Non-OECD mammalian screens and tests (Concptual Framework Levels 3-5)

C.3.16. Pubertal Development and Thyroid Function Assay in Peripubertal Male Rats (male PP assay) (US EPA OPPTS 890.1500)

Status: Assay validated at national level.

994. Modalities detected: (anti)androgen, thyroid, steroidogenesis.

Endpoints: Age and body weight at preputial separation (PPS). Weight of seminal vesicles (+ coagulating gland), ventral prostate, dorsolateral prostate, LABC (Levator ani plus bulbocavernosus muscle complex), epididymides, testes, thyroid, pituitary, adrenals. Histopathologic changes in epididymis, testis, thyroid. Serum testosterone, T4 and TSH.

Background to the assay

995. This assay is designed to identify chemicals that have the potential to interact with androgen receptor (AR-) mediated modalities, thyroid hormone mediated modalities and interference with steroidogenesis. It will also detect chemicals that alter pubertal development via changes in the hypothalamic/pituitary/gonadal (HPG) axis. It will also detect estrogen receptor (ER-) mediated effects, but the accuracy of this is unknown. The principle of the assay is that male rats are dosed with chemical during the period of sexual maturation, starting at postnatal day 23. Route of administration of test substance is via oral gavage. The prepubertal period is a very sensitive age for exposure to agents which alter the endocrine system (US EPA, 2007). Serum androgens in male rats change dramatically during puberty and reproductive organ weights grow rapidly during puberty (Stoker et al., 2000). PPS is an apical measure of the progression of puberty and has been used as the primary endpoint of puberty onset in the rat. It is an androgen-dependent event. The assay has its female counterpart in the peripubertal female rat assay. Male rats achieve sexual maturity at a later age than females (vaginal opening) and therefore the male assay is of longer duration than the female assay (31 days cf. 21 days) and this should be taken into account when comparing the severity of effects obtained in the two assays.

996. The male PP assay was designed to be one of the suite of assays comprising the United States Environmental Protection Agency's (US EPA) Endocrine Disruptor Screening Program (EDSP) "Tier 1" and has been validated in that context (US EPA, 2007). There is no OECD test guideline for the assay. The US EPA guideline (OPPTS 890.1500) was published in October 2009 (US EPA, 2009). Male and female PP assays are considered to be apical assays (i.e. they contain endpoints that may be changed by a number of different modes of action [MOA] and may not be specific to endocrine active substances [EASs]). The animals have intact hypothalamus-pituitary-gonadal/thyroid axes and therefore are a relevant model for human health, although the sensitivity of the assays for ER/AR agonists and antagonists are less than that of the Uterotrophic Bioassay (UT) and Hershberger Bioassay (H). A strength of the PP assays is that (unlike the H and UT assays) they will detect multiple MOA, although it may not be possible to isolate the mechanism of action. The male PP assay is likely to detect (anti)estrogens in addition to androgen/thyroid/steroidogenesis (ATS) modalities. The estrogen methoxychlor was

included in the validation studies of the male assay and gave a weak positive response for some endpoints. Published studies have also demonstrated that the assay responds to strong estrogens such as diethylstilbestrol (Ashby and Lefevre, 2000) and weak estrogens such as nonylphenol (Tan, Kassim and Mohd, 2003). The validation of the male PP assay indicated that sensitivity was high and although it has not been extensively investigated, it showed that the male pubertal assay can be sensitive to dose levels that are near the lowest observed effect level (LOEL) in a developmental toxicity study on the androgen antagonist vinclozolin (US EPA, 2007).

997. A limitation of the validation is that no chemical was shown to be completely negative in the assay. Chloronitrobenzene was included in the validation as a chemical that was expected to be toxic but without endocrine activity, but when tested was positive, delaying PPS, decreasing serum testosterone, decreasing growth of androgen-dependent tissues and reducing T4 levels. It is not known whether these effects were due to non-specificity of the assay or a real effect on endocrine systems. Other chemicals, however, that were positive for one endocrine system were not necessarily positive on others (e.g. perchlorate altered thyroid hormones and thyroid weight but caused no effects on any of the reproductive tract weights or puberty onset). This indicates that false positives are not always seen and helps to reinforce the specificity of the assay. Another possible limitation is the inability to detect specific aromatase inhibitors. Although more general inhibitors of steroidogenesis (including aromatase inhibition), such as ketoconazole, are detected in the assay, specific inhibitors of aromatase only, such as fadrozole, were not (Marty, Crissman and Carney, 2001).

998. Experience with of serum hormone determinations in Level 4 and Level 5 rodent assays has revealed that their detection/measurement in rodent studies can be challenging. A recent workshop on "Practicability of Hormonal Measurements" was organised by the BfR (Germany) and the finding from this workshop will be published (Kucheryavenko et al., 2018). The OECD Expert Group on Reproductive and Developmental Toxicity recommends that to demonstrate proficiency for thyroid hormones measurement, a laboratory should be able to show results from a separate study using a positive control substance. Laboratories may also submit their calibration curves, standard curves, as well as data on the levels of quantification and detection. This group is also establishing a historical control database with thyroid toxicant positive controls.

When/why the assay may be used

999. As mentioned above, the male PP assay may be used as part of the US EPA's EDSP Tier 1 screening battery as an apical assay to detect interaction with multiple endocrine systems. In this context, its use is primarily for hazard determination. It may also be used as a follow-up assay following positive results in *in vitro* assays (e.g. a positive result in the Steroidogenesis Assay). Positive results in an AR *in vitro* assay would preferably be followed by an H assay for reasons of animal welfare – H assays require fewer animals than the male PP assays and are of shorter duration. If there is a need to test in an apical assay, the PP assay may be chosen, realising the caveat that there is some uncertainty regarding its specificity. Depending on the number of doses used, the PP assay may be used for hazard identification/characterisation. The assay could potentially also be used to investigate or supplement higher tier data, possibly to clarify the MOA. One scenario could be if only limited reproductive data are available (e.g. a study not conducted to modern standards or not containing endpoints for sexual development). Data from female and male PP assays could then be used to investigate the occurrence of endocrine effects. A decision

about whether to conduct further animal tests would, however, need to consider whether sufficient supplementary data may be provided by *in vitro* tests.

1000. In order to provide information relevant for assessing whether or not a chemical may fulfil the WHO/IPCS (2002) definition of an endocrine disruptor (ED), the study design has to be sufficiently robust to demonstrate the presence or absence of effects. In the dose selection, the investigator should also consider and ensure that data generated are adequate to fulfil the regulatory requirement across OECD countries as appropriate (e.g. hazard and risk assessment and labelling, ED assessment, etc.). The top dose or concentration should be sufficiently high to give clear systemic (i.e. non endocrine-specific) toxicity in order to ensure that a wide range of exposures (high to low) is tested. However, endocrine effects observed solely in the presence of clear systemic toxicity should be interpreted with caution and may be disregarded when sufficiently justified to be caused by secondary effects which are unlikely to be due to endocrine activity. The reason for this advice is a concern that some EAS-sensitive assays are being run at doses/concentrations of EASs that are too low to trigger direct impacts on the endocrine system. This guidance document is not the place to address this issue directly, but it should be considered when EAS-sensitive test guidelines (TGs) are revised in the future. In addition, the number and spacing of dose/concentration levels should also be adequate to fulfil the objectives of the study (e.g. to demonstrate dose response relationships if this is required).

Introduction to the table of scenarios

1001. Table C.3.16 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. “Existing results” are subdivided into “mechanism” and “effects” data (third and fourth columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

1002. The results of the male PP assay are given in the second column. The assay contains multiple endpoints and it is not possible to provide alternative scenarios for all combinations, therefore some discrimination has been attempted by dividing the endpoints into “apical” and “indicators of hormonal activity”. The terminology used has been chosen to be consistent between both the non-mammalian and mammalian tests. Both groups have similar biological importance, although the “indicators of hormonal activity” in the mammalian assays are serum hormones and are generally, but not always, more variable than “apical endpoints”. “Apical endpoints” are age/body weight at PPS, weights of seminal vesicles, prostate, LABC, epididymides, testes, thyroid, pituitary and adrenals; histopathologic changes in epididymis, testis, thyroid. “Indicators of hormonal activity” are hormones (testosterone, T4 and TSH).

