Hershberger Assay
OCSPP Guideline 890.1400

Standard Evaluation Procedure (SEP)

ENDOCRINE DISRUPTOR SCREENING PROGRAM
U.S. Environmental Protection Agency
Washington, DC 20460

October 2011

TABLE OF CONTENTS

I. INTRODUCTION

A. Use of the Standard Evaluation Procedure

This document was developed by EPA to provide guidance to EPA staff who will be reviewing the data submitted in response to Tier 1 Test Orders issued under the Endocrine Disruptor Screening Program (EDSP). This document provides general guidance and is not binding on either EPA or any outside parties. The use of language such as "will," "is," "may," "can" or "should" in this document does not connote any requirement for either EPA or any outside parties. As such, EPA may depart from the guidance where circumstances warrant and without prior notice. The SEPs are intended to be used in conjunction with the EDSP Test Guideline Series 890 and the Corrections and Clarifications document available on the EDSP web page.

This Standard Evaluation Procedure (SEP) provides guidance on how EPA generally intends to review studies conducted using the OCSPP Guideline 890.1400 for the Hershberger Assay that are submitted to support requirements imposed under the U.S. Environmental Protection Agency's Endocrine Disruptor Screening Program (EDSP). The product of the review will be a Data Evaluation Record (DER) that reflects how well the study was performed and conforms to the Guideline and provides the appropriate conclusions supported by the data. The DER will include, for example, a list of any significant deviations from the protocol as well as their potential impacts, a list of significant information missing from the study report, and any other information about the performance of the study that affects interpretation of the data within the context of the EDSP.

The DER should contain adequate information to provide the EPA with the ability to determine whether the study was performed according to the guideline. The objective of EDSP Tier 1 assays is to characterize the potential of a chemical to interact with the endocrine system.

The Guideline recommends the critical materials, methods, and analyses that lead to successful performance of the assay. If a particular material, method, or analysis is named in the Guideline, it is usually because other materials, methods, or analyses are either known to be inappropriate or at least have not been validated or that there is concern for their potential influence on results. The Agency has posted Corrections and Clarifications on Technical Aspects of the EDSP Tier 1 Assays (OCSPP Test Guideline Series 890) in the docket; the link to this document may be found by way of the EDSP web page (http://www.epa.gov/endo/). It is therefore important to note deviations from specific materials, methods, or analyses in the DER, and provide the Agency's opinion on whether the deviation/deficiency has an impact on the performance and results of the study or the acceptability of the study.

II. THE HERSHBERGER ASSAY

A. Purpose of the Assay

The Hershberger Assay is a screening assay intended to identify substances with androgenic and antiandrogenic activity. Androgens are essential for the initial development of the reproductive system *in utero*, the maturation of the reproductive system and accessory sex organs (ASO), development of secondary and tertiary sex characteristics at puberty, and maintenance of the reproductive system and ASO structures and functions until death (OECD, 2008a). Androgens are needed for feedback regulation of the hypothalamus-pituitary-gonadal (HPG) axis and spermatogenesis. Testosterone has two fundamental and different roles at two different life stages: the essential and irreversible role of testosterone in the formation of the male reproductive tract (internal and external) *in utero*; and the role of testosterone in the pubertal development of the male into a reproductively competent individual (with secondary and tertiary male characteristics) and sustaining those male structures and functions through adulthood (this role is reversible with the removal of testosterone).

Testosterone is partially responsible for testis development and descent; maturation of the epididymides, vas efferens, vas deferens, seminal vesicles and coagulating glands, levator ani-bulbocavernosus (LABC) muscle; and for preputial separation (PPS) at puberty in rodents. Dihydrotestosterone (DHT) is responsible for development of the male urethra and prostate, formation of the penis and scrotum, and male secondary sex characteristics such as scrotal growth, development of scrotal rugae and pigmentation, and penis growth and development.

Androgen agonists and antagonists act as ligands for the androgen receptor (AR) and may activate or inhibit, respectively, gene transcription controlled by the receptor. In addition, some chemicals inhibit the enzyme 5α-reductase, which converts testosterone to DHT in some androgen target tissues. Exogenous chemicals acting as androgens or antiandrogens can disrupt the endocrine system and compromise reproductive and developmental processes. Therefore, the need exists to rapidly assess and evaluate a chemical as a possible androgen agonist or antagonist. *In vitro* assays detecting the binding of a ligand for an androgen receptor (OCSPP 890.1150) provide minimal information regarding potential hazard. Considerations of absorption, distribution, metabolism, and excretion (ADME) require *in vivo* testing.

