Adopted:
7 September 2009

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Hershberger Bioassay in Rats: A Short-term Screening Assay for (Anti)Androgenic Properties

INTRODUCTION

1. The OECD initiated a high-priority activity in 1998 to revise existing guidelines and to develop new guidelines for the screening and testing of potential endocrine disrupters (1). One element of the activity was to develop a Test Guideline for the rat Hershberger Bioassay. After several decades of use by the pharmaceutical industry, this assay was first standardized by an official expert committee in 1962 as a screening tool for androgenic chemicals (2). In 2001-2007, the rat Hershberger Bioassay has undergone an extensive validation programme including the generation of a Background Review Document (23), compilation of a detailed methods paper (3), development of a dissection guide (21) and the conduct of extensive intra- and interlaboratory studies to show the reliability and reproducibility of the bioassay. These validation studies were conducted with a potent reference androgen (testosterone propionate (TP)), two potent synthetic androgens (trenbolone acetate and methyl testosterone), a potent antiandrogenic pharmaceutical (flutamide), a potent inhibitor of the synthesis (finasteride) of the natural androgen (dihydrotestosterone-DHT), several weakly antiandrogenic pesticides (linuron, vinclozolin, procymidone, p,p' DDE), a potent 5α reductase inhibitor (finasteride) and two known negative chemicals (dinitrophenol and nonylphenol) (4) (5) (6) (7) (8). This Test Guideline is the outcome of the long historical experience with the bioassay and the experience gained during the validation test programme and the results obtained therein.

2. The Hershberger Bioassay is a short-term *in vivo* screening test using accessory tissues of the male reproductive tract. The assay originated in the 1930's and was modified in the 1940's to include androgen-responsive muscles in the male reproductive tract (2) (9-15). In the 1960s, over 700 possible androgens were evaluated using a standardized version of the protocol (2) (14), and use of the assay for both androgens and antiandrogens was considered a standard method in the 1960s (2) (15). The current bioassay is based on the changes in weight of five androgen-dependent tissues in the castrate-peripubertal male rat. It evaluates the ability of a chemical to elicit biological activities consistent with androgen agonists, antagonists or 5α-reductase inhibitors. The five target androgen-dependent tissues included in this Test Guideline are the ventral prostate (VP), seminal vesicle (SV) (plus fluids and coagulating glands), levator ani-bulbocavernosus (LABC) muscle, paired Cowper's glands (COW) and the glans penis (GP). In the castrate-peripubertal male rat, these five tissues all respond to androgens with an increase in absolute weight. When these same tissues are stimulated to increase in weight by administration of a potent reference androgen, these five tissues all respond to antiandrogens with a decrease in absolute weight. The primary model for the Hershberger bioassay has been the surgically castrated peripubertal male, which was validated in Phases 1, 2 and 3 of the Hershberger validation program.

3. The Hershberger bioassay serves as a mechanistic *in vivo* screening assay for androgen agonists, androgen antagonists and 5α-reductase inhibitors and its application should be seen in the context of the "OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals"

© OECD, (2009)
You are free to use this material for personal, non-commercial purposes without seeking prior consent from the OECD, provided the source is duly mentioned. Any commercial use of this material is subject to written permission from the OECD.

(Annex 2). In this Conceptual Framework the Hershberger Bioassay is contained in Level 3 as an *in vivo* assay providing data about a single endocrine mechanism, i.e. (anti)androgenicity. It is intended to be included in a battery of *in vitro* and *in vivo* tests to identify substances with potential to interact with the endocrine system, ultimately leading to hazard and risk assessments for human health or the environment.

4. Due to animal welfare concerns with the castration procedure, the intact (uncastrated) stimulated weanling male was sought as an alternative model for the Hershberger Bioassay to avoid the castration step. The stimulated weanling test method was validated (24); however, in the validation studies, the weanling version of the Hershberger Bioassay did not appear to be able to consistently detect effects on androgen-dependent organ weights from weak anti-androgens at the doses tested. Therefore, it was not included in this Test Guideline. However, recognising that its use may provide not only animal welfare benefits but also may provide information on other modes of action, it is available in Guidance Document 115.

INITIAL CONSIDERATIONS AND LIMITATIONS

5. Androgen agonists and antagonists act as ligands for the androgen receptor and may activate or inhibit, respectively, gene transcription controlled by the receptor. In addition, some chemicals inhibit the conversion of testosterone to the more potent natural androgen dihydrotestosterone in some androgen target tissues (5α-reductase inhibitors). Such substances have the potential to lead to adverse health hazards, including reproductive and developmental effects. Therefore, the regulatory need exists to rapidly assess and evaluate a chemical as a possible androgen agonist or antagonist or 5α-reductase inhibitor. While informative, the affinity of a ligand for an androgen receptor as measured by receptor binding or transcriptional activation of reporter genes *in vitro* is not the only determinant of possible hazard. Other determinants include metabolic activation and deactivation upon entering the body, substance distribution to target tissues, and clearance from the body. This leads to the need to screen the possible activity of a chemical *in vivo* under relevant conditions and exposure. *In vivo* evaluation is less critical if the chemical's characteristics regarding Absorption – Distribution – Metabolism – Elimination (ADME) are known. Androgen-dependent tissues respond with rapid and vigorous growth to stimulation by androgens, particularly in castrate-peripubertal male rats. Rodent species, particularly the rat, are also widely used in toxicity studies for hazard characterization. Therefore, the assay version using the castrated peripubertal rat and the five target tissues in this assay are appropriate for the *in vivo* screening of androgen agonists and antagonists and 5α-reductase inhibitors.

6. This Test Guideline is based on those protocols employed in the OECD validation study which have been shown to be reliable and reproducible in intra- and interlaboratory studies (4) (5) (6) (7) (8). Both androgen and antiandrogen procedures are presented in this Guideline.

7. Although there was some variation in the dose of TP used to detect antiandrogens in the OECD Hershberger Bioassay validation programme by the different laboratories (0.2 versus 0.4 mg/kg/d, subcutaneous injection) there was little difference between these two protocol variations in the ability to detect weak or strong antiandrogenic activity. However, it is clear that the dose of TP should not be too high as to block the effects of weak androgen receptor (AR) antagonists or so low that the androgenic tissues display little growth response even without antiandrogen coadministration.