1003. Three possible outcomes for a positive result are therefore envisaged in Table C.3.16:

1. indicators of hormonal activity and apical endpoints positive
2. indicators of hormonal activity positive and apical endpoints negative
3. indicators of hormonal activity negative and apical endpoints positive.

1004. A positive result for apical endpoints could be delayed puberty (preputial separation) or biologically significant reductions in weights of the epididymides, prostate and seminal vesicles accompanied by treatment-related histopathologic changes. A positive result for indicators of hormonal activity could be biologically significant changes in thyroid hormone profiles. The multiple endpoints in this assay mean that there is some redundancy in the

assay, but this is useful as not all chemicals may affect all endpoints associated with a mechanism of action and there may be site-specific differences in response.

1005. Single isolated changes may be indicative of spurious results, but robust dose response information may not be available as the TG only requires two dose levels. The guidance on histopathologic changes in endocrine tests (OECD, 2009) may be helpful in interpretation. Such results should be considered with caution, although it is possible that weak effects have been detected which may then be seen in longer term studies.

1006. A negative result for the male PP assay is taken to be the absence of changes in indicators of hormonal activity and apical endpoints. In the absence of other pertinent lines of evidence, negative results in this test alone cannot be taken as evidence that the substance is not an ED. Further studies will be required as confirmation.

1007. Equivocal results for the guideline are not considered in the table, partly for brevity but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative or whether the result must be put to one side and the test repeated (using the same or a different test). Factors which may have interfered with the result (e.g. composition of the diet used, environmental influences) should be considered. Performance criteria (coefficients of variation for the test endpoints) should be checked for compliance with those in the TG. The assay does not include concurrent positive controls, but attempts have been made to mitigate this by including the performance criteria.

Existing data to be considered

1008. Existing "mechanism" *in vitro* data are assumed to be available from ER-, AR- and steroidogenesis-based assays (Level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available, so judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs et al. (2008; 2013) and OECD (2008). These methods, however, have not yet been validated.

1009. Existing "effects" data refer to *in vivo* effects that may come from H assays where a non-physiological animal model is used. In these cases, it should be remembered that these assays are specifically designed to be sensitive to EDs. Another possibility is that repeat dose oral toxicity studies, reproduction/developmental toxicity screen tests or read-across from analogues may be available. It is unlikely that the male PP assay will be performed if data from robust higher tier reproductive studies are already available as the PP assay offers no advantage over these assays. It is possible, though, that the PP assay has been performed to supplement non-robust higher tier data for the reasons given above. Data may also be available on effects in non-mammalian species, although caution should be used when extrapolating between taxa. A chemical causing endocrine effects in non-mammalian environmental species (fish, for example) may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

1010. When considering the results of the male PP assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include high throughput screening data, read-across data from structural analogues and quantitative structure activity relationship (QSAR). Several QSAR models for ER and AR binding/activation are now available (see Sections B.1.1.1 and B.1.1.2).

Scenarios: Positive and negative results combined with existing data

1011. The scenarios (A to R) presented in Table C.3.16 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend on the regulatory environment, but the considerations given here are generally science based. Although the male PP assay uses rats, the well-conserved nature of the hormonal pathways across taxa should be a strong indication that results in this assay may be relevant to other vertebrate species. Effects in laboratory mammal tests are also highly relevant for environmental mammalian species. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain, exposure route and species-specific metabolism should always be considered. The sensitivity and physiological function of the hormone under investigation in the test species should also be considered. In general, lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage, unless it is apparent that a Level 5 test will be required anyway or will be needed to establish the evidence to conclude on ED properties. Information on some endocrine-related tumours may be detected more comprehensively in carcinogenicity studies (OECD TG 451/453) (Level 4); for example, detection of certain types of thyroid tumors in the absence of reproductive or developmental effects, as well as substances causing tumors in other endocrine-sensitive tissues. At Level 5, the EOGRTS (OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised however, that some jurisdictions may require a two-generation study. Further considerations specific to each scenario are given in the table.

1012. Scenarios A to C represent positive results in the male PP assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive male PP result scenario is divided into the three possible outcomes given above. A positive result in the *in vitro* assays in combination with a positive male PP assay is moderate or strong evidence for E,A,T,S-mediated activity that may or may not be supported by the *in vivo* effects data. In the absence of robust upper-level data, the next step may be to conduct an upper-level test. In the presence of robust Level 5 data, there may be sufficient evidence to conclude concern for endocrine disruption and therefore no need for further screening. Positive results in the male PP assay may also indicate the potential for endocrine mediated effects in lower vertebrates. These could be followed up with partial life cycle tests such as the Fish Sexual Development Test (FSDT) or the Medaka Extended One-Generation Reproduction Test (MEOGRT); or the Larval Amphibian Growth and Development Assay (LAGDA) if effects are on the thyroid hormone system. *In vivo* assays/tests with negative results should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or, alternatively, that the effects do not present a concern for endocrine disruption. The possibility of other (non-E,A,T,S) mechanisms should also not be overlooked (e.g. involving other receptors or endocrine axes).

1013. Scenarios D to F represent positive results in the male PP assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive male PP result scenario is divided into the three possible outcomes given above. Negative results in the *in vitro* assays should be viewed with caution in case a metabolite is responsible for the positive male PP assay. Unless the metabolic profile of the test substance is known, one option may be to conduct these assays with an added

metabolising system. If the metabolic profile is known, then a higher level *in vivo* test may be advisable. The choice of tests will depend on the available *in vivo* effects data. Positive results in the male PP assay may also indicate the potential for endocrine mediated effects in lower vertebrates. These could be followed up with partial life cycle tests such as the FSDT or MEOGRT; or the LAGDA if effects are on the thyroid hormone system. As in Scenarios A to C, *in vivo* assays/tests with negative results should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or, alternatively, that the effects do not present a concern for endocrine disruption.

1014. Scenarios G to I represent positive results in the male PP assay in the presence of various combinations of missing or equivocal data. Positive results in the male PP assay may also indicate the potential for endocrine mediated effects in lower vertebrates. These could be followed up with partial life cycle tests such as the FSDT or MEOGRT; or the LAGDA if effects are on the thyroid hormone system. Each positive male PP result scenario is divided into the three possible outcomes given above. The next step to take in these eventualities will depend on the nature of the other available data and the jurisdiction in which it is being used. In some cases, equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the mode of action (MOA) in question and why the data are considered equivocal, e.g. a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases, but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

1015. Scenarios J to L represent negative results in the male PP assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. As a negative result for the male PP is taken to be negative findings for both indicators of hormonal activity and apical endpoints (unlike the situation with positive outcomes), there is only one possible negative outcome. All three scenarios could also arise from a chemical that is positive in *in vitro* assays, but is metabolised to a non-active metabolite leading to negative results in the male PP assay. This should be considered first when investigating the next step. Endocrine active potency may also explain differences between *in vitro* and *in vivo* results (e.g. a chemical with weak endocrine activity may give a positive result *in vitro* but may be negative *in vivo*). Positive *in vivo* effects data may involve other E,A,T,S, non-E,A,T,S mechanisms (e.g. involving other receptors or endocrine axes), more sensitive endpoints, greater statistical power or life stages that are more sensitive to the substance than the young adult animals in the male PP assay.

1016. Scenarios M to O represent negative results in the male PP assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend on the weight of evidence and may not be possible. Where there are positive *in vivo* effects data, there could still be an E,A,T,S-related mechanism, the effects may be related to length of exposure, route of exposure or exposure at different life stages. Other E,A,T,S or non-E,A,T,S mechanisms may also be involved.

1017. Scenarios P to R represent negative results in the male PP assay in the presence of various combinations of missing or equivocal data. As with the positive result scenarios above (see Paragraph 1 014), the next step to take in these eventualities will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

1018. In all scenarios (A to R), the next step to take to strengthen weight of evidence will depend on the existing information. Table C.3.16 is meant to provide a succinct guide and may not cover all circumstances or possibilities. The scenarios may also suggest that chemicals have simple or single MOA, when in practice they may have multiple endocrine and non-endocrine MOA. In some cases, for example, two opposite modes of simultaneous action (e.g. estrogenic and anti-estrogenic) could, depending on dose, lead to a minimisation or abolition of effects, while in others two different MOA (e.g. estrogenic and anti-androgenic) could potentially reinforce effects. Endocrine pathways interact, mixed effects are common and there are many pathways that cannot be distinguished with currently available TGs. If multiple MOA are suspected, either from the existing results or based on QSAR/read-across/integrated approaches, this should be investigated further if needed for regulatory decision making.

