-androgenic effects in male rats. *Reprod. Toxicol* 30(2), 313-321.

Corley RA, Mast TJ, Carney EW, Rogers JM, and Daston GP (2003): Evaluation of physiologically based models of pregnancy and lactation for their application in children's health risk assessments. Crit Rev.Toxicol, 33, 137-211.

Creasy, D. M. (2003). Evaluation of testicular toxicology: A synopsis and discussion of the recommendations proposed by the Society of Toxicologic Pathology. Birth Defects Research (Part B) 68, 408-415.

Creton S, Saghir SA, Bartels MJ, Billington R, Bus JS, Davies W, Dent MP, Hawksworth GM, Parry S, Travis KZ (2012), Use of toxicokinetics to support chemical evaluation: Informing high dose selection and study interpretation, Regul Toxicol Pharmacol., 62(2):241-7.

Crofton KM, Foss JA, Hass U, Jensen K, Levin ED, Parker SP (2008) Undertaking Positive Control Studies as Part of Developmental Neurotoxicity Testing. A Report from the ILSI Research Foundation/Risk Science Institute Expert Working Group on Neurodevelopmental Endpoints. Neurotox Teratol 30, 266-87 (Jun 2007 Epub ahead of print).

Davies DT. 1993. Assessment of rodent thyroid endocrinology: Advantages and pit-falls. Comparative Haematology International 3:142-152.

Dixon D, Alison R, Bach U, Colman K, Foley GL, Harleman JH, Haworth R, Herbert R, Heuser A, Long G, Mirsky M. Nonproliferative and proliferative lesions of the rat and mouse female reproductive system. Journal of toxicologic pathology. 2014 Nov 29;27(3-4 Suppl):1S.

ECHA, (2023a) European Chemicals Agency "Evaluating results from 55 extended one-generation reproductive toxicity studies under REACH, Final report of the EOGRTS review project," Reference: ECHA-23-R-04-EN. ISBN: 978-92-9468-262-8. DOI: 10.2823/92503

ECHA (2023b): Guidance on information requirements and chemical safety assessment Chapter R.7c: Endpoint specific guidance (https://www.echa.europa.eu/documents/10162/17224/information_requirements_r7c_en.pdf/e2e23a98-adb2-4573-b450-cc0dfa7988e5)

EFSA (2022), Guidance on the use of the benchmark dose approach in risk assessment. *The EFSA Journal,* Volume 20, Issue 10. https://doi.org/10.2903/j.efsa.2022.7584

Elswick, B. A., Welsch, F., and Janszen, D. B. (2000). Effect of different sampling designs on outcome of endocrine disruptor studies. Reprod. Toxicol 14(4), 359-367.

Fenton, S.E., J.T. Hamm, L.S. Birnbaum, and G.L. Youngblood. 2002. Early adverse effects of prenatal exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) in the rat mammary gland *Toxicol. Sci*. 67: 63-74.

Firlit, MG, Schwartz, NB (1977). Uncoupling of vaginal opening and the first ovulation- An indication of an alteration in the pituitary-gonadal axis. Biol. Reprod. 16: 441-444.

Fleeman, T. L., Cappon, G. D., Chapin, R. E., and Hurtt, M. E. (2005). The effects of feed restriction during organogenesis on embryo-fetal development in the rat. Birth Defects Res B Dev. Reprod. Toxicol 74(5), 442-449.

Foster, P.M. (2005). Mode of action: Impaired fetal Leydig cell function—Effects on male reproductive development produced by certain phthalate esters. Crit. Rev. Toxicol. 35:713–719.

Fukuyama, T, Kosaka T, Hayashi K, Miyashita L, Tajima Y, Wada K, Nishino R, Ueda H, Harada T (2013). Immunotoxicity in mice induced by short-term exposure to methoxychlor, parathion, or piperonyl butoxide, Journal of Immunotoxicology, pages 150–159.

Gallavan, R.H. Jr, J.F. Holson, D.G. Stump, J.F. Knapp and V.L. Reynolds (1999), "Interpreting the Toxicologic Significance of Alterations in Anogenital Distance: Potential for Confounding Effects of Progeny Body Weights", Reproductive Toxicology, 13: 383-390.

Gibaldi, M., Perrier, D. (2009), Pharmacokinetics, Revised and Expanded, InformaHealthcare, New York, NY.

Goldman, J.M., Laws, S.C., Balchak, S.K., Cooper, R.L. and Kavlock, R.J. (2000), Endocrine disrupting chemicals: Prepubertal exposures and effects on sexual maturation and thyroid activity in the female rat. A review of the EDSTAC recommendations, Critical Reviews in Toxicology, 30: 135-196.

Gore ER, Gower J, Kuralli E, Sui JL, Enulat D, Herzyk DJ. (2004), Primary antibody responses to keyhole limpet hemocyanin in rat as a model for immunotoxicity evaluation. Toxicology. 197, 23-35

Gore, AC. (2008), Developmental programming and endocrine disruptor effects on reproductive neuroendocrine systems. Front Neuroendocrinol. 29(3), 358-374.

Gray, L.E., Jr. and Ostby, J.S. (1995), In utero 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) alters reproductive morphology and function in female rat offspring, Toxicology and pharmacology 133, 285-294.

Gray Jr L.E. and Gray C.L. (2006). Draft: White paper on statistical assessment of the value gained by sampling multiple F1 pups per litter. Not published. Circulated at the OECD meeting in Paris.

Gray Jr L.E. et al. (2004). Use of the laboratory rat as a model in endocrine disruptor screening and testing. ILAR journal. 45(4): 425-437

Gray, L. E., Jr., Barlow, N. J., Howdeshell, K. L., Ostby, J. S., Furr, J. R., and Gray, C. L. (2009). Transgenerational effects of Di (2-ethylhexyl) phthalate in the male CRL:CD(SD) rat: added value of assessing multiple offspring per litter. *Toxicol Sci* 110 (2), 411-425.

Gupta, PD, Khar, A, Vijayasaradhi, S (1986). Localization of estradiol receptors in rat vaginal epithelial cells. Ind. J. Exp. Biol. 24: 679-682.

Harper, M. J. (1964). Observations on amount and distribution of prenatal mortality in a strain of albino rats. J Reprod. Fertil. 7, 185-209.

Hass U, Scholze M, Christiansen S, Dalgaard M, Vinggaard AM, Axelstad M, Metzdorff SB, Kortenkamp A. Combined exposure to anti-androgens exacerbates disruption of sexual differentiation in the rat. Environ Health Perspect. 2007 Dec;115 Suppl 1(Suppl 1):122-8. doi: 10.1289/ehp.9360.

Heindel JJ, vom Saal FS. 2009. Role of nutrition and environmental endocrine disrupting chemicals during the perinatal period on the aetiology of obesity. Molecular and Cellular Endocrinology 304:90-96.

Herzyk DJ, Gore ER. 2004. Adequate immunotoxicity testing in drug development. Toxicol Lett. 149:115-22.

Hines RN (2008): The ontogenicity of drug metabolism enzymes and implications for adverse drug events. Pharmacology & Therapeutics 2008, 250-267.

Hotchkiss, A. K., Rider, C. V., Blystone, C. R., Wilson, V. S., Hartig, P. C., Ankley, G. T., Foster, P. M., Gray, C. L., and Gray, L. E. (2008). Fifteen years after "Wingspread"--environmental endocrine disrupters and human and wildlife health: where we are today and where we need to go. Toxicol Sci 105(2), 235-259.

Hvid H, Thorup I, Oleksiewicz MB, Sjögren I, Jensen HE (2011). An alternative method for preparation of tissue sections from the rat mammary gland. Exp Toxicol Pathol 63: 317-324

ICH Harmonised Tripartite Guideline (2005). International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use – Guideline on detection of reproductive and developmental toxicity for human pharmaceuticals S5(R3). Current Step 4 version including addendum dated 9 November 2000 incorporated in November 2005, adopted by the regulatory members February 2020.

ILSI (2003). Principles and Practices for Direct Dosing of Pre-weaning Mammals in Toxicity Testing and Research.. A report of the ILSI Risk Science Institute Expert Working Group on Direct Dosing of Preweaning Mammals in Toxicity Testing. ILSI Press, Washington DC.

IPCS (2012). International Programme on Chemical Safety of the World Health Organization. Guidance for Immunotoxicity Risk Assessment for Chemicals. http://www.who.int/entity/ipcs/methods/harmonization/areas/guidance_immunotoxicity.pdf

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.

Kawai R, Ito S, Aida T, Hattori H, Kimura T, Furukawa T, Mori K, Sanbuissho A, Kawada T (2013). Evaluation of primary and secondary responses to a T-cell-dependent antigen, keyhole limpet hemocyanin, in rats. Journal of Immunotoxicology, 10(1): 40–48 .