One such *in vivo* test is the Hershberger assay, which uses male rats castrated around the onset of puberty and determines the effect of exposure to potential androgen agonists, antagonists, and 5α-reductase inhibitors on the weights of androgen-dependent ASO. Castrated young rats are used in this assay to establish low physiological levels of endogenous hormone levels in the androgen-dependent target tissues that are highly responsive to the administration of exogenous testosterone. When evaluating potential antiandrogens, castration with administration of exogenous testosterone ensures a more constant level of hormone exposure in the target tissues and eliminates the HPG feedback loop that can mask antiandrogenic effects on the target tissue (Gray et al, 2005).

B. Study Design

The Hershberger Bioassay is a short-term *in vivo* screening test using accessory tissues of the male reproductive tract (Hershberger et al., 1953). The assay originated in the 1930s and was modified in the 1940s to include androgen-responsive muscles in the male reproductive tract (U.S. EPA, 2009). In the 1960s, over 700 possible androgens were evaluated using a standardized version of the protocol. Consequently the use of the assay for both androgens and antiandrogens was considered a standard method in the 1960s. The current bioassay is based on the changes in weight of five androgen-dependent tissues in the castrated-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 androgen-dependent tissues included in this Test Guideline are the ventral prostate, seminal vesicle (plus fluids and coagulating glands), LABC, paired Cowper's glands, and the glans penis.

In male rats castrated around the time of puberty, these five tissues all respond to androgens with an increase in absolute organ/tissue weight. These same tissues are stimulated to increase in weight by exposure to a potent reference androgen, since the castration removes the primary source of endogenous testosterone. Likewise, these five tissues all respond to antiandrogens with a decrease in absolute organ/tissue weight. The primary model for the Hershberger bioassay has been the surgically-castrated peripubertal male, which was validated by the Organization for Economic Cooperation and Development (OECD) during development of OECD Test Guideline 441. This Test Guideline was used by EPA as the source material for development of the harmonized OCSPP Guideline 890.1400.

Due to animal welfare concerns with the castration procedure, the intact (uncastrated) stimulated weanling male was investigated as a potential alternative model for the Hershberger Bioassay to avoid the castration step (OECD, 2009a, b, c). Although the stimulated weanling test method was validated, the weanling version of the Hershberger Bioassay was unable to consistently detect effects of weak anti-androgens on androgen-dependent organ weights.

III. EVALUATION OF STUDY CONDUCT

This section provides a summary description of the information that would generally be expected to be obtained from a study that had been conducted following the recommendations in the Test Guideline. As described in this section, the DER reviewer is responsible for summarizing how the study was conducted, the extent to which that is consistent with the Guideline, and how, if at all, that affected the validity of the study. This information will factor into the Agency's interpretations of the data contained in the study report. Specific points that are important for the DER to address are highlighted in the individual sections below, as appropriate.

The summary in this section is offered as a general outline to aid in preparation of the DER. The purpose of this section is not to serve as substitute for the Test Guideline, nor to provide any guidance on how the study should be conducted. Rather, the summary is intended to provide context and examples illustrating to the reviewer what the DER would be expected to contain.

A. Test Compound

The purity of the test compound used in the study, its source, and Lot and/or Batch number should be reported in the DER, along with the results of the certificate of analysis. The DER should also include data on the stability of the test chemical in the vehicle for concentrations that bracket those used in the study, and the storage temperature used for the stability analyses. For test substances dosed as suspensions, the homogeneity of the test chemical suspensions should be reported in the DER. Actual concentrations of the test chemical solutions/suspensions are typically reported for at least one preparation day during the study.

B. Controls and Vehicle

The guideline recommends that the reference androgen agonist be testosterone propionate (TP), CAS No 57-82-5 at a dose of 0.2 or 0.4 mg/kg/day. The recommended reference androgen antagonist is flutamide (FT), CAS No 1311-84-7 at a dose of 3 mg/kg/day. Information on the supplier, catalogue and batch number, purity, and CAS number of the positive control for androgen agonists (TP) and antagonists (FLU) should be reported in the DER. Additionally, information on any vehicle (identity, selection, supplier, and lot number) should be included in the methods section of the DER (see Section F below for details).