References

Ashby, J. and P.A. Lefevre (2000), "The peripubertal male rat assay as an alternative to the Hershberger castrated male rat assay for the detection of antiandrogens, oestrogens and metabolic modulators", *Journal of Applied Toxicology*, Vol. 20/1, pp. 35-47, https://doi.org/10.1002/(SICI)1099-1263(200001/02)20:1<35::AID-JAT633>3.0.CO;2-8.

Jacobs, M. et al. (2013), "*In vitro* metabolism and bioavailability tests for endocrine active substances: What is needed next for regulatory purposes?", *ALTEX – Alternatives to Animal Experimentation*, Vol. 30/3, pp. 331-351.

Jacobs, M.N. et al. (2008), "The use of metabolising systems for *in vitro* testing of endocrine disrupters", *Current Drug Metabolism*, Vol. 9/8, pp. 796-826.

Kucheryavenko, O. et al. (2018), "Report from the BfR Expert Hearing on Practicability of Hormonal Measurements", *Archives of Toxicology* (in prep.).

Marty, M.S., J.W. Crissman and E.W. Carney (2001), "Evaluation of the male pubertal onset assay to detect testosterone and steroid biosynthesis inhibitors in CD rats", *Toxicological Sciences*, Vol. 60/1, pp. 285-295, http://dx.doi.org/10.1093/toxsci/60.1.63.

OECD (2009), "Guidance document for histologic evaluation of endocrine and reproductive tests in rodents", OECD Series on Testing and Assessment, No. 106, OECD, Paris, www.oecd.org/env/ehs/testing/43411534.pdf.

OECD (2008), *Detailed Review Paper on the Use of Metabolising Systems for* In Vitro *Testing of Endocrine Disrupters*, OECD Series on Testing and Assessment, No. 97, OECD Publishing, Paris, https://doi.org/10.1787/9789264085497-en.

Stoker, T.E. et al. (2000), "Endocrine disrupting chemicals: Prepubertal exposures and effects on sexual maturation and thyroid function in the male rat. A focus on the EDSTAC recommendations", *Critical Reviews in Toxicology*, Vol. 30/2, pp. 197-252, https://doi.org/10.1080/10408440091159194.

Tan, B.L., N.M. Kassim and M.A. Mohd (2003), "Assessment of pubertal development in juvenile male rats after sub-acute exposure to bisphenol A and nonylphenol", *Toxicology Letters*, Vol. 143/3, pp. 261-270, https://doi.org/10.1016/S0378-4274(03)00172-3.

US EPA (2009), "Endocrine Disrupter Screening Program Test Guidelines OPPTS 890.1500: Pubertal development and thyroid function in intact juvenile/peripubertal male rats", Environmental Protection Agency, Washington, DC, https://www.regulations.gov/document?D=EPA-HQ-OPPT-2009-0576-0010.

US EPA (2007), "Integrated summary report for validation of a test method for assessment of pubertal development and thyroid function in juvenile male rats as a potential screen in the Endocrine Disrupter Screening Program Tier-1 battery", Environmental Protection Agency, Washington, DC.

WHO/IPCS (2002), "Global assessment of the state-of-the-science of endocrine disrupters", Damstra, T. et al. (eds.) WHO/PCS/EDC/02.2, World Health Organization, Geneva, www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en.

Table C.3.16. **Pubertal Development and Thyroid Function Assay in Peripubertal Male Rats (male PP assay) (OPPTS 890.1500): Guidance for scenarios of combinations of results with existing data**

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The guidance offered is not meant to be prescriptive, but provides science-based considerations. It encourages the use of all available data and expert judgement in a weight of evidence approach. Regional and national interpretation of results and "next steps" may vary.

The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data and existing *in vivo* data. The symbol "+" indicates that the data in question represent a positive result, «-" indicates a negative result, and "Eq/0" indicates that the data are either equivocal or are not available.

Existing results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from estrogen receptor (ER-), androgen receptor (AR-) and steroidogenesis-based assays (Level 2). Thyroid hormone receptor (TR) and other assays concerning mechanisms of thyroid disruption may be available, but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". Quantitative structure activity relationship (QSAR) predictions of estrogen and androgen binding/activation may be made for some substances.

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

*** *Note:* three possible outcomes for a positive result are given:

1. indicators of hormonal activity and apical endpoints positive
2. indicators of hormonal activity positive and apical endpoints negative
3. indicators of hormonal activity negative and apical endpoints positive.

"Apical endpoints" are age/body weight at PPS; weights of seminal vesicles, prostate, LABC, epididymides, testes, thyroid, pituitary and adrenals; histopathologic changes in epididymis, testis, thyroid.

"Indicators of hormonal activity" are hormones (testosterone, T4 and TSH).

Scenarios	Result of male PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+ ***	+	+	1) Increased evidence of (anti)-A,T,S activity (weak, moderate or strong). 2) Possible evidence of (anti)-A,T,S activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-A,T,S activity. 3) Moderate or strong (anti)-A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-A,T,S activity.	Perform assay from Level 5 (e.g. Extended One-Generation Reproduction Toxiticy Study [EOGRTS] or two-generation assay).	If existing data are from a Level 5 assay, there may be sufficient information to conclude evidence of concern for endocrine disruption (the EOGRTS provides the most information; however, for endocrine disrupting chemicals [EDCs] with a carcinogenic potential, OECD TG 451-3 may be more sensitive). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate E,A,T,S modalities or other mechanisms. Possible effects on estrogen modality should also be considered. Consider route of exposures for effects data and possible implications of absorption, distribution, metabolism and excretion (ADME) characteristics of the chemical. Hormonal activity possible in lower vertebrates. Consider performing a Fish Sexual Development Test (FSDT), Larval Amphibian Growth and Development Assay (LAGDA) or Medaka Extended One-Generation Reproduction Test (MEOGRT).
B	+	+	–	1) Increased evidence of (anti)-A,T,S activity (weak, moderate or strong). 2) Possible evidence of (anti)-A,T,S activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-A,T,S activity. 3) Moderate or strong (anti)-A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-A,T,S activity.	Perform assay from Level 5 (e.g. EOGRTS or two-generation assay).	Question why there are differences from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay, a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate A,T,S modalities or other mechanisms. Possible effects on estrogen modality should also be considered. Hormonal activity possible in lower vertebrates. Consider performing a FSDT, LAGDA or MEOGRT.

Scenarios	Result of male PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
C	+	+	Eq/0	1) Increased evidence of (anti)-A,T,S activity (weak, moderate or strong). 2) Possible evidence of (anti)-A,T,S activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-A,T,S activity. 3) Moderate or strong (anti)-A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-A,T,S activity.	Perform assay from Level 5 (e.g. EOGRTS or two-generation assay).	Check data on chemical analogues. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate A,T,S modalities or other mechanisms. Possible effects on estrogen modality should also be considered. Consider route of exposure for female Peripubertal (PP) Assay and follow-up assay. Possible implications of ADME characteristics of the chemical. Hormonal activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT. Equivocal results may indicate chemical has multiple modes of action (MOA).
D	+	–	+	1) Increased evidence of (anti)-A,T,S activity (weak, moderate or strong). Acts via non-AR, TR, S mechanism or may require metabolic activation for activity. 2) Possible evidence of (anti)-A,T,S activity, apical endpoints may be less sensitive or unaffected. Acts via non-AR, TR, S mechanism or may require metabolic activation for activity. 3) Moderate or strong (anti)-A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-AR, TR, S mechanism or may require metabolic activation for activity.	Perform *in vitro* estrogen receptor (ER-), androgen receptor (AR-), thyroid hormone receptor (TR-), steroidogenesis (S) assays with added metabolising system.	If existing data are from an adequate Level 5 assay, there is sufficient information to conclude evidence of concern for endocrine disruption (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate A,T,S modalities or other mechanisms. Possible effects on estrogen modality should also be considered. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical. Further mechanistic studies may help determine MOA. Hormonal activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT.