George, JD; Tyl, RW; Hamby, BT; Myers, CB; Marr, MC (2003), One generation extension study of vinclozolin and di-n-butyl phthalate administered by gavage on gestational day 6 to postnatal day 20 in CD (Sprague-Dawley) rats. EPA contract number 68-W-01-023.Work Assignment 2-10. http://www.epa.gov/endo/pubs/edmvs/onegenextensionfinalreport5803.pdf

Kato, H., Iwata, T., Katsu, Y., Watanabe, H., Ohta, Y., and Iguchi, T. (2004), Evaluation of estrogenic activity in diets for experimental animals using *in vitro* assay. J Agric. Food Chem 52(5), 1410-1414.

Keenan, K. P., Smith, P. F., Hertzog, P., Soper, K., Ballam, G. C., and Clark, R. L. (1994), The effects of overfeeding and dietary restriction on Sprague-Dawley rat survival and early pathology biomarkers of aging. Toxicol Pathol. 22(3), 300-315.

Kucheryavenko O, Angeli K, Bonnomet V, Ashraf MW, Bichlmaier I, Simanainen U. Screening for the ovarian toxicity of industrial chemicals–ovarian follicles and corpora lutea counts in rats, an inter-and intra-laboratory comparison. Critical Reviews in Toxicology. 2026 Jun 8:1-20.

Ladics GS. (2007), Primary immune response to sheep red blood cells (SRBC) as the conventional T-cell dependent antibody response (TDAR) test. J. Immunotoxicol. 4, 149-152.

Lanning L.L., Creasy D.M., Chapin R.E., Mann P.C., Barlow N.J., Regan K.S., Goodman D.G. (2002), Recommended Approaches for the Evaluation of Testicular and Epididymal Toxicity. Toxicologic Pathology 30, 507-520.

Latini S (2008), PhD Thesis PACAP function during rat folliculogenesis. Dottorato di Ricerca in Scienze e Technologie Cellulari XX cycle. Rome, December 2008. https://iris.uniroma1.it/retrieve/e3835315-c9c0-15e8-e053-a505fe0a3de9/TESI%20DOTTORATO%20-%20STEFANIA%20LATINI.pdf
Laws, S. C., Stoker, T. E., Ferrell, J. M., Hotchkiss, M. G., and Cooper, R. L. (2007). Effects of altered food intake during pubertal development in male and female wistar rats. Toxicol Sci 100(1), 194-202.

Lee SK, Ou YC, and Yang RS (2002): Comparison of pharmacokinetic interactions and physiologically based pharmacokinetic modeling of PCB 153 and PCB 126 in nonpregnant mice, lactating mice, and suckling pups. Toxicol Sci., 65, 26-34.

Lephart, ED, Mathews, D, Noble, JF, Ojeda, SR (1989). The vaginal epithelium of immature rats metabolizes androgens through an aromatase-like reaction: Changes during the time of puberty. Biol. Reprod. 40: 259-267.

Levin, S., Semler, D., and Ruben, Z. (1993). Effects of two weeks of feed restriction on some common toxicologic parameters in Sprague-Dawley rats. Toxicol Pathol. 21(1), 1-14.

Li T, Boberg J, Johansson HK, Di Nisio V, Christiansen S, Svingen T, Damdimopoulou P. Quantitative analysis of ovarian surface photographs as a tool for assessment of chemical effects on folliculogenesis and ovulation in rats. Reproductive Toxicology. 2023 Aug 1;119:108416.

Long X, Steinmetz R, Ben-Jonathan N, Caperell-Grant A, Young PCM, Nephew KP, Bigsby RM. Strain differences in vaginal responses to the xenoestrogen bisphenol A. Environ Health Perspect 2000; 108:243–247.

Loveless SE, Ladics GS, Smith C, Holsapple MP, Woolhiser MR, White KL Jr., Musgrove DL, Smialowicz RJ, Williams W. (2007). Interlaboratory study of the primary antibody response to sheep red blood cells in outbred rodents following exposure to cyclophosphamide or dexamethasone. J. Immunotoxicol. 4, 233-238.

Lucas JN, Rudmann DG, Credille KM, Irizarry AR, Peter A, Snyder PW (2007). The rat mammary gland: morphologic changes as an indicator of systemic hormonal perturbations induced by xenobiotics. Toxicol Pathol 35: 199-207

Luster, MI et al., 1988. Development of a testing battery to assess chemical-induced immunotoxicity: National Toxicology Program's guidelines for immunotoxicity evaluation in mice. Fundam. Appl. Toxicol. 10, 2-19.

Luster MI, Portier C, Pait DG, White Jr KL, Gennings C, Munson AE, Rosenthal GJ (1992). Risk assessment in immunotoxicology. I. Sensitivity and predictability of immune tests. Fundam Appl Toxicol.

18(2):200-10.

Mangelsdorf, I., J. Buschmann and B. Orthen (2003), “Some Aspects Relating to the Evaluation of the Effects of Chemicals on Male Fertility”, Regulatory Toxicology and Pharmacology, 37, 356-369.

Makris SL, Raffaele K, Allen S, Bowers Wj, Hass U, Alleva E, Calamandrei G, Sheets L, Amcoff P, Delrue N, Crofton KM (2008), A retrospective performance assessment of the developmental neurotoxicity study in support of OECD test guideline 426. Env Health Perspect Aug 117, 17-25

Markaverich, B. M., Alejandro, M. A., Markaverich, D., Zitzow, L., Casajuna, N., Camarao, N., Hill, J., Bhirdo, K., Faith, R., Turk, J., and Crowley, J. R. (2002). Identification of an endocrine disrupting agent from corn with mitogenic activity. Biochem. Biophys. Res Commun. 291(3), 692-700.

Marty, M. S., Johnson, K. A., and Carney, E. W. (2003). Effect of feed restriction on Hershberger and pubertal male assay endpoints. Birth Defects Res B Dev. Reprod. Toxicol 68(4), 363-374.

Mathews, D, Andrews, WW, Parker, R Jr, Ojeda, SR (1987). A role for aromatizable androgens in female rat puberty. Biol. Reprod. 36: 836-843.

Mathieu P, Poirier P, Pibarot P, Lemieux I, Després JP. 2009. Visceral obesity the link among inflammation, hypertension, and cardiovascular disease. Hypertension 53:577-584.

McCoy AT, Bartels MJ, Rick DL, Saghir SA (2012), TK Modeler version 1.0, a Microsoft® Excel®-based modeling software for the prediction of diurnal blood/plasma concentration for toxicokinetic use. Regul Toxicol Pharmacol., 63(2):333-43.

McFadden LG, Bartels MJ, Rick DL, Price PS, Fontaine DD, Saghir SA. (2012), Statistical methodology to determine kinetically derived maximum tolerated dose in repeat dose toxicity studies, Regul Toxicol Pharmacol., 63(2):344-51.
Merlot E., Couret D., and Otten W. 2008. “Prenatal stress, fetal imprinting and immunity”, *Brain, Behaviour, and Immunity* 22:42-51

Moser V.C., Walls I., Zoetis T. (2005) Direct dosing of preweaning rodents in toxicity testing and research: Deliberations of an ILSI RSI expert working group. International Journal of Toxicology 24(2): 87-94.

Muñoz-de-Toro M, Markey CM, Wadia PR, Luque EH, Rubin BS, Sonnenschein C, Soto AM (2005), Perinatal exposure to bisphenol-A alters peripubertal mammary gland development in mice, Endocrinology 146, 4138.

Muhammad, F.S., Goode, A.K., Kock, N.D., Arifin, E.A., Cline, J.M., Adams, M.R., Hoyer, P.B., Christian, P.J., Isom, S., Kaplan, J.R., Appt, S.E. (2009). Effects of 4-vinylcyclohexene diepoxide on peripubertal and adult Sprague-Dawley rats: Ovarian, clinical, and pathologic outcomes. Comparative Medicine, 59(1): 46-59.

Muskhelishvili, L., Wingard, S.K. and Latendresse, J.R. (2005) Proliferating cell nuclear antigen A marker for ovarian follicle counts. Toxicol. Pathol. 33:365-368.

Muskhelishvili, L., Freeman, L.D., Latendresse, J.R. and Bucci, T.J. (2002) An immunohistochemical label to facilitate counting of ovarian follicles. Toxicol Pathol 30:400-402.

Myllynen P, Immonen E, Kummu M and Vähäkangas K (2009): Developmental expression of drug metabolizing enzymes and transporter proteins in human placenta and fetal tissues. Expert Opin. Drug Metab. Toxicol. 5, 1483-1499.

Nakamura N, Kobashi Y, Kim K.S., Park H, Tani Y, Schimazu Y, Zhao T, Nishikawa Y, Omata F, Kawashima M, Yoshida M, Abe T, Saito Y, Senoo Y, Nonaka S, Takina M, Yamamoto C, Kawamura T,

Sugiyama A, Nakayama A, Kaneko Y, Jeong Y.D., Tatematsu D, Akao M, Sato Y, Iwanami S, Fujita Y, Wakui M, Aihara K, Kodama T, Shibuya K, Iwami S, Tsubokura M (2024). Modeling and predicting individual variation in COVID-19 vaccine-elicited antibody response in the general population, PLOS Digital health, https://doi.org/10.1371/journal.pdig.0000497.