8. The growth response of the individual androgen-dependent tissues is not entirely of androgenic origin, i.e. compounds other than androgen agonists can alter the weight of certain tissues. However, the growth response of several tissues concomitantly substantiates a more androgen-specific mechanism. For example, high doses of potent estrogens can increase the weight of the seminal vesicles; however, the other androgen-dependent tissues in the assay do not respond in a similar manner. Antiandrogenic chemicals can act either as androgen receptor antagonists or 5α-reductase inhibitors. 5α-reductase inhibitors have a

© OCDE, (2009)

variable effect, because the conversion to more potent dihydrotestosterone varies by tissue. Antiandrogens that inhibit 5α–reductase, like finasteride, have more pronounced effects in the ventral prostate than other tissues as compared to a potent AR antagonist, like flutamide. This difference in tissue response can be used to differentiate between AR mediated and 5α–reductase mediated modes of action. In addition, the androgen receptor is evolutionarily related to that of other steroid hormones, and some other hormones, when administered at high, supraphysiological dosage levels, can bind and antagonize the growth-promoting effects of TP (13). Further, it also is plausible that enhanced steroid metabolism and a consequent lowering of serum testosterone could reduce androgen-dependent tissue growth. Therefore, any positive outcome in the Hershberger Bioassay should normally be evaluated using a weight of evidence approach, including *in vitro* assays, such as the AR and estrogen receptor (ER) binding assays and corresponding transcriptional activation assays, or from other *in vivo* assays that examine similar androgen target tissues such as the male pubertal assay, 15-day intact adult male assay, or 28-day or 90-day repeat dose studies.

9. Experience indicates that xenobiotic androgens are rarer than xenobiotic antiandrogens. The expectation then is that the Hershberger bioassay will be used most often for the screening of antiandrogens. However, the procedure to test for androgens could, nevertheless, be recommended for steroidal or steroid-like chemicals or for chemicals for which an indication of possible androgenic effects was derived from methods contained in Level 1 or 2 of the conceptual framework (Annex 2). Similarly, adverse effects associated with (anti)androgenic profiles may be observed in Level 5 assays, leading to the need to assess whether a substance operates by an endocrine mode of action.

10. It is acknowledged that all animal-based procedures should conform to local standards of animal care; the descriptions of care and treatment set forth below are minimal performance standards, and will be superseded by local regulations. Further guidance of the humane treatment of animals is given by the OECD (17).

11. As in any bioassay using experimental animals, careful considerations should be given to the necessity to carry out this study. Basically there may be two reasons for such a decision:

- high exposure potential (Level 1 of the Conceptual Framework) or indications for (anti)androgenicity in *in vitro* assays (Level 2) supporting investigations whether such effects may occur *in vivo;*
- effects consistent with (anti)androgenicity in Level 4 or 5 *in vivo* tests supporting investigations of the specific mode of action, e.g., to determine whether the effects were due to an (anti)androgenic mechanism.

12. Definitions used in this Test Guideline are given in Annex 1.

PRINCIPLE OF THE TEST

13. The Hershberger Bioassay achieves its sensitivity by using males with minimal endogenous androgen production. This is achieved through the use of castrated males provided an adequate time after castration for the target tissues to regress to a minimal and uniform baseline weight is allowed. Thus, when screening of potential androgenic activity, there are low endogenous levels of circulating androgens, the hypothalamic - pituitary - gonad axis is rendered unable to compensate via feedback mechanisms, the ability of the tissues to respond is maximized, and the starting tissue weight variability is minimized. When screening of potential anti-androgenic activity, a more consistent tissue weight gain can be achieved when the tissues are stimulated by a reference androgen. As a result, the Hershberger Bioassay requires only 6 animals per dose group whereas other assays with intact pubertal or adult males suggest using 15 males per dose group.

© OCDE, (2009)

14. Castration of peripubertal male rats should be done in an appropriate manner using approved anesthetics and aseptic technique. Analgesics should be administered on the first few days following surgery to eliminate post-surgical discomfort. Castration enhances the precision of the assay to detect weak androgens and antiandrogens by eliminating compensatory endocrine feed-back mechanisms present in the intact animal that can attenuate the effects of administered androgens and antiandrogens and by eliminating the large inter-individual variability in serum testosterone levels. Hence, castration reduces the numbers of animals required to screen for these endocrine activities.

15. When screening for potential androgenic activity, the test substance is administered daily by oral gavage or subcutaneous (sc) injection for a period of ten consecutive days. Test substances are administered to a minimum of two treatment groups of experimental animals using one dose level per group. The animals are necropsied approximately 24 hours after the last dose. A statistically significant increase in two or more target organ weights of the test substance groups compared to the vehicle control group indicates that the test substance is positive for potential androgenic activity (See paragraph 60). Androgens, like trenbolone that cannot be 5α–reduced have more pronounced effects on the LABC and GP versus TP, but all tissues should display increased growth.

16. When screening for potential antiandrogenic activity, the test substance is administered daily by oral gavage or subcutaneous injection for a period of ten consecutive days in concert with daily TP doses (0.2 or 0.4 mg/kg/d) by sc injection. It was determined in the validation programme that either 0.2 or 0.4 mg/kg/d of TP could be used as both were effective in the detection of antiandrogens and, therefore, only one dose should be selected for use in the assay. Graduated test substance doses are administered to a minimum of three treatment groups of experimental animals using one dose level per group. The animals are necropsied approximately 24 hours after the last dose. A statistically significant decrease in two or more target organ weights of the test substance plus TP groups compared to the TP only control group indicates that the test substance is positive for potential antiandrogenic activity (See paragraph 61).

DESCRIPTION OF THE METHOD

Selection of species and strain

17. The rat has been routinely used in the Hershberger Bioassay since the 1930s. Although it is biologically plausible that both the rat and mouse would display similar responses, based upon 70 years of experience with the rat model, the rat is the species of choice for the Hershberger Bioassay. In addition, since Hershberger Bioassay data may be preliminary to a long-term multigenerational study, this allows animals from the same species, strain and source to be used in both studies.