Scenarios	Result of male PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
E	+	–	–	1) Increased evidence of (anti)-A,T,S activity (weak, moderate or strong). Acts via non-AR, TR, S mechanism or may require metabolic activation for activity. Route of exposure may account for the differences from existing data. 2) Possible evidence of (anti)-A,T,S activity, apical endpoints may be less sensitive or unaffected. Acts via non-AR, TR, S mechanism or may require metabolic activation for activity. Route of exposure may account for the differences from existing data. 3) Possible evidence of (anti)-A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-AR, TR, S mechanism or may require metabolic activation for activity. Route of exposure may account for the differences from existing data.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation assay).	Question why there are differences from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay, a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate other mechanisms. Hormonal activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT.
F	+	–	Eq/0	1) Increased evidence of (anti)-A,T,S activity (weak, moderate or strong). Acts via non-AR, TR, S mechanism or may require metabolic activation for activity. 2) Possible evidence of (anti)-A,T,S activity, apical endpoints may be less sensitive or unaffected. Acts via non-AR,TR, S mechanism or may require metabolic activation for activity. 3) Moderate (anti)-E,A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-AR, TR, S mechanism or may require metabolic activation for activity.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation assay).	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate other mechanisms. Check data on chemical analogues. Further mechanistic studies may help determine MOA. If existing data are from an adequate Level 5 assay, question why there are differences. If existing data are from a less sensitive assay, a higher level test may be required. Hormonal activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT. Equivocal results may indicate chemical has multiple MOA.

Scenarios	Result of male PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
G	+	Eq/0	+	1) Increased evidence of (anti)-A,T,S activity (weak, moderate or strong). May act via AR, TR, S mechanism (metabolic activation may be needed). 2) Possible evidence of (anti)-A,T,S activity, apical endpoints may be less sensitive or unaffected. May act via AR, TR, S mechanism (metabolic activation may be needed). 3) Moderate or strong (anti)-A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-A,T,S activity. May act via AR, TR, S mechanism (metabolic activation may be needed).	Perform *in vitro* ER, AR, TR, S assays with added metabolising system (for the "0" scenario, otherwise Eq result available).	If existing data are from a Level 5 asssay, there may be sufficient information to conclude evidence of concern for endocrine disruption (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate A,T,S modalities or other mechanisms. Possible effects on estrogen modality should also be considered. Check data on chemical analogues. Further mechanistic studies may help determine MOA. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical. Equivocal results may indicate chemical has multiple MOA.
H	+	Eq/0	–	1) Increased evidence of (anti)-A,T,S activity (weak, moderate or strong). Acts via unknown mechanism or may require metabolic activation for activity. Route of exposure may account for the differences from existing data. 2) Possible evidence of (anti)-A,T,S activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism. Route of exposure may account for the differences from existing data. 3) Moderate (anti)- A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via unknown mechanism. Route of exposure may account for the differences from existing data.	For the "0" scenario, perform *in vitro* ER, AR, TR, S assays with added metabolising system. (otherwise Eq result available).	Question why there are differences from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay, a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of subtle changes not detected by apical endpoints. Effects on apical endpoints alone may indicate A,T,S modalities or other mechanisms. Possible effects on estrogen modality should also be considered. Check data on chemical analogues. Further mechanistic studies may help determine MOA. Hormonal activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT. Equivocal results may indicate chemical has multiple MOA.

Scenarios	Result of male PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
I	+	Eq/0	Eq/0	1) Increased evidence of (anti)-A,T,S activity (weak, moderate or strong). Acts via unknown mechanism. Unknown potential for adverse effects. 2) Possible evidence of (anti)-E,A,T,S activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. 3) Moderate or strong (anti)-A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. There may be a need for metabolic activation.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation assay).	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate A,T,S modalities or other mechanisms. Possible effects on estrogen modality should also be considered. Check data on chemical analogues. Further mechanistic studies may help determine MOA. Hormonal activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT. Equivocal results may indicate chemical has multiple MOA.
J	–	+	+	No evidence for A,T,S activity in male PP assay. Metabolism or potency explains the difference from existing *in vitro* and *in vivo* data. Effects seen in existing studies are via non-A,T,S mechanism.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation assay).	If existing data are from an adequate Level 5 assay, question why there are differences. If data are from a Hershberger Bioassay (H), this may be more sensitive than male Peripubertal (PP) Assay. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical. Further mechanistic studies may help determine MOA.
K	–	+	–	No evidence for A,T,S activity in male PP assay. Metabolism or potency explains *in vitro*/*in vivo* differences.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation assay).	If existing data are from an adequate Level 5 assay, there may be sufficient information to conclude absence of concern for endocrine disruption (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive). If existing data are from a less sensitive assay, a higher level test may be required. If data are from H assay, need to conduct higher tier assay to conclude absence of concern for endocrine disruption. Further mechanistic studies with metabolism may help determine MOA.

Scenarios	Result of male PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
L	–	+	Eq/0	No evidence for A,T,S activity in male PP assay. Metabolism or potency explains *in vitro*/*in vivo* differences. Unknown potential for adverse effects.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation assay).	Metabolic deactivation of chemical may occur *in vivo* so that possible *in vitro* activity is not realised. Consider possible routes of exposure implications of metabolism. Equivocal results may indicate chemical has multiple MOA.
M	–	–	+	No evidence for A,T,S activity in male PP assay. Effects seen in existing studies are via non-A,T,S mechanism.	Perform assay from Level 5 (e.g. EOGRTS or two-generation assay).	If existing data are from an adequate Level 5 assay, question why there are differences (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive). If data are from H assay, this may be more sensitive than male PP assay. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical.
N	–	–	–	No evidence for A,T,S activity in male PP assay. No evidence for (anti)-A,T,S activity *in vitro*. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	If existing data are from an adequate Level 5 assay, there may be sufficient information to conclude absence of concern for endocrine disruption (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive).
O	–	–	Eq/0	No evidence for A,T,S activity in male PP assay. No evidence for (anti)-A,T,S activity *in vitro*. Unknown potential for adverse effects.	Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	Consider route of exposures and possible implications for ADME characteristics of the chemical in follow-up assay.
P	–	Eq/0	+	No evidence for A,T,S activity in female PP assay. Potential for adverse effects via unknown mechanism.	Perform *in vitro* ER, AR,TR, S assays.	Consider route of exposure and possible implications for differences from existing assay. If data are from H assay, this may be more sensitive than male PP assay. Effects seen in existing studies may be in a more sensitive life stage. Equivocal results may indicate chemical has multiple MOA.

Scenarios	Result of male PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
Q	–	Eq/0	–	No evidence for A,T,S activity in male PP assay. No evidence of adverse effects.	Perform *in vitro* ER, AR,TR, S assays.	If existing data are from an adequate Level 5 assay, there may be sufficient information to conclude absence of concern for endocrine disruption (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive). Further mechanistic studies may strengthen weight of evidence.
R	–	Eq/0	Eq/0	No evidence for A,T,S activity in male PP assay.	Perform *in vitro* ER, AR,TR, S assays, otherwise Eq result available.	Further mechanistic studies may strengthen weight of evidence. Check data on chemical analogues. Equivocal results may indicate chemical has multiple MOA.

C.3.17. Pubertal Development and Thyroid Function Assay in Peripubertal Female Rats (female PP assay) (US EPA OPPTS 890.1450)

Status: Assay validated at national level.

1019. Modalities detected: (anti)estrogen, thyroid, steroidogenesis.

Endpoints: Age and body weight at vaginal opening (VO). Weight of ovaries, uterus, thyroid, pituitary, adrenals. Histopathologic changes in ovaries, uterus, thyroid. Serum T4 and TSH. Age at first vaginal estrus after VO, estrus cyclicity parameters.

Background to the assay

1020. This assay is designed to identify chemicals that have the potential to interact with estrogen receptor (ER-) mediated modalities, thyroid hormone mediated modalities and interference with steroidogenesis. It will also detect chemicals that alter pubertal development via changes in the hypothalamic/pituitary/gonadal (HPG) axis. The principle of the assay is that female rats are dosed with chemical during period of sexual maturation, starting at postnatal day 22. Route of administration of test substance is via oral gavage. The prepubertal period is a very sensitive age for exposure to agents which alter the endocrine system (Goldman et al., 2000). Sexual maturation is determined in females as VO (or patency) and is an estrogen-dependent event that follows the first period of ovarian follicular growth (Goldman et al., 2000). The assay has its male counterpart in the peripubertal (PP) male rat assay. Female rats achieve sexual maturity at an earlier age than males (preputial separation) and therefore the female assay is of shorter duration than the male assay (21 days cf. 31 days) and this should be taken into account when comparing the severity of effects obtained in the two assays.