Odum, J., Tinwell, H., Jones, K., Van Miller, J. P., Joiner, R. L., Tobin, G., Kawasaki, H., Deghenghi, R., and Ashby, J. (2001). Effect of rodent diets on the sexual development of the rat. Toxicol Sci 61(1), 115-127.

OECD (2002a), *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies*, OECD Series on Testing and Assessment, No. 32, OECD Publishing, Paris, https://doi.org/10.1787/9789264078482-en.

OECD (2002b), OECD (2002), *Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies*, OECD Series on Testing and Assessment, No. 35, OECD Publishing, Paris, https://doi.org/10.1787/9789264078499-en. OECD (2004), *Guidance Document for Neurotoxicity Testing*, OECD Series on Testing and Assessment, No. 20, OECD Publishing, Paris, https://doi.org/10.1787/16d23a96-en.

OECD (2007), *Test No. 426: Developmental Neurotoxicity Study, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264067394-en.*

OECD (2009), *Guidance Document for Histologic Evaluation of Endocrine and Reproductive Tests in Rodents*, OECD Series on Testing and Assessment, No. 106, OECD Publishing, Paris, https://doi.org/10.1787/cd68c07d-en.

OECD (2011), *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*, OECD Series on Testing and Assessment, No. 117, OECD Publishing, Paris, https://doi.org/10.1787/1648b49f-en.

OECD (2012), *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 Series on Testing and Assessment, No. 178, OECD Publishing, Paris, https://doi.org/10.1787/9789264221352-en.

OECD (2014), *Guidance Document 116 on the Conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451, 452 and 453: Second edition*, OECD Series on Testing and Assessment, No. 116, OECD Publishing, Paris, https://doi.org/10.1787/9789264221475-en .

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

OECD (2024), *(Q)SAR Assessment Framework: Guidance for the regulatory assessment of (Quantitative) Structure Activity Relationship models and predictions, Second Edition*, OECD Series on Testing and Assessment, No. 405, OECD Publishing, Paris, https://doi.org/10.1787/bbdac345-en.

OECD (2025a), *Test No. 443: Extended One-Generation Reproductive Toxicity Study*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264185371-en .

OECD (2025b), *Guidance Document on the Generation, Reporting and Use of Research Data for*

Regulatory Assessments, OECD Series on Testing and Assessment, No. 417, OECD Publishing, Paris, https://doi.org/10.1787/8d49ec1d-en.

OECD (2026), *Guidance Document on Mammalian Reproductive Toxicity Testing and Assessment*, *Second edition*, OECD Series on Testing and Assessment No. 446, Organisation for Economic Cooperation and Development, Paris. https://one.oecd.org/document/ENV/CBC/MONO(2026)26/en/pdf

Owens, W., Ashby, J., Odum, J., and Onyon, L. (2003). The OECD program to validate the rat uterotrophic bioassay. Phase 2: dietary phytoestrogen analyses. Environ Health Perspect. 111(12), 1559-1567.

Padilla-Banks E, Wendy N. Jefferson, and Retha R. Newbold. 2006. Neonatal Exposure to the Phytoestrogen Genistein Alters Mammary Gland Growth and Developmental Programming of Hormone Receptor Levels Endocrinology. 147: 4871-4882, doi:10.1210/en.2006-038

Palmer, A. K., Ulbrich, B. C (1997), The cult of culling, Fundamental and Applied Toxicology 38 (1), 7-22.

Pedersen T, Peters H (1968). Proposal for a Classification of Oocytes and Follicles in the Mouse Ovary. J Reprod Fert. 17(3):555-557. doi: 10.1530/jrf.0.0170555

Picut, C.A., Swanson, C.L., Scully, K.L., Roseman, V.C., Parker, R.F. and Remick, A.K. (2008) Ovarian follicle counts using proliferating cell nuclear antigen (PCNA) and semi-automated image analysis in rats. Toxicol. Pathol. 36:674-679.
Plagemann A, Harder T, Brunn M, Harder A, Roepke K, Wittrock-Staar M, Ziska T, Schellong K, Rodekamp E, Melchior K, Dudenhausen JW. 2009. Hypothalamic proopiomelanocortin promoter methylation becomes altered by early overfeeding: an epigenetic model of obesity and the metabolic syndrome. Journal of Physiology 587:4963-4976.

Putz, O., Schwartz, C. B., Leblanc, G. A., Cooper, R. L., and Prins, G. S. (2001). Neonatal low- and high-dose exposure to estradiol benzoate in the male rat: II. Effects on male puberty and the reproductive tract. Biol. Reprod. 65(5), 1506-1517.

Rangaraj, N, Gupta, PD (1997). Responsiveness of vaginal cells to estradiol during postnatal development of rat. Recept. Signal Transduct. 7: 291-298.

Richter-Reichhelm, H.-B., Dasenbrock, C., Descotes, G., Emmendörfer, A., Heinrich, U. E., Harlemann, J. H., Hilde-brand, B, Küttler, K., Rühl-Fehlert, C. I., Schilling, K., Schulte, A. E. and Vohr, H.-W. 1995. Validation of a modified 28-Day rat study to evidence effects of test compounds on the immune system. Regulatory Toxicol. Pharmakol., 22, 54-56.

RITA (2003). http://reni.item.fraunhofer.de/reni/trimming/index.php (see also Ruehl-Fehlert et al 2003)

Rodriguez, P, Fernandez-Galaz, C, Tejero, A (1993). Controlled neonatal exposure to estrogens: A suitable tool for reproductive aging studies in the female rat. Biol. Reprod. 49: 387-392.

Roman D, Ulrich P, Paul G, Court M, Vit P, Kehren J, Mahl A. 2004. Determination of the effect of calcineurin inhibitors on the rat's immune system after KLH immunisation. Toxicol Lett. 149:133-40.

Ronti T, Lupattelli G, Mannarino E. 2006. The endocrine function of adipose tissue: an update. Clinical Endocrinology 64:355-365.

Russell, L.D. et al., (1990). Histological and Histopathological Evaluation of the Testis, Cache River Press, Clearwater, Florida

Russo J, Russo IH. (1996). Experimentally induced mammary tumours in rats. Breast Cancer Res Treat. 39(1):7-20.

Saghir Shakil A., Michael J. Bartels, David L. Rick, Alene T. McCoy, Reza J. Rasoulpour, Robert G. Ellis-Hutchings, M. Sue Marty, Claire Terry, Jason P. Bailey, Richard Billington, James S. Bus (2012), Assessment of diurnal systemic dose of agrochemicals in regulatory toxicity testing – An integrated approach without additional animal use, Regul Toxicol Pharmacol., 63(2):321-32
Sanbuissho et al., (2009), "Collaborative work on evaluation of ovarian toxicity by repeated-dose and fertility studies in female rats. The Journal of Toxicological Sciences 34: Special Issue 1, SP1-SP22.

Sakai, T., M. Takahashi, K. Mitsumori, K. Yasuhara, K. Kawashima, H. Mayahara and Y. Ohno (2000). "Collaborative work to evaluate toxicity on male reproductive organs by repeated dose studies in rats - overview of the studies", Journal of Toxicological Sciences, 25, 1-21.

Sato J, Hashimoto S, Doi T, Yamada N, Tsuchitani M. Histological characteristics of the regression of corpora lutea in wistar hannover rats: the comparisons with sprague-dawley rats. Journal of toxicologic pathology. 2014;27(2):107-13.

Scharer, K. (1977). The effect of chronic underfeeding on organ weights of rats. How to interpret organ weight changes in cases of marked growth retardation in toxicity tests? Toxicology 7(1), 45-56.

Schneider, S., Kaufmann, W., Strauss, V., and van, R. B. (2011). Vinclozolin: a feasibility and sensitivity study of the ILSI-HESI F1-extended one-generation rat reproduction protocol. Regul. Toxicol Pharmacol. 59(1), 91-100.

Schwartz, C.L., Christiansen, S., Vinggaard, A.M. *et al.* Anogenital distance as a toxicological or clinical marker for fetal androgen action and risk for reproductive disorders. *Arch Toxicol* **93**, 253–272 (2019). https://doi.org/10.1007/s00204-018-2350-5

Schwartz CL, Christiansen S, Hass U, Ramhøj L, Axelstad M, Löbl NM, Svingen T. On the Use and Interpretation of Areola/Nipple Retention as a Biomarker for Anti-androgenic Effects in Rat Toxicity Studies. Front Toxicol. 2021 Oct 27;3:730752. doi: 10.3389/ftox.2021.730752

Seki, M., Yamaguchi, K., Marumo, H., and Imai, K. (1997). Effects of food restriction on reproductive and toxicological parameters in rats-in search of suitable feeding regimen in long-term tests. J Toxicol Sci 22(5), 427-437.