18. This protocol allows laboratories to select the strain of rat to be used in the assay which should generally be that used historically by the participating laboratory. Commonly used laboratory rat strains may be used; however, strains that mature significantly later than 42 days of age should not be used since castration of these males at 42 days of age could preclude measurement of glans penis weights, which can only be done after the prepuce is separated from the penile shaft. Thus, strains derived from the Fisher 344 rat should not be used, except in rare cases. The Fisher 344 rat has a different timing of sexual development compared with other more commonly used strains such as Sprague Dawley or Wistar strains (16). If such a strain is to be used, the laboratory should castrate them at a slightly older age and be able to demonstrate the sensitivity of the strain used. The rationale for the choice of rat strain should be clearly stated by the laboratory. Where the screening assay may be preliminary to a repeated dose oral study, a reproductive and developmental study, or a long-term study, preferably animals from the same strain and source should be used in all studies.

Housing and feeding conditions

19. All procedures should conform to all local standards of laboratory animal care. These descriptions of care and treatment are minimum standards and will be superseded by more stringent local regulations, when present. The temperature in the experimental animal room should be 22°C (with an approximate range ± 3°C). The relative humidity should be a minimum of 30% and preferably should not exceed a maximum 70%, other than during room cleaning. The aim should be relative humidity of 50-60%. Lighting should be artificial. The daily lighting sequence should be 12 hours light, 12 hours dark.

20. Group housing is preferable to isolation because of the young age of the animals and the fact that rats are social animals. Housing of two or three animals per cage avoids crowding and associated stress that may interfere with the hormonal control of the development of the sex accessory tissue. Cages should be thoroughly cleaned to remove possible contaminants and arranged in such a way that possible effects due to cage placement are minimized. Cages of a proper size (~2000 square centimeters) will prevent overcrowding.

21. Each animal should be identified individually (e.g., ear mark or tag) using a humane method. The method of identification should be recorded.

22. Laboratory diet and drinking water should be provided *ad libitum*. Laboratories executing the Hershberger Bioassay should use the laboratory diet normally used in their chemical testing work. In the validation studies of the Bioassay, no effects or variability were observed that were attributable to the diet. The diet used will be recorded and a sample of the laboratory diet should be retained for possible future analysis.

Performance Criteria for androgen-dependent organ weights

23. During the validation study, there was no evidence that a decrease in body weight affected increases or decreases in the growth of tissue weights for target tissues (i.e., that should be weighted in this study).

24. Among the different strains of rat used successfully in the validation programme, androgen-dependent organ weights are larger in the heavier rat strains than in the lighter strains. Therefore, the Hershberger Bioassay performance criteria do not include absolute expected organ weights for positive and negative controls.

25. Because the Coefficient of Variation (CV) for a tissue has an inverse relationship with statistical power, the Hershberger Bioassay performance criteria are based on maximum CV values for each tissue (Table 1). The CVs are derived from the OECD validation studies. In the case of negative outcomes, laboratories should examine the CVs from the control group and the high dose treatment group to determine if the maximum CV performance criteria have been exceeded.

26. The study should be repeated when: 1) three or more of the ten possible individual CVs in the control and high dose treatment groups exceed the maximums designated for agonist and antagonist studies in Table 1 and 2) at least two target tissues were marginally insignificant, i.e., ρ values between 0.05 and 0.10.

Table 1: Maximum allowable CV.s Determined for the Target Sex Accessory Tissues for the castrate model in the OECD Validation Studies[1].

[1] The threshold CV for a given tissue was identified from a graph of CV values - arranged from smallest sequentially to largest - for all means from all experiments in the validation exercise using a specific model (agonist

© OCDE, (2009)

Tissue	Antiandrogenic effects	Androgenic effects
Seminal vesicles	40%	40%
Ventral prostate	40%	45%
LABC	20%	30%
Cowper.s glands	35%	55%
Glans penis	17%	22%

PROCEDURE

Regulatory compliance and laboratory verification

27. Unlike the uterotrophic assay (TG 440), a demonstration of laboratory competence prior to the initiation of the study is not necessary for the Hershberger assay because concurrent positive (Testosterone Propionate and Flutamide) and negative controls are run as an integral part of the assay.

Number and condition of animals

28. Each treated and control group should include a minimum of 6 animals. This applies to both the androgenic and antiandrogenic protocols.

Castration

29. There should be an initial acclimatization period of several days after receipt of the animals to ensure that the animals are healthy and thriving. Since animals castrated before 42 days of age or postnatal day (pnd) 42 may not display preputial separation, animals should be castrated on pnd 42 or thereafter, not before. The animals are castrated under anesthesia by placing an incision in the scrotum and removing both testes and epididymides with ligation of blood vessels and seminal ducts. After confirming that no bleeding is occurring, the scrotum should be closed with suture or autoclips. Animals should be treated with analgesics for the first few days after surgery to alleviate any post-surgical discomfort. If castrated animals are purchased from an animal supplier, the age of animals and stage of sexual maturity should be assured by the supplier.

Acclimatization after castration

30. The animals should continue acclimation to the laboratory conditions to allow for the regression in the target tissue weights for a minimum of 7 days following castration. Animals should be observed daily, and any animals with evidence of disease or physical abnormalities should be removed. Thus, treatment with initiation of dosing (on study) may commence as early as pnd 49 days of age, but not later than pnd 60. Age at necropsy should not be greater than pnd 70. This flexibility allows a laboratory to schedule the experimental work efficiently.

---

or antagonist). The threshold CV was read from the point at which the increments between to the next highest CVs in the series are dramatically larger than the preceding few CV.s- the "breakpoint". It should be noted that although this analysis identified relatively reliable "breakpoints" for the antagonist model of the assay, CV curves for the agonist assay showed a more uniform increase making identification of a threshold CV by this method somewhat arbitrary.

Body weight and group randomization

31. Differences in individual body weights are a source of variability in tissue weights both within and among groups of animals. Increasing tissue weight variability results in an increased coefficient of variation (CV) and decreases the statistical power of the assay (sometimes referred to as assay sensitivity). Therefore, variations in body weight should be both experimentally and statistically controlled.