1021. The female PP assay was designed to be one of the suite of assays comprising the United States Environmental Protection Agency's Endocrine Disruptor Screening Program "Tier 1" and has been validated in that context (US EPA, 2007). There is no OECD test guideline for the assay. The US EPA guideline (OPPTS 890.1450) was published in October 2009 (US EPA, 2009). Male and female PP assays are considered to be apical assays (i.e. they contain endpoints that may be changed by a number of different modes of action [MOA] and may not be specific to endocrine active substances [EASs]). The animals have intact hypothalamus-pituitary-gonadal/thyroid axes and therefore are a relevant model for human health, although the sensitivity of the assays for estrogen receptor/androgen receptor (ER/AR) agonists and antagonists are less than that of the Uterotrophic Bioassay (UT) and Hersgberger Bioassay (H). A strength of the PP assays is that (unlike the H and UT assays) they will detect multiple MOA, although it may not be possible to isolate the mechanism of action. The female PP assay is likely to detect (anti)androgens in addition to E,T,S modalities, although androgens and anti-androgens were not included in the validation studies of the female assay. The validation of the female PP assay indicated that sensitivity was high and although it has not been extensively investigated, it appeared to provide a good estimate of the no-observed-effect-concentration/lowest-observed-effect-

concentration (NOEL/LOELs) obtained in studies of similar or longer duration (e.g. the LOAEL for ethinylestradiol in the female PP assay was similar to that for reproductive effects in a multigenerational study) (US EPA, 2007).

1022. A limitation of the validation is that no chemical was shown to be completely negative in the assay. Chloronitrobenzene was included in the validation as a chemical that was expected to be toxic but without endocrine activity, but when tested was positive in the assay, delaying VO, reducing uterine weight, reducing T4 levels and increasing TSH levels. It is not known whether these effects were due to non-specificity of the assay or a real effect on endocrine systems. Other chemicals, however, that were positive for one endocrine system were not necessarily positive on others (e.g. propylthiouracil altered thyroid hormones and thyroid weight but caused no effects on any of the reproductive tract weights or puberty onset). This indicates that false positives are not always seen and helps to reinforce the specificity of the assay.

1023. Experience with serum hormone determinations in Level 4 and Level 5 rodent assays has revealed that their detection/measurement in rodent studies can be challenging. A recent workshop on "Practicability of Hormonal Measurements" was organised by the BfR (Germany) and the finding from this workshop will be published (Kucheryavenko et al., 2018). The OECD Expert Group on Reproductive and Developmental Toxicity recommends that to demonstrate proficiency for thyroid hormones measurement, a laboratory should be able to show results from a separate study using a positive control substance. Laboratories may also submit their calibration curves, standard curves, as well as data on the levels of quantification and detection. This group is also establishing a historical control database with thyroid toxicant positive controls.

When/why the assay may be used

1024. As mentioned above, the female PP assay may be used as part of the US EPA's Tier 1 screening battery as an apical assay to detect interaction with multiple endocrine systems. In this context, its use is primarily for hazard determination. It may also be used as a follow-up assay following positive results in *in vitro* assays (e.g. a positive result in the Steroidogenesis Assay). Positive results in an ER *in vitro* assay would preferably be followed by a UT assay for reasons of animal welfare – UT assays require fewer animals than the female PP assays and are of shorter duration. If there is a need to test in an apical assay, then the PP assay may be chosen, realising the caveat that there is some uncertainty regarding the specificity of the PP assay. Depending on the number of doses used, the PP assay may be used for hazard identification/characterisation. The assay could potentially also be used to investigate or supplement higher tier data, possibly to clarify the MOA. One scenario could be if only limited reproductive data are available (e.g. a study not conducted to modern standards or not containing endpoints for sexual development). Data from female and male PP assays could then be used to investigate the occurrence of endocrine effects. A decision about whether to conduct further animal tests would, however, need to consider whether sufficient supplementary data may be provided by *in vitro* tests.

1025. In order to provide information relevant for assessing whether or not a chemical may fulfil the WHO/IPCS (2002) definition of an endocrine disruptor (ED), the study design has to be sufficiently robust to demonstrate the presence or absence of effects. In the dose selection, the investigator should also consider and ensure that data generated are adequate to fulfil the regulatory requirement across OECD countries as appropriate (e.g. hazard and risk assessment and labelling, ED assessment, etc.). The top dose or concentration should be sufficiently high to give clear systemic (i.e. non endocrine-specific) toxicity in order

to ensure that a wide range of exposures (high to low) is tested. However, endocrine effects observed solely in the presence of clear systemic toxicity should be interpreted with caution and may be disregarded when sufficiently justified to be caused by secondary effects which are unlikely to be due to endocrine activity. The reason for this advice is a concern that some EAS-sensitive assays are being run at doses/concentrations of EASs that are too low to trigger direct impacts on the endocrine system. This guidance document is not the place to address this issue directly, but it should be considered when EAS-sensitive test guidelines (TGs) are revised in the future. In addition, the number and spacing of dose/concentration levels should also be adequate to fulfil the objectives of the study (e.g. to demonstrate dose response relationships if this is required).

Introduction to the table of scenarios

1026. Table C.3.17 gives guidance on a further step to take in the event of a positive (+) or negative (-) result and in the presence of positive (+), negative (-) or equivocal/absent (Eq/0) existing results. "Existing results" are subdivided into "mechanism" and "effects" data (third and fourth columns). The table is divided horizontally into a series of scenarios that represent all the combinations of these events.

1027. The results of the female PP assay are given in the second column. The assay contains multiple endpoints and it is not possible to provide alternative scenarios for all combinations, therefore some discrimination has been attempted by dividing the endpoints into "apical" and "indicators of hormonal activity". The terminology used has been chosen to be consistent between both the non-mammalian and mammalian tests. Both groups have similar biological importance, although the "indicators of hormonal activity" in the mammalian assays are serum hormones and are generally, but not always, more variable than "apical endpoints". "Apical endpoints" are age/body weight at VO, estrus cyclicity parameters, weights of ovaries, uterus, thyroid, pituitary and adrenals; histopathologic changes in ovaries and uterus. "Indicators of hormonal activity" are hormones (T4 and TSH).

1028. Three possible outcomes for a positive result are therefore envisaged in Table C.3.17:

1. indicators of hormonal activity and apical endpoints positive
2. indicators of hormonal activity positive and apical endpoints negative
3. indicators of hormonal activity negative and apical endpoints positive.

1029. A positive result for apical endpoints could be delayed puberty (VO) or biologically significant reductions in uterine weights, accompanied by treatment-related histopathologic changes. A positive result for indicators of hormonal activity could be biologically significant changes in hormone profiles. The multiple endpoints in this assay mean that there is some redundancy in the assay, but this is useful as not all chemicals may affect all endpoints associated with a mechanism of action and there may be site-specific differences in response.

1030. Single isolated changes may be indicative of spurious results, but robust dose response information may not be available, as the TG only requires two dose levels. The guidance on histopathologic changes in endocrine tests (OECD, 2009) may be helpful in interpretation. Such results should be considered with caution, although it is possible that these endpoints may have detected weak effects that were not detected by the apical endpoints in this study but may then be detected in longer term studies.

1031. A negative result for the female PP assay is taken to be the absence of changes in both endocrine relevant indicators of hormonal activity and apical endpoints. In the absence of other pertinent lines of evidence, negative results in this test alone cannot be taken as evidence that the substance is not an ED. Further studies will be required as confirmation.

1032. Equivocal results for the guideline are not considered in the table, partly for brevity, but also because equivocal results are by nature uncertain. A decision must eventually be reached about whether the endocrine endpoints tend to be positive or negative or whether the result must be put to one side and the test repeated (using the same or a different test). Factors which may have interfered with the result (e.g. composition of the diet used, environmental influences) should be considered. Performance criteria (coefficients of variation for the test endpoints) should be checked for compliance with those in the TG. The assay does not include concurrent positive controls, but attempts have been made to mitigate this by including the performance criteria.