Siemens Medical Solutions Diagnostics (Diagnostic Products Corporation, DPC), Coat-A-Count TKT3 (total T3) and TKT4 (total T4), Veterinary application documents of T3 (March 26, 1993, ZV106 A) and T4 (March 24, 1993, ZV103 A).

Smith, H. W., Winstead, C. J., Stank, K. K., Halstead, B. W., and Wierda, D. (2003). A predictive F344 rat immunotoxicology model: cellular parameters combined with humoral response to NP-CgammaG and KLH. Toxicology 194(1-2), 129-145.

Spencer, S. J., Galic, M. A., and Pittman, Q. J. (2011). Neonatal programming of innate immune function. Am. J Physiol Endocrinol. Metab 300(1), E11-E18.

Steinmetz R, Mitchner NA, Grant A, Allen DL, Bigsby RM, Ben-Jonathan N. The xenoestrogen bisphenol A induces growth, differentiation, and c-fos gene expression in the female reproductive tract. Endocrinology 1998; 139:2741–2747.

Takayama S, Akaike M, Kawashima K, Takahashi M, Kurokawa Y; A Collaborative Study in Japan on Optimal Treatment Period and Parameters for Detection of Male Fertility Disorders in Rats Induced by

Medical Drugs. J. Amer. Coll. Toxicol., 14: 266-292, 1995.

Taketa Y., 2021. Luteal toxicity evaluation in rats. J Toxicol Pathol. 2021 Nov 18;35(1):7–17. doi: 10.1293/tox.2021-0058

Terry, K. K., Chatman, L. A., Foley, G. L., Kadyszewski, E., Fleeman, T. L., Hurtt, M. E., and Chapin, R. E. (2005). Effects of feed restriction on fertility in female rats. Birth Defects Res B Dev. Reprod. Toxicol 74(5), 431-441.

The ICICIS Group Investigators. (1998). Report of validation study of assessment of direct immunotoxicity in the rat. Toxicology 125, 183–201.

Thigpen, J. E., Setchell, K. D., Padilla-Banks, E., Haseman, J. K., Saunders, H. E., Caviness, G. F., Kissling, G. E., Grant, M. G., and Forsythe, D. B. (2007). Variations in phytoestrogen content between different mill dates of the same diet produces significant differences in the time of vaginal opening in CD-1 mice and F344 rats but not in CD Sprague-Dawley rats. Environ Health Perspect. 115(12), 1717-1726.

Thigpen, J. E., Setchell, K. D., Saunders, H. E., Haseman, J. K., Grant, M. G., and Forsythe, D. B. (2004). Selecting the appropriate rodent diet for endocrine disruptor research and testing studies. ILAR. J 45(4), 401-416.

Tonk, E. C., de Groot, D. M., Penninks, A. H., Waalkens-Berendsen, I. D., Wolterbeek, A. P., Piersma, A. H., and van Loveren, H. (2011). Developmental immunotoxicity of di-n-octyltin dichloride (DOTC) in an extended one-generation reproductive toxicity study. Toxicol Lett. 204(2-3), 156-163.

Tyl, R.W., K. Crofton, A. Moretto, V. Moser, L.P. Sheets and T.J. Sobotka (2008), “Identification and Interpretation of Developmental Neurotoxicity Effects: a Report from the ILSI Research Foundation/Risk Science Institute Expert Working Group on Neurodevelopmental Endpoints”, Neurotoxicology and Teratology, 30: 349-381.

Ulbrich, B. and A.K. Palmer (1995), “Detection of Effects on Male Reproduction - a Literature Survey”, Journal of the American College of Toxicologists, 14, 293-327.

Ulrich P, Paul G, Perentes E, Mahl A, Roman D. 2004. Validation of immune function testing during a 4-week oral toxicity study with FK506. Toxicol Lett. 149:123-31.

US EPA (2012). Benchmark Dose Technical Guidance Document https://www.epa.gov/risk/benchmark-dose-technical-guidance

US EPA (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

US EPA. (2006). A framework for assessing health risk of environmental exposure to children. Office of Research and Development, National Center for Environmental Assessment, Washington, DC; EPA/600/R-05/093F.

US EPA (2007). 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

Vandenberg L.N., Maffini M.V., Schaeberle C.M., Ucci A.A., Sonnenschein C., Rubin B.S., Soto A.M. (2008). Perinatal exposure to the xenoestrogen bisphenol-A induces mammary intraductal hyperplasias

in adult CD-1 mice. Reproductive Toxicology, 26 (3-4), pp. 210-219

Vandebriel RJ, Hakkert B-, Pennings J.L.A., Rossi LH https://pubmed.ncbi.nlm.nih.gov/39041360/ - full-view-affiliation-3, Bichlmaier I -, Bil W (2024). Evaluating the conduct and reporting of the T-cell dependent antibody response in the Extended One-Generation Reproductive Toxicity study provided under the EU REACH regulation. Crit Rev Toxicol. Aug;54(7):476-484.

Wang X, Coppi A, Lebrec H. 2018. T-cell-dependent antibody response assay: biology, methods, and application. In: McQueen CA, editor. Comprehensive toxicology. 3rd ed. Amsterdam: Elsevier. Doi: 10.1016/B978-0-12-801238-3.64235-6.484

White, K. L., Jr., Sheth, C. M., and Peachee, V. L. (2007). Comparison of primary immune responses to SRBC and KLH in rodents. J Immunotoxicol. 4(2), 153-158.

Willoughby, CR et al. (2000). Two-generation reproduction studies of di-(C7-C9 alkyl) phthalate and di-(C9-C11 alkyl) phthalate in the rat. Reproductive toxicology 14:427-450.

Wiltse, J. (2005). Mode of action: Inhibition of histone deacetylase, altering WNT-dependent gene expression, and regulation of betacatenin—Developmental effects of valproic acid. Crit. Rev. Toxicol. 35:727–738.

Wise D (2024). Rodent anogenital distance recommendations. Birth Defects Research 2024; 116:e2347. Doi: 10.1002/bdr2.2347

Yang, J-M; Salmon, A.G., and Marty, M.A. 2009. Development of TEFs for PCB congeners by using an alternative biomarker —Thyroid hormone levels. Regul Toxicol Pharmacol. Dec 31.

Yki-Järvinen H. 2010. Liver fat in the pathogenesis of insulin resistance and type 2 diabetes. Digestive Diseases 28:203-209.

Yoon M, and Barton HA (2008): Predicting maternal rat and pup exposures: How different are they? Toxicol Sci., 102, 15-32.

Yoshida, M., Sanbuissyo, A., Hisada, S., Takahashi, M., Ohno, Y., and Nishikawa, A. (2009) Morphological characterization of the ovary under normal cycling in rats and its viewpoints of ovarian toxicity detection. J. Toxicol. Sci. 34 Suppl 1: SP 189–197.You, L., Sar, M., Bartolucci, E.J., McIntyre, B.S. & Sriperumbudur, R. Modulation of mammary gland development in prepubertal male rats exposed to genistein and methoxychlor. Toxicol. Sci. 66, 216–225 (2002).

Zimmermann P, Curtis N (2019). Factors That Influence the Immune Response to Vaccination. Clin Microbiol Rev. Mar 13;32(2):e00084-18.

Zoeller, R.T., and Crofton, K.M. (2005). Mode of action: Developmental thyroid hormone insufficiency—neurological abnormalities resulting from exposure to propylthiouracil. Crit. Rev. Toxicol. 35:771–781.

Zoeller RT, Bansal R., Parris C. 2005. Bisphenol-A, an environmental contaminant that acts as a thyroid hormone receptor antagonist *in vitro*, increases serum thyroxin, and alters RC3/neurogranin expression in the developing rat brain. Endocrinology 146:607-612.

Zoeller RT, Rovet J. 2004. Timing of thyroid hormone action in the developing brain: clinical observations and experimental findings, J. Neuroendocrinology. **16:809-818**.

Zoeller RT, Tyl RW, Tan SW. 2007. Current and potential rodent screens and tests for thyroid toxicants. Critical Reviews in Toxicology 37:55-95.

Zorrilla LM, Gibson EK, Jeffay SC, Crofton KM, Setzer WR, Cooper RL and Stoker TE. 2009. The effects of triclosan on puberty and thyroid hormones in male Wistar rats. Toxicological Sciences 107:56-64.

Annex 1. Detailed list of endpoints and examinations included in the EOGRTS.

Ax 1.1 This Annex contains tables detailing the endpoints and examinations required in TG 443. The same information is provided in TG 443 but these tables are intended to provide it in a way that makes study design and protocol drafting easier. Annex Table 1.1 lists the endpoints required in the parental generation (P). Annex Table 1.2 lists the endpoints required for litters, prior to pup selection at weaning. Annex Table 1.3 lists the endpoints required in the offspring generation(s) (F1 and F2 if triggered), once pup selection to cohorts has taken place. Endpoints are listed in Annex Table 1.3 according to cohort but it should be noted that although cohorts 1, 2 and 3 are nominally designated reproductive, neurotoxicity and immunotoxicity endpoints respectively, in practice, some endpoints are covered by all cohorts and some are in different cohorts. Optional endpoints are not listed here but are given in TG 443.