32. Experimental control involves producing small variations in body weight within and among the study groups. First, unusually small or large animals should be avoided and not placed in the study cohort. At study commencement the weight variation of animals used should not exceed ± 20 % of the mean weight (e.g. 175g ± 35g for castrated peripubertal rats). Second, animals should be assigned to groups (both control and treatment) by randomized weight distribution, so that mean body weight of each group is not statistically different from any other group. The block randomization procedure used should be recorded.

33. Because toxicity may decrease the body weight of treated groups relative to the control group, the body weight on the first day of test substance administration could be used as the statistical covariate, not the body weight at necropsy.

Dosage

34. In order to establish whether a test substance can have androgenic action *in vivo*, two dose groups of the test substance plus positive and vehicle (negative) controls (See paragraph 3) are normally sufficient, and this design is therefore preferred for animal welfare reasons. If the purpose is either to obtain a dose-response curve or to extrapolate to lower doses, at least 3 dose groups are needed. If information beyond identification of androgenic activity (such as an estimate of potency) is required, a different dosing regime should be considered. To test for antiandrogens, the test substance is administered together with a reference androgen agonist. A minimum of 3 test groups with different doses of the test chemical and a positive and a negative control (See paragraph 44) should be used. Except for treatment with the test substance, animals in the control group should be handled in an identical manner to the test group subjects. If a vehicle is used in administering the test substance, the control group should receive the vehicle in the highest volume used with the test groups.

35. All dose levels should be proposed and selected taking into account any existing toxicity and (toxico-) kinetic data available for the test substance or related materials. The highest dose level should first take into consideration the LD_{50} and/or acute toxicity information in order to avoid death, severe suffering or distress in the animals (17) (18) (19) (20) and, second, take into consideration available information on the doses used in subchronic and chronic studies. In general, the highest dose should not cause a reduction in the final body weight of the animals greater than 10% of control weight. The highest dose should be either 1) the highest dose that ensures animal survival and that is without significant toxicity or distress to the animals after 10 consecutive days of administration up to a maximal dose of 1000 mg/kg/day (See paragraph 36) or 2) a dose inducing (anti)androgenic effects, whichever is lower. As a screen, large intervals, e.g., one half log units (corresponding to a dose progression of 3.2) or even one log units, between dosages are acceptable. If there are no suitable data available, a range finding study (See paragraph 37) may be performed to aid the determination of the doses to be used.

Limit dose level

36. If a test at the limit dose of 1000 mg/kg body weight/day and a lower dose using the procedures described for this study, fails to produce a statistically significant change in reproductive organ weights,

 © OCDE, (2009)

then additional dose levels may be considered unnecessary. The limit dose applies except when human exposure data indicate the need for a higher dose level to be used.

Considerations for range finding

37. If necessary, a preliminary range finding study can be carried out with a few animals to select the appropriate dose groups [modified OECD guidelines for acute toxicity testing (TG 420, TG 423, TG 425)]. The objective in the case of the Hershberger Bioassay is to select doses that ensure animal survival and that are without significant toxicity or distress to the animals after ten consecutive days of chemical administration up to a limit dose of 1000 mg/kg/d as noted in paragraphs 35 and 36. In this respect an OECD Guidance Document (17) may be used defining clinical signs indicative of toxicity or distress to the animals. If feasible within this range finding study after ten days of administration, the target tissues may be excised and weighed approximately 24-hours after the last dose is administered. These data could then be used to assist the selection of the doses in the main study.

Reference substances and vehicle

38. The reference androgen agonist should be Testosterone Propionate (TP), CAS No 57-82-5. The reference TP dosage may be either 0.2 mg/kg-bw/d or 0.4 mg/kg-bw/d. The reference androgen antagonist should be Flutamide (FT), CAS No 1311-84-7. The reference FT dosage should be 3 mg/kg-bw/d, and the FT should be coadministered with the reference TP dosage.

39. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first. However, since many androgen ligands or their metabolic precursors tend to be hydrophobic, the most common approach is to use a solution/suspension in oil (e.g. corn, peanut, sesame or olive oil). Test substances can be dissolved in a minimal amount of 95% ethanol or other appropriate solvents and diluted to final working concentrations in the test vehicle. The toxic characteristics of the solvent should be known, and should be tested in a separate solvent-only control group. If the test substance is considered stable, gentle heating and vigorous mechanical action can be used to assist in dissolving the test substance. The stability of the test substance in the vehicle should be determined. If the test substance is stable for the duration of the study, then one starting aliquot of the test substance may be prepared, and the specified dosage dilutions prepared daily using care to avoid contamination and spoilage of the samples.

Administration of doses

40. TP should be administered by subcutaneous injection, and FT by oral gavage.

41. The test substance is administered by oral gavage or subcutaneous injection. Animal welfare considerations and the physical/chemical properties of the test substance need to be taken into account when choosing the route of administration. In addition, toxicological aspects like the relevance to the human route of exposure to the chemical (e.g. oral gavage to model ingestion, subcutaneous injection to model inhalation or dermal adsorption) and existing toxicological information and data on metabolism and kinetics (e.g. need to avoid first pass metabolism, better efficiency via a particular route) should be taken into account before extensive, long-term testing is initiated if positive results are obtained by injection.

42. The animals should be dosed in the same manner and time sequence for ten consecutive days at approximately 24 hour intervals. The dosage level should be adjusted daily based on the concurrent daily measures of body weight. The volume of dose and time that it is administered should be recorded on each day of exposure. Care should be taken in order not to exceed the maximum dose described in paragraph 35 to allow a meaningful interpretation of the data. Reduction of body weight, clinical signs, and other findings should be thoroughly assessed in this respect. For oral gavage, a stomach tube or a suitable

intubation cannula should be used. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. Local animal care guidelines should be followed, but the volume should not exceed 5 ml/kg body weight, except in the case of aqueous solutions where 10 ml/kg body weight may be used. For subcutaneous injections, doses should be administered to the dorsoscapular and or lumbar regions via sterile needle (e.g. 23- or 25-gauge) and a tuberculin syringe. Shaving the injection site is optional. Any losses, leakage at the injection site or incomplete dosing should be recorded. The total volume injected per rat per day should not exceed 0.5 ml/kg body weight.