Existing data to be considered

1033. Existing "mechanism" *in vitro* data are assumed to be available from ER-, AR- and steroidogenesis-based assays (Level 2). Assays may also be available for interference with thyroid modalities. In practice, it is possible that data from all of these assays may not be available, so judgement will need to be used to decide which assays to perform. Although the current *in vitro* test guidelines do not incorporate metabolic activation, published information on use of metabolic activation systems is available in Jacobs et al. (2008; 2013) and OECD (2008). These methods, however, have not yet been validated.

1034. Existing "effects" data refer to *in vivo* effects that may come from UT assays where a non-physiological animal model is used. In these cases, it should be remembered that these assays are specifically designed to be sensitive to EDs. The immature rodent UT assay is also sensitive to activities other than ER (ant)agonism, including changes resulting from energy intake (Odum et al., 2004). Another possibility is that repeat dose oral toxicity studies, reproduction/developmental toxicity screen tests or read-across from analogues may be available. It is unlikely that the female PP assay will be performed if data from robust higher tier reproductive studies are already available, as the PP assay offers no advantage over these assays. It is possible, though, that the PP assay has been performed to supplement non-robust higher tier data for the reasons given above. Data may also be available on effects in non-mammalian species, although caution should be used when extrapolating between taxa. A chemical causing endocrine effects non-mammalian environmental species (fish, for example) may also have endocrine effects in mammals, but the physiological consequences of the effects are likely to be different.

1035. When considering the results of the female PP assay, all available data should be used in order to reach a conclusion and a weight of evidence approach taken. This may include high throughput screening data, read-across data from structural analogues and quantitative structure activity relationship (QSAR). Several QSAR models for ER and AR binding/activation are now available (see Sections B.1.1.1 and B.1.1.2).

Scenarios: Positive and negative results combined with existing data

1036. The scenarios (A to R) presented in Table C.3.17 represent all the possibilities of positive or negative results in combination with the presence or absence of existing data. The action taken will also depend on the regulatory environment, but the considerations given here are generally science based. Although the female PP assay uses rats, the

well-conserved nature of the hormonal pathways across taxa should be a strong indication that results in this assay may be relevant to other vertebrate species. Effects in laboratory mammal tests are also highly relevant for environmental mammalian species. Wherever possible, the recommended "next step which could be taken" avoids unnecessary animal testing. However, sometimes conducting an animal test will be indicated and then the relevance of species, strain, exposure route route and species-specific metabolism should always be considered. The sensitivity and physiological function of the hormone under investigation in the test species should also be considered. In general, lower level tests should be conducted before higher level tests in order to avoid unnecessary animal usage, unless it is apparent that a Level 5 test will be required anyway or will be needed to establish the evidence to conclude on ED properties. Information on some endocrine-related tumours may be detected more comprehensively in carcinogenicity studies (OECD TG 451/453) (Level 4); for example, detection of certain types of thyroid tumors in the absence of reproductive or developmental effects, as well as substances causing tumors in other endocrine-sensitive tissues. At Level 5, the Extended One-Generation Reproductive Toxicity Study (EOGRTS – OECD TG 443) is the most sensitive reproduction assay for detecting endocrine disruption because it includes evaluation of a number of endocrine endpoints not included in the two-generation study (OECD TG 416) adopted in 2001. It is recognised, however, that some jurisdictions may require a two-generation study. Further considerations specific to each scenario are given in the table.

1037. Scenarios A to C represent positive results in the female PP assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive female PP result scenario is divided into the three possible outcomes given above. A positive result in the *in vitro* assays in combination with a positive female PP assay is moderate or strong evidence for estrogen/androgen/thyroid/steroidogenesis (E,A,T,S-) mediated activity that may or may not be supported by the *in vivo* effects data. In the absence of robust upper-level data, the next step may be to conduct an upper-level test. In the presence of robust Level 5 data, there may be sufficient evidence to conclude concern for endocrine disruption and therefore no need for further screening. Positive results in the female PP assay may also indicate the potential for endocrine mediated effects in lower vertebrates. These could be followed up with partial life cycle tests such as the Fish Sexual Development Test (FSDT) or the Medaka Extended One-Generation Reproduction Test (MEOGRT); or the Larval Amphibian Growth and Development Assay (LAGDA) if effects are on the thyroid hormone system. *In vivo* assays/tests with negative results should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or, alternatively, that the effects do not present a concern for endocrine disruption. The possibility of other (non-E,A,T,S) mechanisms should also not be overlooked (e.g. involving other receptors or endocrine axes).

1038. Scenarios D to F represent positive results in the female PP assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Each positive female PP result scenario is divided into the three possible outcomes given above. Negative results in the *in vitro* assays should be viewed with caution in case a metabolite is responsible for the positive female PP assay. Unless the metabolic profile of the test substance is known, one option may be to conduct these assays with an added metabolising system. If the metabolic profile is known, then a higher level *in vivo* test may be advisable. The choice of tests will depend on the available *in vivo* effects data. Positive results in the female PP assay may also indicate the potential for endocrine mediated effects in lower vertebrates. These could be followed up with partial life cycle tests such as the

FSDT or MEOGRT; or the LAGDA if effects are on the thyroid hormone system. As in Scenarios A to C, *in vivo* assays/tests with negative results should be interpreted with caution as they may either indicate that the tests used do not have sufficient power to detect weak effects or, alternatively, that the effects do not present a concern for endocrine disruption.

1039. Scenarios G to I represent positive results in the female PP assay in the presence of various combinations of missing or equivocal data. Positive results in the female PP assay may also indicate the potential for endocrine mediated effects in lower vertebrates. These could be followed up with partial life cycle tests such as the FSDT or MEOGRT; or the LAGDA if effects are on the thyroid hormone system. Each positive female PP result scenario is divided into the three possible outcomes given above. The next step to take in these eventualities will depend on the nature of the other available data and the jurisdiction in which it is being used. In some cases, equivocal data may be viewed as positive whilst in others it may or may not contribute to the weight of evidence. The interpretation may also depend on the MOA in question and why the data are considered equivocal, e.g. a study that is equivocal for thyroid effects may still be of value in evaluating (anti)androgenic effects. In all three scenarios, the recommended first step is to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. Equivocal and missing data are alternative scenarios and two possibilities for the next step are given in most cases, but the nature of equivocal data means that decisions need to be taken on a case-by-case basis. In all cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

1040. Scenarios J to L represent negative results in the female PP assay in the presence of positive *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. As a negative result for the female PP is taken to be negative findings for both indicators of hormonal activity and apical endpoints (unlike the situation with positive outcomes), there is only one possible negative outcome. All three scenarios could also arise from a chemical that is positive in *in vitro* assays, but is metabolised to a non-active metabolite leading to negative results in the female PP assay. This should be considered first when investigating the next step. Endocrine active potency may also explain differences between *in vitro* and *in vivo* results (e.g. a weak chemical may give a positive result *in vitro* but may be negative *in vivo*). Positive *in vivo* effects data may involve other E,A,T,S, non-E,A,T,S mechanisms (e.g. involving other receptors or endocrine axes), more sensitive endpoints, greater statistical power or life stages that are more sensitive to the substance than the young adult animals in the female PP assay.

1041. Scenarios M to O represent negative results in the female PP assay in the presence of negative *in vitro* mechanistic data and positive, negative or equivocal *in vivo* effects data. Negative results for all tests (Scenario N) may be sufficient to enable a conclusion of no concern for endocrine disruption. This will depend on the weight of evidence and may not be possible. Where there are positive *in vivo* effects data, there could still be an E,A,T,S-related mechanism, the effects may be related to length of exposure, route of exposure or exposure at different life stages. Other E,A,T,S or non-E,A,T,S mechanisms may also be involved.

1042. Scenarios P to R represent negative results in the female PP assay in the presence of various combinations of missing or equivocal data. As with the positive result scenarios above (see Paragraph 1 039), the next step to take in these eventualities will have to be decided on a case-by-case basis. However, the recommended first step is generally to obtain reliable mechanistic (*in vitro*) data rather than proceed further with *in vivo* testing. In all

cases, the role of metabolism, route of exposure and data from structural analogues should be considered before deciding on the next step.