Table A 1.1. Endpoints and examinations required in the parental generation(s) (P) (excluding litter observations).

Endpoints/examinations	**P males (all, unless otherwise specified)**	**P females (all, unless otherwise specified)**
	Frequency of observations	**Frequency of observations**
PHASE OF STUDY: IN-LIFE		
General observations: - Behavioural changes - All signs of toxicity - Morbidity - Mortality	Once a day	Once a day
Body weight	On 1st day of dosing and at least weekly thereafter.	On 1st day of dosing and at least weekly thereafter. During lactation, on the same days as the pups.
Clinical observations including: - Changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (*e.g.*, lacrimation, piloerection, pupil size, unusual respiratory pattern). - Changes in gait, posture, response to handling, presence of clonic or tonic movements, stereotypy (*e.g.* excessive grooming, repetitive circling) or bizarre behaviour (*e.g.* self-mutilation, walking backwards).	Once a week (*e.g.* when animals are weighed).	
Food consumption (or water consumption, if substance administered in the drinking water).	At least weekly (same day as weighing)	
Oestrous cyclicity (by vaginal cytology)	NA	At the beginning of treatment period until confirmation of mating

		or end of mating period.
Mating and pregnancy parameters including: - Precoital interval and duration of pregnancy. - Signs of dystocia, abnormal nesting behaviour, nursing performance.	NA	As often as is applicable.

PHASE OF STUDY: TERMINAL

Body weight.	At termination	
Macroscopic examination of all organs for abnormalities	At termination	
Organ weight: - Uterus (with oviducts and cervix) - Ovaries - Testes - Epididymides (total and cauda for the samples used for sperm counts) - Prostate (dorsolateral and ventral part combined) - Seminal vesicles with coagulating glands and their fluids (as one unit). - Brain - Liver - Kidneys - Heart - Spleen - Thymus - Adrenal glands - Pituitary - Thyroid - Other known target organs	 At termination At termination At termination At termination At termination At termination At termination At termination At termination At termination At termination At termination Post-fixation At termination	 At termination At termination
Histopathology of fixed organs: - Uterus (with oviducts and cervix) - Ovaries - Testes (detailed histopathological examination of one testis) - Epididymides (detailed histopathological examination of one epididymis) - Prostate (dorsolateral and ventral) - Seminal vesicles (and coagulating glands) - Brain - Liver - Kidneys - Heart - Spleen - Thymus - Adrenal glands - Pituitary - Thyroid (and parathyroid) - Peripheral nerve - Muscle - Spinal cord - Eye (and optic nerve)	Post termination, both sexes (as applicable). In all high dose and controls. Additionally, in lower doses if treatment related-change is shown. Include reproductive organs of all animals suspected of reduced fertility.	

- Gastrointestinal tract - Urinary bladder - Lung - Trachea - Bone marrow - Vas deferens (males) - Mammary gland (males and females) - Vagina - Other known target organs		
Examination of the uteri for presence and number of implantation sites.	NA	At termination
Oestrous cycle stage (by vaginal cytology)	NA	At termination
Sperm parameters: - Enumeration of cauda epididymis sperm reserves. - Evaluation of sperm motility and morphology. - Evaluation of spermatid count	At or post termination.	NA
Clinical biochemistry (including): - Glucose - Total cholesterol - Urea - Creatinine - Total protein - Albumin - Two enzymes indicative of hepatocellular effects	10 animals/sex/group at termination (post-fasting).	
Haematology : - Haematocrit, - Haemoglobin concentration - Erythrocyte count - Total and differential leukocyte count - Platelet count - Blood clotting time/potential	10 animals/sex/group at termination (post-fasting).	
Thyroid hormones (T4 and TSH)	10 animals/sex/group at termination (post fasting) or at a pre-termination bleed.	

NA : not applicable

Table A 1.2. Endpoints and examinations required in F1 litters (and F2 if triggered) up to weaning.

Endpoints/examinations	F1 (and F2) males (all, unless otherwise specified)	F1 (and F2) females (all, unless otherwise specified)
	Frequency of observations	**Frequency of observations**
PHASE OF STUDY: IN-LIFE		
General observations: - Behavioural changes - All signs of toxicity - Morbidity - Mortality	Once a day	
Body weight	On PND 0 or PND 1 and regularly thereafter (at least on PND 4, 7, 14 & 21)	
Litter examination including: - Number and sex of pups, stillbirths and live births. - Presence of gross anomalies (externally visible abnormalities, including cleft palate; subcutaneous haemorrhages; abnormal skin colour or texture; presence of umbilical cord; lack of milk in stomach; presence of dried secretions).	As soon as possible after birth. Live pups are to be counted on PND 4, 7, 14 and 21	
Clinical examination of the neonates, e.g. - Qualitative assessment of body temperature, state of activity and reaction to handling. - External abnormalities, changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity. - Changes in gait, posture, response to handling. -Clonic or tonic movements, stereotypy or bizarre behaviour. - Abnormalities of genital organs. *e.g.* hypospadias or cleft penis	As often as is applicable and when weighed.	
Anogenital distance in pups and bodyweight (both on the same day)	Measured together with bodyweight in all pups (i.e., all dose groups, including concurrent controls) at least once, on the same day of age, between PND 0 and PND 4	
Counting the number of retained nipples/areolae in male pups (see GD 151, Section 3).	The presence of areolae in male pups should be assessed and recorded between PND 12 and 14, on the same PND for all pups (i.e., all dose groups, including concurrent controls) when the nipples are clearly visible in female littermates	Report qualitatively that the nipples are visible in the female littermates
PHASE OF STUDY: TERMINAL		
Surplus pups after standardisation at PND 4		
Body weight, gross necropsy, thyroid hormones (T4 and TSH).	At termination	

Surplus pups not allocated to Cohorts (at weaning) and F2 generation (if triggered)	
Body weight.	At termination
Macroscopic examination of all organs for abnormalities	At termination
Organ weight and/or retention for possible histopathology : - Brain - Spleen - Thymus - Mammary tissues - Other organs as appropriate	10 pups/sex/group at termination
Thyroid hormones (T4 and TSH)	10 animals/sex/group at termination

NA : not applicable

Table A 1.3. Endpoints and examinations required in the offspring generation(s) post weaning, when selection of pups to cohorts has taken place

Endpoints/examinations	Cohort 1	Cohort 2	Cohort 3
PHASE OF STUDY: IN-LIFE. ENDPOINTS NOT SPECIFIC TO COHORTS			
	Frequency of observations		
General observations: - Behavioural changes - All signs of toxicity - Morbidity - Mortality	Once a day		
Body weight	All cohorts. At least weekly and on day of attainment of vaginal opening or balano-preputial separation.		
Clinical observations including: - Changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (*e.g.*, lacrimation, piloerection, pupil size, unusual respiratory pattern). - Changes in gait, posture, response to handling, presence of clonic or tonic movements, stereotypy (*e.g.* excessive grooming, repetitive circling) or bizarre behaviour (*e.g.* self-mutilation, walking backwards). - Abnormalities of genital organs *e.g.* hypospadias or cleft penis	All cohorts. When animals are weighed		
Food consumption (or water consumption, if substance administered in the drinking water)	All cohorts. At least weekly.		
Sexual maturity: - vaginal opening (females) - balano-preputial separation (males)	All cohorts, except cohort 2. Daily examination until achieved. Body weight should be recorded at the time of sexual maturity.		
PHASE OF STUDY: IN-LIFE. ENDPOINTS SPECIFIC TO COHORTS 1, 2 OR 3			
	Frequency of observations	**Frequency of observations**	**Frequency of observations**

Oestrous cyclicity (by vaginal cytology)	Cohort 1A. Daily from onset of vaginal opening until 1st oestrus - Daily, for 2 weeks from around PND 75 Cohort 1B. If mated: From pairing until confirmation of mating.	Not determined	Not determined
Mating and pregnancy parameters including: - Precoital interval and duration of pregnancy. - Signs of dystocia, abnormal nesting behaviour, nursing performance.	Cohort 1B. If mated: As often as is applicable	Not determined	Not determined
Litter parameters including: - Number and sex of pups, stillbirths and live births - Presence of gross anomalies (externally visible abnormalities, including cleft palate; subcutaneous haemorrhages; abnormal skin colour or texture; presence of umbilical cord; lack of milk in stomach; presence of dried secretions).	Cohort 1B. F2 pups: As often as is applicable	Not determined	Not determined
Assessment of neurotoxicity: - Auditory startle test - Functional observation battery - Motor activity (determined at least once)	Not determined	Cohort 2A. PND 24 or 25. Between PND 63 and 75.	Not determined

PHASE OF STUDY: TERMINAL. ENDPOINTS NOT SPECIFIC TO COHORTS

Body weight	All cohorts. At termination.
Oestrous cycle stage (by vaginal cytology)	All cohorts. At termination.
Examination of external organs (especially sex organs) for structural abnormalities	All cohorts. At termination.
Macroscopic examination of all internal organs for abnormalities	All cohorts. At termination.

