# DRAFT OECD GUIDELINE FOR THE TESTING OF CHEMICALS

# The Hershberger Bioassay in Rats

## 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 validated by an official expert committee in 1962 as a screening tool for androgenic chemicals (2). Over the last five years, the rat Hershberger Bioassay has undergone an extensive validation program including the compilation of a detailed methods paper (3), dissection guide and the conduct of extensive intra- and interlaboratory studies to show the reliability and reproducibility of the bioassay 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 a natural androgen (dihydrotestosterone-DHT) synthesis (finasteride), several weakly antiandrogenic pesticides (linuron, vinclozolin, procymidone, p,p' DDT), a potent 5 alpha reductase inhibitor (finasteride) and two known negative chemicals (dinitrophenol and nonylphenol (4)(5)(6)(7)(8). This Test Guideline XXX is the outcome of the long historical experience with the bioassay and the experience gained during the validation test programme and the results obtained thereby.

2. The Hershberger Bioassay is a short-term screening test that originated in the 1930's using accessory tissues of the male reproductive tract and modified in the 1940's to include androgen-responsive muscles in the male reproductive tract (2)(9-16). In the 1960s, a standardized protocol was used to evaluate over 700 possible androgens (2)(14), and the assay's use for both androgens and antiandrogens was considered a standard method in the 1960s (2)(15)(16). 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 or antagonists. The five androgen-dependent tissue included in the Test Guideline XXX 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.

3. This bioassay serves as an *in vivo* screening assay for androgens and antiandrogens and its application should be seen in the context of the "OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals". 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.

4. The Hershberger Bioassay is an *in vivo* screening assay embedded within a battery of tests ultimately leading to hazard and risk assessments for human health or the environment. As for all screening assays it should not fail to identify chemicals with intrinsic (anti)androgenic

activity (no false negatives), while a low rate of false positives may be allowed for. The above mentioned OECD validation program demonstrated sensitivity of the test procedure for antagonists and agonists as well as good intra- and interlaboratory reproducibility, and demonstrated a low rate of false positives with two negative compounds.

## 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 or transcriptional activation of reporter genes *in vitro* are not the only determinants 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, unless 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 rat and these five target tissues in this assay are appropriate for the *in vivo* screening of androgen agonists and antagonists and 5α-reductase inhibitors.

6. This Guideline is based on those protocols employed in the OECD validation study which have been shown to be reliable and repeatable in intra- and interlaboratory studies (4)(5)(6)(7)(8).

7. Although there was some variation in the dose of TP used to detect antiandrogens in the OECD Hershberger Bioassay validation program by the different laboratories (0.2 versus 0.4 mg/kg/d sc) 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 androgen origin, i.e. compounds other than androgen agonists can alter the weight of certain tissues. The growth response of several tissues is then more specific. 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 inhibitor. 5α-reductase inhibitors have a variable affect, because the conversion to more potent dihydrotestosterone varies by tissue. For example, the VP is relatively dependent on 5α-reductase activity, and the LABC is much less so. In addition, the androgen receptor is evolutionarily related to other steroid hormones, and some other hormones, when administered at high, supraphysiological dosage levels, can bind and antagonize the growth-promoting effects of

TP. 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 ER binding assays and corresponding transcriptional activation assays, or from other *in vivo* assays such as similar androgen target tissues in the male pubertal assay and 91-day repeat dose studies.

9. Experience indicates that xenobiotic androgens are rare and that xenobiotic antiandrogens are more common. The expectation then is that the Hershberger bioassay will be used most often for the screening of androgen antagonists. 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 in-development methods contained in level 1 or 2 of the conceptual framework. 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 will 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. Definitions used in this Test Guideline are given in ANNEX 1.

## PRINCIPLE OF THE TEST

12. The Hershberger Bioassay achieves its sensitivity by using castrated males and providing adequate time after castration for the target tissues to regress to a minimal and uniform baseline weight. Thus, for androgens, there are low endogenous levels of circulating androgens, the hypothalamic-pituitary- gonad axis is unable to compensate via feedback mechanisms, the ability of the tissues to respond is maximized, and the starting tissue weight variability is minimized. For antiandrogens, 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.

13. Castration of peripubertal male rats must be done in a humane manner using appropriate 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 is humane and it reduces the numbers of animals required to screen for these endocrine activities.

14. For androgens, the test substance is administered daily by oral gavage for a period of ten consecutive days. Graduated test substance doses 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 dose responsive, statistically significant

increase in the 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. Androgens, like trenbolone that cannot be 5 alpha reduced have more pronounced effects on the LABC and GP versus TP, but all tissues should display increased growth.