C. Dose Formulations

For the dose formulations, an aqueous solution/suspension is typically considered first. However, many androgen ligands are hydrophobic, necessitating solution/ suspension in oil (e.g. corn, peanut, sesame or olive oil). The vehicle choice and any justification (if other than water) should be provided in the DER. 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. It is desirable that the toxic characteristics of the solvent be known, and also be tested in a separate solvent-only control group. Information regarding the dose formulation preparation, storage, and analyses (for stability, homogeneity, and concentration) should be included in the appropriate section of the DER. As long as conditions (storage time and temperature) are similar, demonstration of homogeneity and stability of the dose formulations do not need to be conducted concurrent with the study. However, analysis of appropriate concentration conducted on dose formulations used during the assay should be documented in the DER.

D. Test Animals

The rat has been routinely used in the Hershberger Bioassay since the 1930s, and this species was validated by the OECD (OECD, 2008b). Strains that mature substantially later than PND 42 are not recommended because castration of these males at PND 42 can prevent PPS and preclude measurement of the glans penis. Sprague Dawley and Wistar strains are preferred. Fisher 344 rats are not recommended because they sexually mature later. The DER will include: the species and strain used and the rationale for this choice; the source and supplier of the animals; and the number and age of the animals at receipt, dosing initiation, and necropsy.

E. Animal Husbandry

The Test Guideline recommends that all procedures for animal care and use be documented and in compliance with regulatory standards. The recommended environmental conditions include: temperature 22 ± 3ºC; relative humidity 30-70% (preferably 50-60%); and 12-hour light/dark cycle. Group housing of two or three rats per cage is preferred, in a cage of approximately 2000 cm^2. Laboratory diet and water are typically provided *ad libitum*. The guideline recommends that performing laboratories use the laboratory diet typically used in other studies.

The Corrections and Clarifications on Technical Aspects of the EDSP Tier 1 Assays (OCSPP Test Guideline Series 890) states that tap water is variable and may contain potentially endocrine active substances. As such, tap water is generally not a recommended source of drinking water. Acceptable sources of drinking water typically include deionized, double-distilled water and charcoal-filtered water. Other sources may also be acceptable. However, the presence of soluble organic chemical contaminants such as natural and artificial hormones in drinking water has the potential to introduce variability into (and potentially compromise) the results (e.g. result in false negatives or false positives). Consequently, if an alternative source of water has been used, EPA recommends that the laboratory document that such contaminants have been removed from the drinking water.

F. Experimental Design

For the agonist component where two dose levels are recommended, the guideline recommends that the highest dose level be at or just below the Maximum Tolerated Dose (MTD) but need not exceed the Limit Dose (1000 mg/kg/day); however, typically, the Agency also considers the toxicity profile of the chemical (i.e., cholinesterase inhibition, target organ toxicity, etc.) in dose selection. EPA recommends that the second dose level typically be spaced to produce a lesser degree of toxicity relative to the high dose unless justification is provided for testing at a different level. The DER should contain the rationale for the selection of doses.

For the antagonist component where three dose levels are recommended, it is recommended that a similar rationale be used: if the highest dose level results in toxic effects and provided it does not produce an incidence of fatalities which would prevent a meaningful evaluation, then it is recommended that the intermediate dose levels be spaced to produce a gradation of toxic effects; and that the lowest dose level produce no evidence of toxicity. Previously conducted toxicity studies can provide the necessary information to determine the appropriate doses of the test substance. However, a dose range-finding study may be necessary to aid in the dose-selection for the agonist component of the assay.

G. Assay Procedures

1. Preparation of Animals for Treatment

After an initial acclimation period of several days upon receipt, the rats are castrated on PND 42 or thereafter (not before) and provided a recovery period of at least 7 days following castration to allow for regression in the target tissues. The 10 days of dosing can be initiated as early as PND 49, but not later than PND 60. It is recommended that age at necropsy not be greater than PND 70. Information regarding the age at castration and the age at initiation of treatment should be included in the DER under the appropriate section in the methods regarding animal strain, age, and source. The DER should evaluate this information in terms of the age of onset of puberty for the strain used (i.e., Sprague-Dawley or Wistar rats).