PHASE OF STUDY: TERMINAL. ENDPOINTS SPECIFIC TO COHORTS 1, 2 OR 3

Organ weight: - Uterus (with oviducts and cervix) - Ovaries - Testes - Epididymides (total and cauda for the samples used for sperm counts cohort 1A) - Prostate (dorsolateral and ventral part combined) - Seminal vesicles with coagulating glands and their fluids (as one unit) - Brain - Liver - Kidneys	 Cohort 1A &1B Cohort 1A &1B Cohort 1A &1B Cohort 1A &1B Cohort 1A &1B Cohort 1A &1B Cohort 1A Cohort 1A Cohort 1A	 Cohort 2A & 2B. Brain weight only determined.	Not determined

- Heart - Spleen - Thymus - Lymph nodes - Adrenal glands - Pituitary - Thyroid (post-fixation) - Other known target organs	Cohort 1A Cohort 1A Cohort 1A Cohort 1A Cohort 1A Cohort 1A &1B Cohort 1A Cohort 1A &1B Cohort 1A &1B		
Histopathology of fixed organs: - Uterus (with oviducts and cervix) - Ovaries (including enumeration of primordial and small growing follicles, corpora lutea) - Testes (detailed histopathological examination of one testis cohort 1A) - Epididymides (detailed histopathological examination of one epididymis cohort 1A) - Prostate (dorsolateral and ventral) - Seminal vesicles (and coagulating glands) - Brain - Liver - Kidneys - Heart - Spleen - Thymus - Adrenal glands - Pituitary - Thyroid (and parathyroid) - Peripheral nerve - Muscle - Spinal cord - Eye (and optic nerve) - Gastrointestinal tract - Urinary bladder - Lung - Trachea - Bone marrow - Vas deferens (males) - Vagina - Mammary gland (males and females) - Known target organs	Cohort 1A. Histopathology of all tissues for high dose and control. In lower dose groups if treatment related findings. Cohort 1B. Histopathology in case of suspected repro or endocrine toxicant and/or if results from cohort 1A are equivocal. Organs preserved are vagina, uterus, cervix, ovaries, testes, epididymides, seminal vesicles, coagulating glands, prostate, pituitary, other identified target organs.	Neurohistopathology assessed in Cohorts 2A and 2B (see below).	Not determined
Assessment of neurohistopathology : (Using qualitative and quantitative methods) - Olfactory bulbs - Cerebral cortex - Hippocampus - Basal ganglia - Thalamus - Hypothalamus - Mid-brain (thecum, tegmentum, cerebral peduncles) - Brain-stem - Cerebellum - Eyes, (retina and optic nerve)	Not determined	Cohort 2A. between PND 75 and 90. Cohort 2B: on PND21 or 22. All endpoints required for cohort 2A. Eyes, peripheral nerve, muscle and spinal cord not required for cohort 2B.	Not determined

- Peripheral nerve - Muscle - Spinal cord		Note that assessment is done for control and high dose animals and in lower doses if treatment related changes are shown.	
Examination of the uteri for presence and number of implantation sites (at termination)	Cohort 1B. If mated.	Not determined	Not determined
Sperm parameters: - Enumeration of cauda epididymis sperm reserves. - Evaluation of sperm motility and morphology. - Evaluation of spermatid count.	Cohort 1A: At or post termination.	Not determined	Not determined
Clinical biochemistry (including): - Glucose - Total cholesterol - Urea - Creatinine, - Total protein, - Albumin - Two enzymes indicative of hepatocellular effects	Cohort 1A: 10 animals/sex/group at termination.	Not determined	Not determined
Haematology : - Haematocrit, - Haemoglobin concentration, - Erythrocyte count, - Total and differential leukocyte count	Cohort 1A. 10 animals/sex/group at termination	Not determined	Not determined
Thyroid hormones (T4 and TSH)	Cohort 1A. 10 animals/sex/group at termination or at a pre-termination bleed.	Not determined	Not determined
Assessment of immunotoxicity: - primary and secondary IgM and IgG antibody response to a T cell dependent antigen (immunization with antigen is part of the test)	Not determined	Not determined	Cohort 3. Please see paragraph 52 and Figure 2 of TG 443 for details of the immunization and sample collection scheme.
Assessment of immunotoxicity: - Splenic lymphocyte subpopulation analysis (CD4+ and CD8+ T lymphocytes, B lymphocytes and NK cells) using one half of the spleen. - Weight of lymph nodes associated with and distant from the route of exposure. - Histopathology on the collected lymph nodes and bone marrow.	Cohort 1A. 10 animals/sex/ group at termination.	Not determined	Not determined

Annex 2. Analysis of the effect of decreased body weight (via dietary modulation) on absolute organ weight in rats

Objective

Ax2.1. The objective of this analysis is to aid the interpretation of rat organ weight data from the EOGRTS when body weight (BW) is reduced, by providing numerical data from literature studies.

Methods

Ax2.2. A literature search was conducted to identify papers that studied the effect of dietary modulation on organ weight in rats. Data were then extracted and analysed from suitable papers. Papers considered to be suitable were those where feed restriction or where different proprietary diets that had been modified with the intention of producing different body weights had been used. Papers using diets with nutrient deficiencies or supplements were excluded.

Although the protocols for dietary modulation varied, the common parameters of terminal body weights and organ weights were compared across studies. As the EOGRTS involves dosing for at least one month, only studies of approximately ≥ one month were selected in order to make the data more relevant. The studies are a mixture of reproductive and non-reproductive with varying methods of feed restriction/dietary modulation. In this analysis, however, these variables are acknowledged but not accounted for.

Ax2.3. Organ weight and terminal body weight data for adult rats were selected from the publications. The protocols used in the studies are summarised below. The percentage decreases in terminal body weight compared to the relevant control group were then calculated. The treatment groups across all the studies were then assigned to categories of 5%, 10%, 15%, 20%, 30% and 40% decrease in terminal body weight compared to control values. The percentage decreases in organ weights compared to the relevant control group were then calculated as

$$\frac{\text{(control organ weight – treated organ weight) x 100}}{\text{control organ weight}}$$

and the data tabulated according to the body weight decrease category. Organ weight data from categories of terminal body weight reduction less than 5% or greater than 40% of control values were not used in this analysis (using rounded figures). The boundaries were set based on the questionable relevance of effects below 5% and that substance-induced effects in the EOGRTS above 40% would be highly unlikely and undesirable if dose setting has been done properly.

Descriptions of protocols

Ax2.4. The papers selected for this analysis are listed below. The relevant protocol details are summarised but it should be noted that all the papers used in this analysis contain many other experiments.

Scharer, 1977.

> Wistar rats (males only) were feed restricted (FR) for up to 13 weeks. Rats had an initial BW of 110 g (age not given) at onset of treatment. Food allowance was varied to achieve a target BW of 250g after 13 weeks treatment. Weeks at which FR took place varied by group. Groups used in Tables 1-4 are: Ad libitum (ad lib) (control group); FR from weeks 0-13 to achieve BW of 250g; FR from weeks 5-13 to achieve BW of 250g; FR from weeks 10-13 to achieve BW of 250g. N=12 rats per sex per group.

Chapin et al, 1993.

> SD rats (males and females) were maintained at 90, 80 & 70% of control BW by FR. Rats were approximately 10 weeks old at onset of treatment. Food allowance was based on a pilot study initially, then varied to achieve the target weights. Males were killed after 15 weeks treatment (at 25 weeks old), females were mated and killed at GD 14 (age 27 weeks old), total treatment time for females was 17 weeks. Groups used in Tables 1-4 are: Ad lib (control group); FR to achieve 90% of control BW; FR to achieve 80% of control BW; FR to achieve 70% of control BW; N=20 rats per sex per group.

Keenan et al, 1994.

> SD rats (males and females) were maintained on different proprietary diets (PMI 5002 and PMI 5002-9) with ad lib and feed restricted groups for each diet. PMI 5002-9 had a lower metabolisable energy value than PMI 5002. PMI 5002 was considered to be the “control” group for comparative purposes and statistical evaluation. Rats were 36 days old, males weighed 115-175g and females 91-156g, at onset of treatment. Animals were killed after 52 weeks treatment. Groups used in Tables 1-4 are: Ad lib PMI 5002 (control group); PMI 5002 FR by feeding for 6.5h per day; PMI 5002 FR by feeding 65% of control food consumption; ad lib PMI 5002-9; PMI 5002-9 FR to achieve the same caloric intake as the PMI 5002 FR 65% group. N=10 rats per sex per group.