15. For antiandrogens, the test substance is administered daily by oral gavage for a period of ten consecutive days in concert with daily TP doses (0.2 or 0.4 mg/kg/d) by sc injection. Graduated test substance doses 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 dose responsive, statistically significant decrease in the 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. Antiandrogens that inhibit 5 alpha reductase, like finasteride, have more pronounced effects in the ventral prostate than other tissues as compared to a potent AR antagonist, like flutamide.

16. The specifics for both androgen and antiandrogen procedure are presented herein.

17. 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) and 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., were the effects due to an (anti)androgenic mechanism.

## DESCRIPTION OF THE METHOD

### Selection of species and strain

18. 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. Commonly used laboratory strains of healthy animals should be employed.

19. 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. If such a strain were used, then the laboratory should castrate them at a slightly older age and be able to demonstrate the sensitivity of the strain used.

### Housing and feeding conditions

20. 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 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.

21. 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 will be retained for possible future analysis.

22. Due to the young age of the immature animals and the fact that rats are social animals, group housing of 2-3 rats per cage is preferably to isolation. Three animals or less 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.

23. Each animal will be identified individually (e.g., ear mark or tag). The method of identification will be recorded.

24. Although high levels of phytoestrogens in laboratory diets have not been shown to affect the endpoints in the Hershberger Bioassay, as a precaution the dietary recommendations should comply with those described for the Uterotrophic Bioassay. The presence of phytoestrogens results primarily from the inclusion of soy and alfalfa products in the laboratory diets. Similarly, since there are many dietary (metabolizable energy, fat content, vitamin supplementation, potential contaminants and etc) and other environmental (cage material, bedding, etc) factors that also could potentially impact on the outcome of the Hershberger Bioassay, these factors need to controlled and considered as a source of variation if situations arise in which positive or negative controls fail to respond normally. Here, as in the Uterotrophic Bioassay, limited quantities of dietary phytoestrogens are acceptable and do not reduce the sensitivity of the bioassay. As a guide, dietary levels of phytoestrogens should not exceed 350 µg of genistein equivalents/gram of laboratory diet for immature male rats.

25. When chemicals are administered at dose levels that produce changes in body weight below the MTD (maximum reduction in necropsy weight of 10%), body weight is not an important variable in the Hershberger Bioassay. It is evident that among the different strains of rats used successfully in the validation program that androgen-dependent organ weight are larger in the heavier rat strains than in the lighter strains. For this reason, the Hershberger Bioassay performance criteria do not include absolute expected organ weights for positive and negative controls. Although these organ weights covary among strains with body weights, within a strain in a specific study the variation should be relatively low and the correlation with body weight is negligible. Hence, the Hershberger Bioassay performance criteria include CV values for each organ but not absolute organ weights. Laboratories should examine the data from each treatment group to determine if these performance criteria are met. In cases where the CVs of the control groups exceed the expected historical CV values by more than two-fold the ability to detect treatment effects is seriously compromised and the study may need to be repeated after the source of variability is identified. In some cases, high CV values result from errors in data

recording. For this reason, individual data should be examined for biologically implausible results whereas, in some other cases, additional training in conduct of necropsy may be required.

## PROCEDURE

### Regulatory compliance and laboratory verification

26. The assay should be conducted in conformity with the OECD Good Laboratory Practice and Quality Assurance Procedures (18).

27. On a regular basis or prior to the study, the responsiveness of the test system (animal model) should have been verified in the laboratory using appropriate dose(s) of a reference androgen: testosterone propionate (CAS No. 57-82-5) (TP) to examine whether a statistically significant increase in the five male accessory tissues as compared to the vehicle control group is achieved. If this is not the case, then the experimental design should be modified.

### Strain

28. This protocol allows laboratories to select the strain of rat to be used in the validation of the assay. The selection should be the strain used historically by the participating laboratory, but should not include strains like the Fisher 344 rat. The Fisher 344 rat has a different timing of sexual development compared to other more commonly used strains such as Sprague Dawley or Wistar strains. 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

### Number and condition of animals

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

### Castration

30. 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 can 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

31. 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 will be observed daily, and any animals with evidence of disease or physical abnormalities will

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.

**Body weight and group randomization**

32. Differences in individual body weights are a source of variability in tissues 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.

33. 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. A study commencement the weight variation of animals used should not exceed ± 20 % of the mean weight (e.g. 175g ± 35g). 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.

34. 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**

35. Generally, a minimum of two test groups and a control group 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.

36. All dose levels should be proposed and selected taking into account any existing toxicity and (toxico-) kinetic data available for the test compound 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) and second take into consideration available information on the maximum tolerated dose in subchronic and chronic studies. In general, the MTD should not cause a reduction in the final body weight of the animals greater than 10% of control weight. The highest dose should represent the limit dose, a maximum tolerated dose (MTD) or a dose inducing (anti)androgenic effects. Thereafter, one or more reduced dose levels should be selected with a view to demonstrating any dosage related response and identifying a no-observed-effects-level (NOEL). 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 may be performed to aid the determination of the doses to be used.