Specific procedures for androgen agonists

43. For the test for androgen agonists, the vehicle is the negative control, and the TP-treated group is the positive control. Biological activity consistent with androgen agonists is tested by administering a test substance to treatment groups at the selected doses for 10 consecutive days. The weights of the five sex accessory tissues from the test substance groups are compared to the vehicle group for statistically significant increases in weight.

Specific procedures for androgen antagonists and 5α-reductase inhibitors

44. For the test for androgen antagonists and 5α-reductase inhibitors, the TP-treated group is the negative control, and the group coadministered with reference doses of TP and FT is the positive control. Biological activity consistent with androgen antagonists and 5α-reductase inhibitors is tested by administering a reference dose of TP and administering the test substance for 10 consecutive days. The weights of the five sex accessory tissues from the TP plus test substance groups are compared to the reference TP-only group for statistically significant decreases in weights.

OBSERVATIONS

Clinical observations

45. General clinical observations should be made at least once a day and more frequently when signs of toxicity are observed. Observations should be carried out preferably at the same time(s) each day and considering the period of anticipated peak effects after dosing. All animals should be observed for mortality, morbidity and general clinical signs such as changes in behaviour, skin, fur, eyes, mucous membranes, occurrence of secretions and excretions and autonomic activity (e.g. lacrimation, piloerection, pupil size, unusual respiratory pattern).

46. Any animal found dead should be removed and disposed of without further data analysis. Any mortality of animals prior to necropsy should be included in the study record together with any apparent reasons for mortality. Any moribund animals should be humanely terminated. Any moribund and subsequently euthanized animals should be included in the study record with apparent reasons for morbidity.

Body weight and food consumption

47. All animals should be weighed daily to the nearest 0.1 g, starting just prior to initiation of treatment i.e., when the animals are allocated into groups. As an optional measurement, the amount of food consumed during the treatment period may be measured per cage by weighing the feeders. The food consumption results should be expressed in grams per rat per day.

© OCDE, (2009)

Dissection and measurement of tissue and organ weights

48. Approximately 24 hours after the last administration of the test substance, the rats should be euthanized and exsanguinated according to the normal procedures of the conducting laboratory, and necropsy carried out. The method of humane killing should be recorded in the laboratory report.

49. Ideally, the necropsy order should be randomized across groups to avoid progression directly up or down dose groups that could affect the data. Any finding at necropsy, i.e., pathological changes/visible lesions should be noted and reported.

50. The five androgen-dependent tissues (VP, SV, LABC, COW, GP) should be weighted. These tissues should be excised, carefully trimmed of excess adhering tissue and fat, and their fresh (unfixed) weights determined. Each tissue should be handled with particular care to avoid the loss of fluids and to avoid desiccation, which may introduce significant errors and variability by decreasing the recorded weights. Several of the tissues may be very small or difficult to dissect, and this will introduce variability. Therefore, it is important that persons carrying out the dissection of the sex accessory tissues are familiar with standard dissection procedures for these tissues. A standard operating procedure (SOP) manual for dissection is available from the OECD (21). Careful training according to the SOP guide will minimize a potential source of variation in the study. Ideally the same prosector should be responsible for the dissection of a given tissue to eliminate inter-individual differences in tissue processing. If this is not possible, the necropsy should be designed such that each prosector dissects a given tissue from all treatment groups as opposed to one individual dissecting all tissues from a control group, while someone else is responsible for the treated groups. Each sex accessory tissues should be weighed without blotting to the nearest 0.1 mg, and the weights recorded for each animal.

51. Several of the tissues may be very small or difficult to dissect, and this will introduce variability. Previous work has indicated a range of coefficient of variations (CVs) that appears to differ based upon the proficiency of the laboratory. In a few cases, large differences in the absolute weights of the tissues such as the VP and COWS have been observed within a particular laboratory.

52. Liver, paired kidney, and paired adrenal weights are optional measurements. Again, tissues should be trimmed free of any adhering fascia and fat. The liver should be weighed and recorded to the nearest 0.1 g, and the paired kidneys and paired adrenals should be weighed and recorded to the nearest 0.1 mg. The liver, kidney and adrenals are not only influenced by androgens; they also provide useful indices of systemic toxicity.

53. Measurement of serum luteinising hormone (LH), follicular stimulating hormone (FSH) and testosterone (T) is optional. Serum T levels are useful to determine if the test substance induces liver metabolism of testosterone, lowering serum levels. Without the T data, such an effect might appear to be via an antiandrogenic mechanism. LH levels provides information about the ability of an antiandrogen to not only reduce organ weights, but also to affect hypothalamic-pituitary function, which in long term studies can induce testis tumors. FSH is an important hormone for spermatogenesis. Serum T4 and T3 also are optional measures that would provide useful supplemental information about the ability to disrupt thyroid hormone homeostasis. If hormone measurements are to be made, the rats should be anesthetized prior to necropsy and blood taken by cardiac puncture, and the method of anesthesia should be chosen with care so that it does not affect hormone measurement. The method of serum preparation, the source of radioimmunoassay or other measurement kits, the analytical procedures, and the results should be recorded. LH levels should be reported as ng per ml of serum, and T should also be reported as ng per ml of serum.

54. The dissection of the tissues is described as follows with a detailed dissection guide with photographs published as supplementary materials as part of the validation programme (21). A dissection video is also available from the Korea Food and Drug Administration web page (22).