Seki et al, 1997.

> SD rats (males and females) were feed restricted for 13 weeks. Rats were 6-weeks old at onset of treatment. Rats were feed restricted such that the amount of feed they received was 85%, 70% and 55 % of the food consumption (FC) of the ad lib control group. Animals were killed after 13 weeks treatment. Groups used in Tables 1-4 are: Ad lib (control group); FR to 85% of control food consumption; FR to 70% of control food consumption; FR to 55% of control food consumption. N=10 rats per sex per group.

Odum et al (2001)

> Pregnant Wistar rats were maintained on different proprietary diets (PMI 5001 and RM1/RM3) from day 1 of pregnancy. RM1/RM3 had a lower metabolisable energy value than PMI 5001. There was no feed restriction, but the diets resulted in body weight differences. PMI 5001 was designated to be the “control” group for comparative purposes and statistical evaluation. After weaning, F1 pups were then maintained on the diets. Males only were killed on PND 68. Groups used in Tables 1-4 are: Ad lib PMI 5001 (control group); ad lib RM1/RM3. N=5 or 6 litters (19 pups) per sex per group.

Marty et al., 2003.

> CD rats (males only) were feed restricted from PND 23. Food was restricted to achieve and maintain a 10% reduction in body weight compared to controls. Rats were killed at PND 45 or 59. Treatment duration was 22, or 36 days. Groups used in Tables 1-4 are: Ad lib from PND 22-45

(control group); FR from PND 22-45 to achieve a 10% reduction in body weight (90% control BW); Ad lib from PND 22-59 (control group); FR from PND 22-59 to achieve a 10% reduction in body weight (90% control BW). N=12 rats per group.

Carney et al., 2004.

Pregnant SD rats were feed restricted from GD 7 through lactation. Rats were feed restricted such that FC was reduced by 10, 30 and 50 % of that consumed by an ad lib control group. F1 male and female pups were killed at 21 weeks old. Groups used in Tables 1-4 are: Ad lib (control group); FR such that FC was 30% less than controls (70% of control FC); FR such that FC was 50% less than controls (50% of control FC). N=24-26 litters. Numeric terminal body weights are not given in the publication but the authors supplied additional data.

Laws et al., 2007.

Male Wistar rats were feed restricted from PND 23 to 53, female rats were feed restricted from PND 22 to 41. Rats were fed 10, 20, 30 and 40 % less than the amount consumed by control rats. Male rats were killed at PND 53 and female rats at PND 41. Groups used in Tables 1-4 are: Ad lib (control group); FR such that FC was 20% less than controls (80% of control FC); FR such that FC was 30% less than controls (70% of control FC); FR such that FC was 40% less than controls (60% of control FC). N=13 rats per group.

Results

Ax2.5. The different methods of feed restriction/ dietary modulation resulted in varying terminal body weights which provided a reasonable distribution of body weight reduction across the categories created for subsequent analysis of organ weights. Body weight decreases relative to controls in male and female rats are shown in Annex Tables 2.1 and 2.2 respectively. In the studies of Chapin et al (1993) and Marty et al (2003), where feed intake was adjusted to achieve target body weights, the degree of body weight decrease matched the target. Where animals were fed a fixed percentage of the control feed (Carney et al, 2004; Laws et al, 2007) then the percentage body weight decreases are more variable. The studies of Scharer (1977), Odum et al (2001) and Marty et al (2003) did not include adult female groups that were comparable to the other studies and therefore have been excluded from Annex Table 2.2. In male rats, all the studies produced body weight decreases that were within the rounded boundaries of 5-40%. In female rats three groups exceeded the upper boundary and therefore were not used in the subsequent analysis.

Table A 2.1. Terminal body weight changes and assignment to categories for male rats in the selected studies. Further details of treatment are provided in the text. Note that only results used in the analysis for Annex Table 2.3 are shown.

Age at termination; (duration of treatment)	Treatment	Body weight at term-ination	% Decrease in BW relative to controls	Category of BW decrease for sub-sequent analysis	Reference
Not specified, initial BW was 110g; (13w)	Ad lib (control) FR 0-13w to 250g BW FR 5-13w to 250g BW FR 10-13w to 250g BW	366 243[a] 251[a] 269[a]	- 33.6[a] 31.4[a] 26.5[a]	- 30% 30% 30%	Scharer
25 weeks old; (15 weeks)	Ad lib (control) FR to 90% control BW FR to 80% control BW FR to 70% control BW	670.8 592.8* 536.8* 468.5*	- 11.6* 20.0* 30.2*	- 10% 20% 30%	Chapin
57 weeks old; (52	Ad lib PMI 5002 (control)	734	-	-	Keenan

weeks)	FR PMI 5002 6.5h/d FR PMI 5002 to 65% control FC[b] Ad lib PMI 5002-9 FR PMI 5002-9 to caloric intake[b]	559* 570* 640* 431*	23.8* 35.0* 12.8* 41.3*	20% 30% 10% 40%	
19 weeks old; (13 weeks)	Ad lib (control) FR to 85% control FC FR to 70% control FC FR to 55% control FC	574.2 466.3 402.4 317.5	- 18.8 29.9 44.7	- 20% 30% 40%	Seki
68 days old; (F1, GD1-PND 68)	Ad lib PMI 5001 diet (control) Ad lib SDS RM3/RM1	404.5 362.6*	- 10.4*	- 10%	Odum
45 days old; (22 days) and 59 days old; (36 days)	Ad lib (control), to age 45d FR to 90% control BW, to age 45d Ad lib (control), to age 59d FR to 90% control BW, to age 59d	228.8 205.2* 340.1 306.6*	- 10.3* - 9.9*	- 10% - 10%	Marty
21 weeks old; (F1, GD7-PND147	Ad lib (control) FR to 70% control FC FR to 50% control FC	574.7 468.3* 331.8*	- 18.5* 42.3%*	- 15% 40%	Carney
53 days old; (31 days)	Ad lib (control) FR to 80% control FC FR to 70% control FC FR to 60% control FC	296.9 283.8 259.8* 235.4*	- 4.4 12.5* 20.7*	- 5% 15% 20%	Laws

FR: Feed restriction; FC: food consumption; BW: body weight
*Statistical significance was taken from the reference (publication)
[a]Statistical significance was not determined
[b]Caloric intake was determined for group "FR PMI 5002 to 65% control FC" and was matched in group "FR PMI 5002-9 to caloric intake[b]"

Table A 2.2. Terminal body weight changes and assignment to categories for female rats in the selected studies. Further details of treatment are provided in the text. Note that only results used in the analysis for Table 4 are shown (with the exception of groups annotated NA.

Age at termination; (duration of treatment)	Treatment	Body weight at term-ination	% Decrease in BW relative to controls	Category of BW decrease for sub-sequent analysis	Reference
27 weeks old; (17 weeks)	Ad lib (control) FR to 90% control BW FR to 80% control BW FR to 70% control BW	431.3 389.4* 347.6* 283.0*	- 9.7* 19.4* 34.4*	- 10% 20% 30%	Chapin
57 weeks old; (52 weeks)	Ad lib PMI 5002 (control) FR PMI 5002 6.5h/d FR PMI 5002 to 65% control FC[b] Ad lib PMI 5002-9 FR PMI 5002-9 to caloric intake[b]	456 309* 286* 380* 244*	- 32.2* 60.1* 16.7* 46.5*	- 30% NA[c] 15% NA[c]	Keenan
19 weeks old; (13 weeks)	Ad lib (control) FR to 85% control FC FR to 70% control FC FR to 55% control FC	321.6 253.8* 220.5* 171.5*	- 21.1* 31.4* 46.7*	- 20% 30% NA[c]	Seki
21 weeks old; (F1, GD7-	Ad lib (control) FR to 70% control FC	285.3 239.2*	- 16.2*	- 15%	Carney

PND147	FR to 50% control FC	180.6*	36.7%*	40%	
41 days old; (21 days)	Ad lib (control) FR to 80% control FC FR to 70% control FC FR to 60% control FC	154.8 147.4 136.0* 123.6*	- 4.8 12.1* 20.2*	- 5% 15% 20%	Laws

FR: Feed Restriction; FC: food consumption; BW: body weight
*Statistical significance was taken from the reference (publication)
[a]Statistical significance was not determined
[b]Caloric intake was determined for group "FR PMI 5002 to 65% control FC" and was matched in group "FR PMI 5002-9 to caloric intake[b]"
[c]NA: not applicable. Only body weight decreases up to approximately 40% of control values were analysed.

Ax2.6. The percentage decreases in absolute organ weights compared to control values, across the studies analysed are presented in Annex Tables 2.3 and 2.4. Each column of the tables represents a category of body weight decrease. Data was not available from every study for every organ.

Table A 2.3. Decrease in organ weights occurring in male rats with decreased body weights, as a consequence of dietary modulation. Body weights decreases across studies are grouped into % intervals as shown in Annex Table 2.1. Rows where study data for an organ weight were not reported are not shown.