**Limit test**

37. If a test at one dose level of at least 1000 mg/kg body weight/day using the procedures described for this study, fails to produce a statistically significant change in reproductive organ weights, then additional dose levels may be considered unnecessary. The limit test applies except when human exposure data indicate the need for a higher dose level to be used.

**Considerations for range finding**

38. If necessary, a preliminary range finding study can be carried out with few animals [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. 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 mandatory 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 main study design (select an acceptable maximum and lower doses and recommend the number of dose groups).

**Laboratory health and safety requirements**

39. The test substances should be treated as possible reproductive and developmental toxicants with appropriate precautions instituted to protect laboratory personnel, e.g. necessary training, labeling and storage procedures, protective handling procedures during dose preparation and dose administration, and appropriate protective equipment.

**Reference substances and vehicle**

40. The reference androgen agonist will 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 could 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.

41. It is recommended that, wherever possible, the use of an aqueous solution/suspension be considered first. However, 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 must 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**

42. TP will be administered by subcutaneous injection, and FT by oral gavage.

43. The test compound is administered by oral gavage or subcutaneous injection. Animal welfare considerations and the physical/chemical properties of the test material 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.

44. The animals will be dosed in the same manner and time sequence for ten consecutive

days at approximately 24 hour intervals. The dosage level will be adjusted daily based on the concurrent daily measures of body weight. The volume of dose and time that it is administered will be recorded on each day of exposure. Care must be taken in order not to exceed the maximum tolerated dose 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**

45. 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 chemical groups are compared to the vehicle group for statistically significant increases in weight.

**Specific procedures for androgen antagonists**

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

## OBSERVATIONS

### Clinical observations

47. 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 are to 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).

48. Any animal found dead will be removed and disposed of without further data analysis. Any mortality of animals prior to necropsy will be included in the study record together with any apparent reasons for mortality.

### Body weight and food consumption

49. 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.

**Dissection and measurement of tissue and organ weights**

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

51. The order in which the animals are necropsied will be designed such that one animal from each of the groups is necropsied in a random fashion before necropsy of the second animal from each group. In this way, all the animals in the same treatment group are not necropsied at once.

52. The five androgen-dependent tissues (VP, SV, LABC, COW, GP) are mandatory measurements. These tissues will 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 (2)(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 weighing 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 weighs a given tissue from all treatment groups as opposed to one individual weighing all tissues from a control group, while someone else is responsible for the treated groups. Each sex accessory tissues will be weighed without blotting to the nearest 0.1 mg, and the weights recorded for each animal.

53. Liver, paired kidney, and paired adrenal weights are optional measurements. Again, tissues should be trimmed free of any adhering fascia and fat. The liver will be weighed and recorded to the nearest 0.1 g, and the paired kidneys and paired adrenals will be weighed and recorded to the nearest 0.1 mg.

54. The serum hormones LH and T are optional measurements. In those cases, the rats will 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. Serum T4 and T3 also are optional measures. These would provide useful supplemental information about the ability to disrupt thyroid hormone homeostasis. 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.

55. 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 program (2)(21).

- 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. 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 mandatory 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 VP right lobe from the urethra and dissect the lobe from the urethra. While still holding right VP lobe of the ventral prostate, 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 SCVG, 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.

56. If the evaluation of each chemical requires necropsy of more animals than is reasonable for a single day, necropsy may be staggered on two consecutive days. In this case the work could be divided so that necropsy of 3 animals per treatment per day (1 cage) takes place on the first day with the dosing and necropsy being delayed by one day in the second half of the animals.

57. Carcasses will be disposed of in an appropriate manner following necropsy.

## REPORTING

### Data

58. Data will 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). The data will be summarized in tabular form. The data will 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.

59. **A final report shall include:**

**Testing facility:**

- Name of facility, location
- Principal investigator and other personnel and their study responsibilities
- Dates study began and ended

**Test substance:**

- Identity, purity, 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

**Vehicle:**

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

**Test animals and animal husbandry procedures:**

- Species/strain used
- Source or supplier of animals, including full address
- Number and age of animals supplied
- Housing conditions (temperature, lighting, and so on), diet used, source of diet, diet lot, bedding and bedding source;
- Caging conditions and number of animals per cage;
- Age at castration and time 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, and test substance groups;
- Mean and standard deviation of the body weights for each group throughout the study;
- Necropsy procedures, including means of exsanguinations and any anesthesia
- If serum analyses are performed, the RIA procedure, source of RIA kits, procedure for scintillation counting, and standardization.