1043. In all scenarios (A to R), the next step to take to strengthen weight of evidence will depend on the existing information. Table C.3.17 is meant to provide a succinct guide and may not cover all circumstances or possibilities. The scenarios may also suggest that chemicals have simple or single MOA, when in practice they may have multiple endocrine and non-endocrine MOA. In some cases, for example, two opposite modes of simultaneous action (e.g. estrogenic and anti-estrogenic) could, depending on dose, lead to a minimisation or abolition of effects, while in others two different MOA (e.g. estrogenic and anti-androgenic) could potentially reinforce effects. Endocrine pathways interact, mixed effects are common and there are many pathways that cannot be distinguished with currently available TGs. If multiple MOA are suspected, either from the existing results or based on QSAR/read-across/integrated approaches, this should be investigated further if needed for regulatory decision making.

References

Goldman, J.M. et al. (2000), "Endocrine-disrupting chemicals: Prepubertal exposures and effects on sexual maturation and thyroid activity in the female rat. A focus on the EDSTAC recommendations", *Critical Reviews in Toxicology*, Vol. 30/2, pp. 135-196, https://doi.org/10.1080/10408440091159185.

Jacobs, M. et al. (2013), "*In vitro* metabolism and bioavailability tests for endocrine active substances: What is needed next for regulatory purposes?", *ALTEX – Alternatives to Animal Experimentation*, Vol. 30/3, pp. 331-351.

Jacobs, M.N. et al. (2008), "The use of metabolising systems for *in vitro* testing of endocrine disrupters", *Current Drug Metabolism*, Vol. 9/8, pp. 796-826.

Kucheryavenko, O. et al. (2018), "Report from the BfR Expert Hearing on Practicability of Hormonal Measurements", *Archives of Toxicology* (in prep.).

Odum, J. et al. (2004), "Cumulative dietary energy intake determines the onset of puberty in female rats", *Environmental Health Perspectives*, Vol. 112/15, pp. 1472-1480, http://dx.doi.org/10.1289/ehp.7039.

OECD (2008), *Detailed Review Paper on the Use of Metabolising Systems for* In Vitro *Testing of Endocrine Disrupters*, OECD Series on Testing and Assessment, No. 97, OECD Publishing, Paris, https://doi.org/10.1787/9789264085497-en.

OECD (2009), "Guidance document for histologic evaluation of endocrine and reproductive tests in rodents", OECD Series on Testing and Assessment, No. 106, OECD, Paris, www.oecd.org/env/ehs/testing/43411534.pdf.

US EPA (2009), "Endocrine Disrupter Screening Program Test Guidelines OPPTS 890.1450: Pubertal development and thyroid function in intact juvenile/peripubertal female rats", Environmental Protection Agency, Washington, DC.

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 Disrupter Screening Program Tier-1 battery", Environmental Protection Agency, Washington, DC.

WHO/IPCS (2002), "Global assessment of the state-of-the-science of endocrine disrupters", Damstra, T. et al. (eds.) WHO/PCS/EDC/02.2, World Health Organization, Geneva, www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en.

Table C.3.17. **Pubertal Development and Thyroid Function Assay in Peripubertal Female Rats (female PP assay) (US EPA OPPTS 890.1450): Guidance for scenarios of combinations of results with existing data**

This table represents possible conclusions to be drawn from assay data, and a next step which could be taken if further evidence is required about possible endocrine disrupting properties and/or effects. The guidance offered is not meant to be prescriptive, but provides science-based considerations. It encourages the use of all available data and expert judgement in a weight of evidence approach. Regional and national interpretation of results and "next steps" may vary.

The conclusions are grouped into a series of scenarios (A-R), each scenario representing a different combination of assay results, existing *in vitro* data and existing *in vivo* data. The symbol "+" indicates that the data in question represent a positive result, «-" indicates a negative result, and "Eq/0" indicates that the data are either equivocal or are not available.

Existing results: * "Mechanism (*in vitro* mechanistic data)" assumes that mechanistic data are available from estrogen receptor (ER-), androgen receptor (AR-) and steroidogenesis-based assays (Level 2). Thyroid hormone receptor (TR) and other assays concerning mechanisms of thyroid disruption may be available, but they are not in common use. In practice, data from all assays may not be available and therefore this must be taken into account when deciding on the "next step". Quantitative structure activity relationship (QSAR) predictions of estrogen and androgen binding/activation may be made for some substances.

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

*** *Note:* three possible outcomes for a positive result are given:

1. indicators of hormonal activity and apical endpoints positive
2. indicators of hormonal activity positive and apical endpoints negative
3. indicators of hormonal activity negative and apical endpoints positive.

"Apical endpoints" are age/body weight at vaginal opening, estrus cyclicity parameters, weights of ovaries, uterus, thyroid, pituitary and adrenals; histopathologic changes in ovaries and uterus.

"Indicators of hormonal activity" are hormones (T4 and TSH).

Scenarios	Result of female PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
A	+ ***	+	+	1) Increased evidence of (anti)-E,T,S activity (weak, moderate or strong). 2) Possible evidence of (anti)-E,T,S activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-E,T,S activity. 3) Moderate or strong (anti)-E,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-E,T,S activity.	Perform assay from Level 5 (e.g. Extended One-Generation Reproduction Toxiticy Study [EOGRTS] or two-generation) assay.	If existing data are from a Level 5 assay, there may be sufficient information to conclude evidence of concern for endocrine disruption (the EOGRTS provides the most information; however, for endocrine disrupting chemicals [EDCs] with a carcinogenic potential, OECD TG 451-3 may be more sensitive). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate estrogen/androgen/thyroid/steroidogenesis (E,A,T,S) modalities or other mechanisms. Possible effects on androgen modality should also be considered. Consider route of exposures for effects data and possible implications of absorption, distribution, metabolism and excretion (ADME) characteristics of the chemical. Endocrine activity possible in lower vertebrates. Consider performing a Fish Sexual Development Test (FSDT), Larval Amphibian Growth and Development Assay (LAGDA) or Medaka Extended One-Generation Reproduction Test (MEOGRT).
B	+	+	–	1) Increased evidence of (anti)-E,T,S activity (weak, moderate or strong). 2) Possible evidence of (anti)-E,T,S activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-E,T,S activity. 3) Moderate or strong (anti)-E,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-E,T,S activity.	Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	Question why there are differences from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay,n a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate E,T,S modalities or other mechanisms. Possible effects on androgen modality should also be considered. Endocrine activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT.

Scenarios	Result of female PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
C	+	+	Eq/0	1) Increased evidence of (anti)-E,T,S activity (weak, moderate or strong). 2) Possible evidence of (anti)-E,T,S activity, apical endpoints may be less sensitive or unaffected. Increased evidence of (anti)-E,T,S activity. 3) Moderate or strong (anti)-E,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-E,T,S activity.	Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	Check data on chemical analogues. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate E,T,S modalities or other mechanisms. Possible effects on androgen modality should also be considered. Consider route of exposure for female Peripubertal (PP) Assay and follow-up assay. Possible implications of ADME characteristics of the chemical. Endocrine activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT. Equivocal results may indicate chemical has multiple modes of action (MOA).
D	+	–	+	1) Increased evidence of (anti)-E,T,S activity (weak, moderate or strong). Acts via non-endocrine receptor (ER), thyroid hormone receptor (TR), steroidogenesis (S) mechanism or requires metabolic activation for activity. 2) Possible evidence of (anti)-E,T,S activity, apical endpoints may be less sensitive or unaffected. Acts via non-ER, TR, S mechanism or may require metabolic activation for activity. 3) Moderate or strong (anti)-E,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-ER, TR, S mechanism or may require metabolic activation for activity.	Perform *in vitro* ER, androgen receptor (AR), TR, S assays with added metabolising system.	If existing data are from an adequate Level 5 assay, there is sufficient information to conclude evidence of concern for endocrine disruption (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate E,T,S modalities or other mechanisms. Possible effects on androgen modality should also be considered. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical. Further mechanistic studies may help determine MOA. Endocrine activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT.