- With the ventral surface of the animal upwards, determine if the prepuce of the penis has separated from the glans penis. If so, then retract the prepuce and remove the glans penis, weigh (nearest 0.1 mg), and record the weight;
- Open the abdominal skin and wall, exposing the viscera. If the optional organs are weighed, remove and weigh liver to nearest 0.1 g, remove the stomach and intestines, remove and weigh the paired kidneys and paired adrenals to the nearest 0.1 mg. This dissection exposes the bladder and begins the dissection of the target male accessory tissues.
- To dissect the VP, separate bladder from the ventral muscle layer by cutting connective tissue along the midline. Displace the bladder anteriorly towards the seminal vesicles (SV), revealing the left and right lobes of the ventral prostate (covered by a layer of fat). Carefully tease the fat from the right and left lobes of the VP. Gently displace the VP right lobe from the urethra and dissect the lobe from the urethra. While still holding the VP right lobe, gently displace the VP left lobe from the urethra and then dissect; weigh to nearest 0.1 mg and record the weight.
- To dissect the SVCG, displace the bladder caudally, exposing the vas deferens and right and left lobes of the seminal vesicles plus coagulating glands (SVCG). Prevent leakage of fluid by clamping a hemostat at the base of the SVCGs, where the vas deferens joins the urethra. Carefully dissect the SVCGs, with the hemostat in place trim fat and adnexa away, place in a tared weigh-boat, remove the hemostat, and weigh to the nearest 0.1 mg and record the weight.
- To dissect the levator ani plus bulbocavernosus muscles (LABC), the muscles and the base of the penis are exposed. The LA muscles wrap around the colon, while the anterior LA and BC muscles are attached to the penile bulbs. The skin and adnexa from the perianal region extending from the base of the penis to the anterior end of the anus are removed. The BC muscles are gradually dissected from the penile bulb and tissues. The colon is cut in two and, the full LABC can be dissected and removed. The LABC should be trimmed of fat and adnexa, weighed to the nearest 0.1 mg, and record the weight.
- After the LABC has been removed, the round Cowper's or bulbourethral glands (COW) are visible at the base of, and slightly dorsal to, the penile bulbs. Careful dissection is required to avoid nicking the thin capsule in order to prevent fluid leakage. Weigh the paired COW to the nearest 0.1 mg, and record the weight.
- In addition, if fluid is lost from any gland during the necropsy and dissection, this should be recorded.

55. If the evaluation of each chemical requires necropsy of more animals than is reasonable for a single day, the study start may be staggered on two consecutive days, resulting in the staggering of the necropsy and the related work over two days. If staggered in this manner,_one-half of the animals per treatment group should be used per day.

56. Carcasses should be disposed of in an appropriate manner following necropsy.

© OCDE, (2009)

REPORTING

Data

57. Data should be reported individually (i.e. body weight, accessory sex tissue weights, optional measurements and other responses and observations) and for each group of animals (means and standard deviations of all measurement taken). The data should be summarized in tabular form. The data should show the number of animals at the start of the test, the number of animals found dead during the test or found showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration and severity.

58. A final report should include:

Testing facility

- Name of facility, location
- Study director and other personnel and their study responsibilities
- Dates the study began and ended, i.e., first day of test substance administration and last day of necropsy, respectively.

Test substance

- Source, lot/batch number, identity, purity, full address of the supplier and characterization of the test substance(s)
- Physical nature and, where relevant, physicochemical properties;
- Storage conditions and the method and frequency of dilution preparation
- Any data generated on stability
- Any analyses of dosing solutions/suspensions.

Vehicle

- Characterization of the vehicle (identity, supplier and lot #)
- Justification of the vehicle choice (if other than water)

Test animals and animal husbandry procedures

- Species/strain used and rationale for choice
- Source or supplier of animals, including full address
- Number and age of animals supplied
- Housing conditions (temperature, lighting, and so on)
- Diet (name, type, supplier, lot number, content and if known, phytoestrogens levels)
- Bedding (name, type, supplier, content)
- Caging conditions and number of animals per cage;

Assay Conditions

- Age at castration and duration of acclimatization after castration;
- Individual weights of animals at the start of the study (to nearest 0.1 g);
- Randomization process and a record of the assignment to vehicle, reference, test substance groups, and cages

© OCDE, (2009)

- Mean and standard deviation of the body weights for each group for each weigh day throughout the study;
- Rationale for dose selection
- Route of administration of test substance and rationale for the choice of exposure route
- If an assay for antiandrogenicity, the TP treatment (dose and volume),
- Test substance treatment (dose and volume),
- Time of dosing
- Necropsy procedures, including means of exsanguinations and any anesthesia
- If serum analyses are performed, details of the method should be supplied. For example, if RIA is used, the RIA procedure, source of RIA kits, kit expiration dates, procedure for scintillation counting, and standardization should be reported.

Results

- Daily observations for each animal during dosing, including:
    - Body weights (to the nearest 0.1 g),
    - Clinical signs (if any),
    - Any measurement or notes of food consumption.
- Necropsy observations for each animal, including:
    - Date of necropsy,
    - Animal treatment group,
    - Animal ID,
    - Prosector,
    - Time of day necropsy and dissection are performed,
    - Animal age,
    - Final body weight at necropsy, noting any statistically significant increase or decrease,
    - Order of animal exsanguination and dissection at necropsy,
    - Weights of the five target androgen dependent tissues:
        - Ventral prostate (to the nearest 0.1 mg)
        - Seminal vesicles plus coagulating glands, including fluid (paired, to nearest 0.1 mg)
        - Levator ani plus bulbocavernosus muscle complex (to nearest 0.1 mg)
        - Cowper's glands (fresh weight – paired, to nearest 0.1 mg).
        - Glans penis (fresh weight to nearest 0.1 mg),
    - Weights of optional tissues, if performed:
        - Liver (to nearest 0.1 g)
        - Kidney (paired, to nearest 0.1 mg)
        - Adrenal (paired, to nearest 0.1 mg)
    - General remarks and comments
- Analyses of serum hormones, if performed.
    - Serum LH (optional – ng per ml of serum), and
    - Serum T (optional – ng per ml of serum)
- General remarks and comments

Data summarization

Data should be summarized in tabular form containing the sample size for each group, the mean of the value, and the standard error of the mean or the standard deviation. Tables should include necropsy body weights, body weight changes from the beginning of dosing until necropsy, target accessory sex tissues weights, and any optional organ weights.