**Results:**

- Daily observations for each animal during dosing, including:
    - Body weights (to the nearest 0.1 g),
    - Clinical signs (if any),
    - If an assay for antiandrogenicity, the TP treatment (dose and volume),
    - Test substance treatment (dose and volume),
    - Time of dosing
    - 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,

- Order of animal exsanguination and dissection at necropsy,
- Weights of five mandatory sex accessory tissues, glands and liver:
    - 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)
    - Glans penis (fresh weight to nearest 0.1 mg), and
    - Cowper's glands (fresh weight – paired, to nearest 0.1 mg).
    - Liver (to nearest 0.1 g)
- Weights of optional tissues, if performed:
    - 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, mandatory tissues weights (VP, SVCG, LABC, GP, and COW), and any optional organ weights.

**Analysis of results**

60. Body and organ weights at necropsy should first be analyzed for their statistical characteristics such as homogeneity beginning with one-way ANOVAs. The ANOVA model could include initial bodyweight on the first day of dosing as a covariate. If any data display heterogeneity of variance, then appropriate data transformations should be employed. Often log transformation is required because the variance is proportional to the mean. If the study was conducted in phases, then the analysis is a two-way ANOVA with Phase and Treatment as main effects, again, bodyweight is used as a covariate. For organ and tissue weight data, bodyweight at necropsy could be used as a covariate in the model if the treatment did not affect growth rates, but the analysis should not include "relative organ" weights, generated by dividing the organ weight by body weight due to the invalid statistical assumptions underlying this data manipulation. Often log transformation of serum hormone data, if collected, is required because the variance is proportional to the mean.

61. For androgen agonism, the test substance groups will be compared to the vehicle control. A statistically significant increase in tissue weight ($p \leq 0.05$) will be considered a positive androgen agonist result. Individual treatment group comparisons to the control group can be done using two tailed t-test or a Dunnett's one tailed test. Androgens, like trenbolone that cannot be 5 alpha reduced have more pronounced effects on the LABC and GP versus TP, but all tissues should display increased growth.

62. For androgen antagonism, the test substance with co-administered reference androgen groups will be compared to the reference androgen control. A statistically significant decrease in tissue weight ($p \leq 0.05$) will be considered a positive antagonist result. If more than one set of

comparisons is required, all comparisons will be conducted separately for each test group against its control. Individual treatment group comparisons to the control group can be done using two tailed t-test or a Dunnett's one tailed test. Antiandrogens that inhibit 5 alpha reductase, like finasteride, have more pronounced effects in the ventral prostate than other tissues as compared to a potent AR antagonist, like flutamide. In the validation studies, the GP was quite responsive to low doses of androgens, while on the other hand, the GP was relatively insensitive to lower doses of weak antiandrogens versus the other androgen-dependent tissues.

63. Those groups attaining statistical significance should be identified.

64. Data should be summarized in tabular form containing the mean, standard error of the mean and sample size for each group. Individual data tables should also be included. The individual values, mean, SE 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 historical CV values 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. Historically, the CVs for the SV, VP, LABC, Cowper's glands and glans penis are about 25%, 25%, 13%, 25% and 12% respectively for antiandrogenic effects and 40%, 40%, 18%, 18% and 35% respectively for androgenic effects. 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.

65. 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, being recording errors.

66. Comparison of study results with historical data 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.

## 68. ANNEX 1
## DEFINITIONS

**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.

**Evident toxicity** is a general term describing clear signs of toxicity following administration of test substance. These should be sufficient for hazard assessment and should be such that an increase in the dose administered can be expected to result in the development of severe toxic signs and probable mortality.

**NOEL** is the abbreviation for no-observed-effect level. In the Hershberger Bioassay, this is the highest dose level where no statistically significant change in any androgen-dependent organ weight is observed due to treatment.

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

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

**Date of birth** is postnatal day 0.

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

**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.

**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.

## 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. 2002. 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. ENV/JM/TG/EDTA(2002)1/REV2/ADD1.

(5) OECD. 2005. 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. ENV/JM/TG/EDTA(2003)4/ ADD1.

(6) There is a draft report on Phase-3 by Gelbke.

(7) Owens, W, Zeiger E, Walker M, Ashby J, Onyon L, Gray, Jr, LE. 2006. The OECD program 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) There is a manuscript on Phase-2 which should be submitted about 1 Sept 2006. A manuscript for Phase-3 is being drafted.

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

(16) Dorfman RI. 1969b. Androgens and anabolic agents. In: *Methods in Hormone Research, volume IIA*. (Dorfman RI, ed.) NewYork:Academic Press. 151-220.

(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 program 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.

Prepared by **LE Gray Jr, JW Owens and G Timm.**