Scenarios	Result of female PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
E	+	–	–	1) Increased evidence of (anti)-E,T,S activity (weak, moderate or strong). Acts via non-ER, TR, S mechanism or may require metabolic activation for activity. Route of exposure may account for the differences from existing data. 2) Possible evidence of (anti)-E,T,S activity, apical endpoints may be less sensitive or unaffected. Acts via non-ER, TR, S mechanism or may require metabolic activation for activity. Route of exposure may account for the differences from existing data. 3) Possible evidence of (anti)-E,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-ER, TR, S mechanism or may require metabolic activation for activity. Route of exposure may account for the differences from existing data.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	Question why there are differences from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay, a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate other mechanisms. Endocrine activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT.
F	+	–	Eq/0	1) Increased evidence of (anti)-E,T,S activity (weak, moderate or strong). Acts via non-ER, TR, S mechanism or may require metabolic activation for activity. 2) Possible evidence of (anti)-E,T,S activity, apical endpoints may be less sensitive or unaffected. Acts via non-ER,TR, S mechanism or may require metabolic activation for activity. 3) Moderate (anti)-E,A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via non-ER, TR, S mechanism or may require metabolic activation for activity.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate other mechanisms. Check data on chemical analogues. Further mechanistic studies may help determine MOA. If existing data are from an adequate Level 5 assay, question there are why differences. If existing data are from a less sensitive assay, a higher level test may be required. Endocrine activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT. Equivocal results may indicate chemical has multiple MOA.

Scenarios	Result of female PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
G	+	Eq/0	+	1) Increased evidence of (anti)-E,T,S activity (weak, moderate or strong). May act via ER, TR, S mechanism (metabolic activation may be needed). 2) Possible evidence of (anti)-E,T,S activity, apical endpoints may be less sensitive or unaffected. May act via ER, TR, S mechanism (metabolic activation may be needed). 3) Moderate or strong (anti)-E,A,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Increased evidence of (anti)-E,T,S activity. May act via ER, TR, S mechanism (metabolic activation may be needed).	Perform *in vitro* ER, AR, TR, S assays (for the "0" scenario, otherwise Eq result available) OR Perform *in vitro* ER, AR, TR, S assays with added metabolising system.	If existing data are from a Level 5 assay, there may be sufficient information to conclude evidence of concern for endocrine disruption (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive). Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate E,T,S modalities or other mechanisms. Possible effects on A modality should also be considered. Check data on chemical analogues. Further mechanistic studies may help determine MOA. Consider route of exposures for effects data and possible implications of ADME characteristics of the chemical. Endocrine activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT. Equivocal results may indicate chemical has multiple MOA.
H	+	Eq/0	–	1) Increased evidence of (anti)-E,T,S activity (weak, moderate or strong). Acts via unknown mechanism or may require metabolic activation for activity. Route of exposure may account for the differences from existing data. 2) Possible evidence of (anti)-E,T,S activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism. Weak activity does not result in adverse effects. 3) Moderate (anti)-E,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via unknown mechanism. Weak activity does not result in adverse effects.	For the "0" scenario, perform *in vitro* ER, AR, TR, S assays with added metabolising system (otherwise Eq result available).	Question why there are differences from existing data. Consider route of exposures and possible implications of ADME characteristics of the chemical. If existing data are from a less sensitive assay, then a higher level test may be required. Effects on indicators of hormonal activity alone may be indicative of subtle changes not detected by apical endpoints. Effects on apical endpoints alone may indicate E,T,S modalities or other mechanisms. Possible effects on A modality should also be considered. Check data on chemical analogues. Further mechanistic studies may help determine MOA. Endocrine activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT. Equivocal results may indicate chemical has multiple MOA.

Scenarios	Result of female PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
I	+	Eq/0	Eq/0	1) Increased evidence of (anti)-E,T,S activity (weak, moderate or strong). Acts via unknown mechanism. Unknown potential for adverse effects. 2) Possible evidence of (anti)-E,T,S activity, apical endpoints may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. 3) Moderate or strong (anti)-E,T,S activity, indicators of hormonal activity may be less sensitive or unaffected. Acts via unknown mechanism. Unknown potential for adverse effects. There may be a need for metabolic activation.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	Effects on indicators of hormonal activity alone may be indicative of changes not detected by apical endpoints. Effects on apical endpoints alone may indicate E,T,S modalities or other mechanisms. Possible effects on androgen modality should also be considered. Check data on chemical analogues. Further mechanistic studies may help determine MOA. Endocrine activity possible in lower vertebrates. Consider performing an FSDT, LAGDA or MEOGRT. Equivocal results may indicate chemical has multiple MOA.
J	–	+	+	No evidence for E,T,S activity in female PP assay. Metabolism or potency may explains the difference from existing *in vitro* and *in vivo* data. Effects seen in existing studies are via non-E,T,S mechanism.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	If existing data are from an adequate Level 5 assay, question why there are differences. If data are from Uterotrophic Bioassays (UT) then this may be more sensitive than female PP assay. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical. Further mechanistic studies may help determine MOA.
K	–	+	–	No evidence for E,T,S activity in female PP assay. Metabolism or may potency explain *in vitro*/*in vivo* differences.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	If existing data are from an adequate Level 5 assay, there may be sufficient information to conclude absence of concern for endocrine disruption (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive). If existing data are from a less sensitive assay, then a higher level test may be required. If data are from UT assay, then need to conduct higher tier assay to conclude absence of concern for endocrine disruption. Further mechanistic studies with metabolism may help determine MOA.

Scenarios	Result of female PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
L	–	+	Eq/0	No evidence for E,T,S activity in female PP assay. Metabolism or potency may explain *in vitro*/*in vivo* differences. Unknown potential for adverse effects.	Perform *in vitro* ER, AR, TR, S assays with added metabolising system OR Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	Metabolic deactivation of chemical may occur *in vivo* so that potential *in vitro* activity is not realised. Consider possible routes of exposure implications of metabolism. Equivocal results may indicate chemical has multiple MOA.
M	–	–	+	No evidence for E,T,S activity in female PP assay. Effects seen in existing studies are via non-E,T,S mechanism.	Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	If existing data are from an adequate Level 5 assay, question why there are differences. If data are from UT assay, then this may be more sensitive than female PP assay. Effects seen in existing studies may be in a more sensitive life stage. Consider route of exposures and possible implications of ADME characteristics of the chemical.
N	–	–	–	No evidence for E,T,S activity in female PP assay. No evidence for (anti)-E,T,S activity *in vitro*. No evidence of adverse effects.	Possibly no need for further testing. If there is uncertainty, may perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	If existing data are from an adequate Level 5 assay, there may be sufficient information to conclude absence of concern for endocrine disruption (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive).
O	–	–	Eq/0	No evidence for E,T,S activity in female PP assay. No evidence for (anti)-E,T,S activity *in vitro*. Unknown potential for adverse effects.	Perform assay from Level 5 (e.g. EOGRTS or two-generation) assay.	Consider route of exposures and possible implications for ADME characteristics of the chemical in follow-up assay.
P	–	Eq/0	+	No evidence for E,T,S activity in female PP assay. Potential for adverse effects via unknown mechanism.	Perform *in vitro* ER, AR,TR, S assays.	Consider route of exposure and possible implications for differences from existing assay. If data are from Hershberger Bioassay (H), then this may be more sensitive than female PP assay. Effects seen in existing studies may be in a more sensitive life stage. Equivocal results may indicate chemical has multiple MOA.

Scenarios	Result of female PP assay	Existing results		Possible conclusions: 1) Indicators of hormonal activity and apical endpoints positive 2) Indicators of hormonal activity positive and apical endpoints negative 3) Indicators of hormonal activity negative and apical endpoints positive	Next step which could be taken to strengthen weight of evidence if necessary	Other considerations
		Mechanism (*in vitro* mechanistic data)*	Effects (*in vivo* effects of concern)**			
Q	–	Eq/0	–	No evidence for E,T,S activity in female PP assay. No evidence of adverse effects.	Perform *in vitro* ER, AR,TR, S assays.	If existing data are from an adequate Level 5 assay, there may be sufficient information to conclude absence of concern for endocrine disruption (the EOGRTS provides the most information; however, for EDCs with a carcinogenic potential, OECD TG 451-3 may be more sensitive). Further mechanistic studies may strengthen weight of evidence.
R	–	Eq/0	Eq/0	No evidence for E,T,S activity in female PP assay.	Perform *in vitro* ER, AR, TR, S assays, otherwise Eq result available.	Further mechanistic studies may strengthen weight of evidence. Check data on chemical analogues. Equivocal results may indicate chemical has multiple MOA.