Rat age at termin-ation of study	LIVER (Males): % decrease in organ weight						
	5% BW decrease[1]	10% BW decrease	15% BW decrease	20% BW decrease	30% BW decrease	40% BW decrease	Ref
Not specified, adult					48.1* 54.5* 52.4*		Scharer
25 w		15.6 *		30.4 *	43.7*		Chapin
19 w				28.6	41.3	55.6	Seki
68 d (F1)		11.6					Odum
45 d 59 d		22.7* 15.8*					Marty
21 w (F1)			32.4*			54.2*	Carney
53 d	14.5*		22.0*	32.8*			Laws

Rat age at termin-ation of study	KIDNEY WEIGHT (Males): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
Not specified, adult					32.6* 34.0* 23.5*		Scharer
25 w		5.5*		15.0 *	26.8*		Chapin
19 w				12.8*	29.1*	42.9	Seki
68 d (F1)		28.6					Odum
21 w (F1)			24.4*			46.6*	Carney
53 d	11.5*		15.6*	27.1*			Laws

Rat age at termin-ation	THYMUS WEIGHT (Males): % decrease in organ weight

of study	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
19 w				0.0	11.6	13.1	Seki
21 w (F1)			23.0*			30.6*	Carney

Rat age at termin-ation of study	SPLEEN WEIGHT (Males): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
Not specified, adult						32.0* 38.5* 38.6*	Scharer
19 w				12.7*	21.9*	38.9*	Seki
21 w (F1)			19.8*			38.3*	Carney

Rat age at termin-ation of study	ADRENALS WEIGHT (Males): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
Not specified, adult					11.4 -5.7 -2.1		Scharer
19 w				10.4*	9.2	23.9*	Seki
21 w (F1)			12.3*			10.5*	Carney
53 d	11.6		18.5*	26.8*			Laws

Rat age at termin-ation of study	BRAIN WEIGHT (Males): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
Not specified, adult					7.4* 6.5* 6.3*		Scharer
57 w		0.9		3.4	0.0	1.7	Keenan
19 w				1.1	3.1	4.4*	Seki
21 w (F1)			3.5*			12.4*	Carney

Rat age at termin-ation of study	THYROID WEIGHT (Males): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
Not specified, adult					21.7* 23.0* 18.9*		Scharer
57 w		2.9		14.3	28.6	31.4*	Keenan
19 w				19.8	27.8*	42.3*	Seki

Rat age at termin-ation of study	TESTIS WEIGHT (Males): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
Not specified, adult					8.5 11.5 13.2		Scharer
25 w		1.1		2.1	-3.2		Chapin
57 w		-13.3		-0.3	-15.4	-10.8	Keenan
19 w				4.7	3.6	5.1	Seki
68 d (F1)		0.0					Odum

45 d 59d		-1.1 3.0					Marty
21 w (F1)			11.9			24.3*	Carney
53 d	1.4		2.1	4.9			Laws

Rat age at termin-ation of study	EPIDIDYMIDES WEIGHT (Males): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
25 w		1.4		0.6	1.1		Chapin
68 d (F1)		3.1					Odum
45 d 59d		2.4* 4.7*					Marty
21 w (F1)			10.8			23.9*	Carney
53 d	3.6		3.2	10.4*			Laws

Rat age at termin-ation of study	PROSTATE WEIGHT (Males): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
Not specified, adult					31.5* 53.9* 47.5*		Scharer
25 w		19.0*		21.9*	36.2*		Chapin
57 w		-13.8		0.0	21.5	21.5	Keenan
19 w				1.6	18.6*	41.5*	Seki
68 d (F1)		4.9					Odum
45 d 59d		17.8* 7.0*					Marty
21 w (F1)			17.3*			40.0*	Carney
53 d	6.7		12.6	23.3*			Laws

Rat age at termin-ation of study	SEMINAL VESICLES WEIGHT (Males): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
25 w		10.3*		15.5*	24.0*		Chapin
68 d (F1)		6.5					Odum
45 d 59d		22.7* 8.6*					Marty
21 w (F1)			20.5*			43.6*	Carney
53 d	2.7		11.4	30.7*			Laws

[1]Percentage decrease in terminal body weight with respect to control body weight
BW : body weight
* Statistical significance was taken from reference (publication)
Blank cells indicate no data available

Table A 2.4. Decrease in organ weights occurring in female rats with decreased body weights as a consequence of dietary modulation.

Rat age at termin-ation of study	LIVER WEIGHT (Females): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
25 w		8.3		16.6 *	37.3*		Chapin
				21.4*	28.2*	47.8*	Seki
21 w (F1)			28.5*			50.0*	Carney
53 d	13.4		16.8*	29.5*			Laws

Rat age at termin-ation of study	KIDNEY WEIGHT (Females): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
25 w		4.0		13.3 *	24.3*		Chapin
				14.0*	20.7*		Seki
21 w (F1)			16.5*			41.4*	Carney
53 d	7.4*		13.4*	20.8*			Laws

Rat age at termin-ation of study	THYMUS WEIGHT (Females): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
25 w				17.9*	14.7		Seki
21 w (F1)			23.8			31.0*	Carney
53 d							Laws

Rat age at termin-ation of study	SPLEEN WEIGHT (Females): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
25 w				11.2	16.1*		Seki
21 w (F1)			15.5*			32.3*	Carney
53 d							Laws

Rat age at termin-ation of study	ADRENALS WEIGHT (Females): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
25 w				4.2	16.3*		Seki
21 w (F1)			-6.6			27.9*	Carney
53 d	7.1		13.6*	17.0*			Laws

Rat age at termin-ation of study	BRAIN WEIGHT (Females): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
25 w			0.0		-2.4		Keenan
21 w (F1)				-1.9	2.3		Seki
53 d			3.1*			12.6*	Carney

Rat age at termin-ation of study	THYROID WEIGHT (Females): % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref

68 d (F1)			-32.0		20.0*		Keenan
21 w (F1)				20.9*	13.7		Seki

Rat age at termin-ation of study	OVARIES WEIGHT: % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
25 w		1.5		0.7	21.1*		Chapin
			-14.7		1.5		Keenan
				-0.1	15.0*		Seki
21 w (F1)			2.6			13.2*	Carney
53 d	3.3		21.7*	31.7*			Laws

Rat age at termin-ation of study	UTERUS WEIGHT: % decrease in organ weight						
	5% BW[1]	10% BW	15% BW	20% BW	30% BW	40% BW	Ref
25 w			30.6*		13.9		Keenan
				5.8	11.6		Seki
21 w (F1)			5.9			32.6*	Carney
53 d	6.0		10.8	28.3			Laws

[1]Percentage decrease in terminal body weight with respect to control body weight
BW : body weight
* Statistical significance was taken from reference (publication)
Blank cells indicate no data available

Discussion

Ax2.7. This analysis demonstrated that reductions in absolute liver and kidney weights were strongly associated with body weight reduction (body weight reductions ranging from 5-40% resulted in liver and kidney weight reductions of 10-50% relative to control values). A relationship between thymus or adrenal weights and body weight reductions could not be determined due to inconsistencies in the data. Body weight reductions of 15% and 40% resulted in spleen weight reductions of approximately 20% and 35%, respectively. Brain weight was largely unchanged with reductions in body weight of up to 30% relative to controls, while a body weight reduction of 40% led to a brain weight reduction of up to 12%. For the thyroid, there was some evidence that body weight reduction was associated with thyroid weight reduction. No change in thyroid weight was seen at body weight reductions of up to 15% but body weight reduction of 30% resulted in thyroid weight reductions of approximately 20%.

Ax2.8. In male rats, the weights of testes and epididymides were largely unchanged with reductions in body weight until 40% body weight reduction occurred. With a 15% body weight reduction, testes and epididymides weights were reduced 2-10%, while a 40% body weight reduction resulted in testes and epididymides weights being reduced by 24% in F1 animals. Prostate and seminal vesicle weight varied more with body weight. At 10% body weight reduction, prostate and seminal vesicle weights were reduced 0-20% and at 40% body weight reduction, prostate and seminal vesicle weights were reduced 20-45%.

Ax2.9. In female rats, the weights of ovaries were reduced with reductions in body weight but there was no consistent pattern. With a 15% body weight reduction, ovary weight was reduced by 0-20% whilst at 40% body weight reduction ovary weight was reduced by 13%. Uterus weight also had no consistent pattern. At 15% body weight reduction, uterus weight was reduced by 5-30% whilst at 40% body weight reduction, uterus weight was reduced by 32%. Uterus weight is also greatly influenced by the stage of the oestrous cycle.

Ax2.10. The tabulation of data in this manner may aid in the interpretation of changes in organ weights in the presence of body weight reduction occurring in the EOGRTS. It should be stressed however, that all

relevant data from the study should be considered together when drawing conclusions on the potential for a substance to cause adverse effects in the EOGRTS.