© OCDE, (2009)

Discussion of the results

Analysis of results

59. Necropsy body and organ weights should be statistically analyzed for characteristics such as homogeneity of variance with appropriate data transformations as needed. Treatment groups should be compared to a control group using techniques such as ANOVA followed by pairwise comparisons (e.g. Dunnett's one tailed test) and the criterion for statistical difference, for example, $p \leq 0.05$. Those groups attaining statistical significance should be identified. However, "relative organ" weights should be avoided due to the invalid statistical assumptions underlying this data manipulation.

60. For androgen agonism, the control should be the vehicle-only test group. The mode of action characteristics of a test substance can lead to different relative responses amongst the tissues, for example trenbolone, which cannot be 5 alpha-reduced, has more pronounced effects on the LABC and GP than does TP. A statistically significant increase ($p \leq 0.05$) in any two or more of the five target androgen-dependent tissue weights (VP, LABC, GP, CG and SVCG) should be considered a positive androgen agonist result, and all the target tissues should display some degree of increased growth combined evaluation of all accessory sex organs (ASO) tissue responses could be achieved using appropriate multivariate data analysis. This could improve the analysis, especially in cases where only a single tissue gives a statistically significant response.

61. For androgen antagonism, the control should be the reference androgen (testosterone propionate only) test group. The mode of action characteristics of a test substance can lead to different relative responses amongst the tissues, for example 5 alpha α-reductase inhibitors, like finasteride, have more pronounced effects on the ventral prostate than other tissues as compared to potent AR antagonists, like flutamide. A statistically significant reduction ($p \leq 0.05$) in any two or more of the five target androgen-dependent tissue weights (VP, LABC, GP, CG and SVCG) relative to TP treatment alone should be considered a positive androgen antagonist result and all the target tissues should display some degree of reduced growth. Combined evaluation of all ASO tissue responses could be achieved using appropriate multivariate data analysis. This could improve the analysis, especially in cases where only a single tissue gives a statistically significant response.

62. Data should be summarized in tabular form containing the mean, standard error of the mean (standard deviation would also be acceptable) and sample size for each group. Individual data tables should also be included. The individual values, mean, SE (SD) and CV values for the control data should be examined to determine if they meet acceptable criteria for consistency with expected historical values. CVs that exceed CV values listed in Table 1 (see paragraphs 25 and 26) for each organ weight should determine if there are errors in data recording or entry or if the laboratory has not yet mastered accurate dissection of the androgen-dependent tissues and further training/practice is warranted. Generally, CVs (the standard deviation divided by the mean organ weight) are reproducible from lab to lab and study to study. Data presented should include at least; ventral prostate, seminal vesicle, levator ani plus bulbocavernosus, Cowper's glands, glans penis, liver, and body weights and body weight change from the beginning of dosing until necropsy. Data also may be presented after covariance adjustment for body weight, but this should not replace presentation of the unadjusted data. In addition, if preputial separation (PPS) does not occur in any of the groups, the incidence of PPS should be recorded and statistically compared to the control group using Fisher Exact test.

63. When verifying the computer data entries with the original data sheets for accuracy, organ weight values that are not biologically plausible or vary by more than three standard deviations from that treatment group means should be carefully scrutinized and may need to be discarded, likely being recording errors.

64. Comparison of study results with OECD CV values (in Table 1) is often an important step in interpretation as to the validity of the study results. Historical data for vehicle control groups should be maintained in the laboratory. Historical data for responses to positive reference substances, such as TP and FT, should also be maintained in the laboratory. Laboratories may also periodically test the response to known weak androgen agonists and antagonists and maintain these data. These data can be compared to available data OECD data to ensure that the laboratory's methods yield sufficient statistical precision and power.

© OCDE, (2009)

ANNEX 1

DEFINITIONS

Androgenic is a term used to describe a positive influence on the growth of androgen-dependent tissues

Antiandrogenic is the capability of a chemical to suppress the action of TP in a mammalian organism.

Date of birth is postnatal day 0.

Dose is the amount of test substance administered. For the Hershberger Bioassay, the dose is expressed as weight of test substance per unit body weight of test animal per day (e.g. mg/kg body weight/day).

Dosage is a general term comprising of dose, its frequency and the duration of dosing.
.

Moribund is a term used to describe an animal in a dying state, i.e., near the point of death.

Postnatal day X is the Xth day of life after the day of birth.

Sensitivity is the capability of a test method to correctly identify chemicals having the property that is being tested for.

Specificity is the capability of a test method to correctly identify chemicals not having the property that is being tested for.

Validation is a scientific process designed to characterize the operational requirements and limitations of a test method and to demonstrate its reliability and relevance for a particular purpose.

© OCDE, (2009)

ANNEX 2

Note: Document prepared by the Secretariat of the Test Guidelines Programme based on the agreement reached at the 6th Meeting of the EDTA Task Force

OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals

Level 1 Sorting & prioritization based upon existing information	- physical & chemical properties, e.g., MW, reactivity, volatility, biodegradability, - human & environmental exposure, e.g., production volume, release, use patterns - hazard, e.g., available toxicological data	
Level 2 *In vitro* assays providing mechanistic data	- ER, AR, TR receptor binding affinity - Transcriptional activation - Aromatase and steroidogenesis *in vitro* - Aryl hydrocarbon receptor recognition/binding - QSARs	-High Through Put Prescreens - Thyroid function - Fish hepatocyte VTG assay - Others (as appropriate)
Level 3 *In vivo* assays providing data about single endocrine Mechanisms and effects	- Uterotrophic assay (estrogenic related) - Hershberger assay (androgenic related) - Non -receptor mediated hormone function - Others (e.g. thyroid)	- Fish VTG (vitellogenin) assay (estrogenic related)
Level 4 *In vivo* assays providing data about multiple endocrine Mechanisms and effects	- enhanced OECD 407 (endpoints based on endocrine mechanisms) - male and female pubertal assays - adult intact male assay	- Fish gonadal histopathology assay - Frog metamorphosis assay
Level 5 *In vivo* assays providing data on effects from endocrine & other mechanisms	- 1-generation assay (TG415 enhanced)[1] - 2-generation assay (TG416 enhanced)[1] - reproductive screening test (TG421 enhanced)[1] - combined 28 day/reproduction screening test (TG 422 enhanced)[1] 1 Potential enhancements will be considered by VMG mamm	- Partial and full life cycle assays in fish, birds, amphibians & invertebrates (developmental and reproduction)

VMG mamm: Validation Management Group on Mammalian Testing and Assessment

© OCDE, (2009)

Notes to the Framework

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

Note 2: In level 5,ecotoxicology should include endpoints that indicate mechanisms of adverse effects, and potential population damage

Note 3: When a multimodal model covers several of the single endpoint assays, that model would replace the use of those single endpoint assays

Note 4: The assessment of each chemical should be based on a case by case basis, taking into account all available information, bearing in mind the function of the framework levels.

Note 5: The framework should not be considered as all inclusive at the present time. At levels 3,4 and 5 it includes assays that are either available or for which validation is under way. With respect to the latter, these are provisionally included. Once developed and validated, they will be formally added to the framework.

Note 6: Level 5 should not be considered as including definitive tests only. Tests included at that level are considered to contribute to general hazard and risk assessment.

© OCDE, (2009)

LITERATURE

(1) OECD. 1998. Report of the First Meeting of the OECD Endocrine Disrupter Testing and Assessment (EDTA) Task Force, 10th-11th March 1998, ENV/MC/CHEM/RA(98)5.

(2) Dorfman RI. 1962. Standard methods adopted by official organization. Academic Press, NY.

(3) Gray LE Jr, Furr J and Ostby JS. 2005. Hershberger assay to investigate the effects of endocrine disrupting compounds with androgenic and antiandrogenic activity in castrate-immature male rats. In: *Current Protocols in Toxicology* 16.9.1-16.9.15. J Wiley and Sons Inc.

(4) OECD. 2006. Final OECD report of the initial work towards the validation of the rat Hershberger assay. Phase 1. Androgenic response to testosterone propionate and anti-androgenic effects of flutamide. Environmental Health and Safety, Monograph Series on Testing and Assessment N° 62. ENV/JM/MONO(2006)30.

(5) OECD. 2008. Report of the OECD Validation of the Rat Hershberger Bioassay: Phase 2: Testing of Androgen Agonists, Androgen Antagonists and a 5α-Reductase Inhibitor in Dose Response Studies by Multiple Laboratories. Environmental Health and Safety, Monograph Series on Testing and Assessment N° 86. ENV/JM/MONO(2008)3.

(6) OECD. 2007. Report of the Validation of the Rat Hershberger Assay: Phase 3: Coded Testing of Androgen Agonists, Androgen Antagonists and Negative Reference Chemicals by Multiple Laboratories. Surgical Castrate Model Protocol. Environmental Health and Safety, Monograph Series on Testing and Assessment N° 73. ENV/JM/MONO(2007)20.

(7) Owens, W, Zeiger E, Walker M, Ashby J, Onyon L, Gray, Jr, LE. 2006. The OECD programme to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses. Phase 1: Use of a potent agonist and a potent antagonist to test the standardized protocol. Env. Health Persp. 114:1265-1269.

(8) Owens W, Gray LE, Zeiger E, Walker M, Yamasaki K, Ashby J, Jacob E. 2007. The OECD program to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses: phase 2 dose-response studies. Environ Health Perspect. 115(5):671-8.

(9) Korenchevsky V. 1932. The assay of testicular hormone preparations. Biochem J 26:413-422.

(10) Korenchevsky V, Dennison M, Schalit R. 1932. The response of castrated male rats to the injection of the testicular hormone. Biochem J 26:1306-1314.

(11) Eisenberg E, Gordan GS. 1950. The levator ani muscle of the rat as an index of myotrophic activity of steroidal hormones. J Pharmacol Exp Therap 99:38-44.

(12) Eisenberg E, Gordan GS, Elliott HW. 1949. Testosterone and tissue respiration of the castrate male rat with a possible test for mytrophic activity. Endocrinology 45:113-119.

(13) Hershberger L, Shipley E, Meyer R. 1953. Myotrophic activity of 19-nortestosterone and other steroids determined by modified levator ani muscle method. Proc Soc Exp Biol Med 83:175-180.

(14) Hilgar AG, Vollmer EP. 1964. Endocrine bioassay data: Androgenic and myogenic. Washington DC: United

States Public Health Service.

(15) Dorfman RI. 1969. Androgens and anabolic agents. In: Methods in Hormone Research, volume IIA. (Dorfman RI, ed.) New York:Academic Press, 151-220.

(16) Massaro EJ. 2002. Handbook of Neurotoxicology, volume I. New York: Humana Press, p 38.

(17) OECD. 2000. Guidance document on the recognition, assessment and use of clinical signs as humane endpoints for experimental animals used in safety evaluation. Environmental Health and Safety Monograph Series on Testing and Assessment No 19. ENV/JM/MONO(2000)7.

(18) OECD. 1982. Organization for Economic Co-operation and Development - Principles of Good Laboratory Practice, ISBN 92-64-12367-9, Paris.

(19) OECD. 2001. Acute oral toxicity – up-and-down procedure. OECD Guideline for the testing of chemicals 425.

(20) OECD. 2001. Guidance document on acute oral toxicity . Environmental Health and Safety Monograph Series on Testing and Assessment No 24. ENV/JM/MONO(2001)4.

(21) Supplemental materials for Owens et al. 2006. The OECD programme to validate the rat Hershberger bioassay to screen compounds for in vivo androgen and antiandrogen responses. Phase 1: Use of a potent agonist and a potent antagonist to test the standardized protocol. Env. Health Persp. 114:1265-1269. See, section II, The dissection guidance provided to the laboratories: http://www.ehponline.org/docs/2006/8751/suppl.pdf.

(22) Korea Food and Drug Administration. Visual reference guide on Hershberger assay procedure, including a dissection video. http://rndmoa.kfda.go.kr/endocrine/reference/education_fr.html .

(23) OECD 2008. Background Review Document on the Rodent Hershberger Bioassay. Environmental Health and Safety Monograph Series on Testing and Assessment No90. ENV/JM/MONO(2008)17.

(24) OECD 2008. Draft Validation report of the Intact, Stimulated, Weanling Male Rat Version of the Hershberger Bioassay.