Animals at the extremes of the body weight range are excluded and the remaining animals randomly assigned to the treatment groups, stratified by body weight, so that there were no statistically significant differences in initial mean body weight for each group. At the commencement of the study, the variability in initial body weights should be within 20% of the mean group weight. The reviewer should include information regarding animal assignment in the methods section of the DER.

2. Dose Administration

The DER will include a description of the dose administration for the test formulations, TP, and flutamide in the DER (i.e., the route of administration and its rationale, dose schedule, and dose volume). The guideline recommends TP be administered by subcutaneous (s.c.) injection, at a dose of 0.2-0.4 mg/kg bw/day, and flutamide via oral gavage at a dose of 3 mg/kg bw/day. The positive control for antiandrogens is TP + flutamide, which are co-administered. The route of administration for the test substance may be via oral gavage or s.c. injection, taking into account considerations such as: animal welfare; physical/chemical properties; relevance to expected route of exposure for humans; and existing toxicology data on metabolism and kinetics. Animals are dosed for 10 consecutive days at approximately 24-hour intervals, with the dose level adjusted daily based on the concurrent body weight measurement. The dose volume for oral gavage administration typically does not exceed 10 mL/kg body weight for aqueous solutions or 5 mL/kg for oil suspensions or solutions. The guideline recommends subcutaneous injection be administered in the dorsoscapular and/or lumbar region and not exceed 0.5 mg/kg body weight.

3. Treatment Groups

A description of the study design, including treatment groups (positive controls, negative controls, and test substance), dose levels, and number of animals is included in the methods section of the DER. The DER template provides a table format for summarizing this information.

For the test for androgen agonists, the vehicle is the negative control, and the TP-treated group is the positive control. The five androgen-dependent tissues include the ventral prostate, seminal vesicle (plus fluids and coagulating glands), LABC, paired Cowper's glands, and the glans penis. The weights of the five tissues are compared to the vehicle controls for statistically significant increases. Two dose groups of the test substance plus positive and vehicle (negative) controls are normally sufficient to determine if a chemical is an androgen agonist, and this design is therefore preferred for animal welfare reasons.

Anti-androgen activity, via AR antagonism or 5α-reductase inhibition, is tested by administering a reference dose of TP and administering the test substance for 10 consecutive days. For the test for androgen antagonists and 5α-reductase inhibitors, the TP-treated group is the negative control, and the group co-administered reference doses of TP and flutamide is the positive control. The weights of the five tissues from the TP plus test substance groups are compared to the reference TP-only group for statistically significant decreases. A minimum of three test groups with different doses of the test chemical and a positive and a negative control are used to determine if a chemical is an anti-androgen.

Table 1. Study Design

Test group	Dose (mg/kg/day)	# of Males
Androgen Agonist Assay		
Vehicle control (negative control)	0	6
Low Dose	#	6
High Dose	#	6
Testosterone propionate (TP) (s.c.), positive control	# [0.2 or 0.4]	6
Anti-Androgen Assay		
Vehicle control	0	6
Testosterone propionate (negative control, s.c.)	# [0.2 or 0.4]	6
Low Dose (+TP)	#	6
Mid Dose (+TP)	#	6
High Dose (+TP)	#	6
Flutamide(oral gavage) + TP (s.c.), positive control	3	6

4. Observations

Consistent with good scientific practice, clinical observations are conducted at least once daily for mortality and signs of toxicity. If any clinical signs of toxicity are reported, the nature, incidence, severity, and duration are included in the DER, along with an assessment of whether the findings were considered adverse and related to treatment. As per the guideline, body weights are obtained daily, however food consumption expressed as g/animal/day (on a cage basis) is optional. Body weights and body weight change from the beginning of dosing until necropsy should also be included in the DER.

5. Necropsy and Organ Weights

The guideline recommends that approximately 24 hours after the final administration of the test substance, the rats be euthanized, exsanguinated, and necropsied according to standard laboratory procedures. Weights of the five androgen-dependent tissues (ventral prostate, seminal vesicles, LABC, Cowper's gland, and glans penis) are critical measurements. Individual data for organ weights are reported to the nearest 0.1 mg. The organ weight data will generally be presented after covariance adjustment for initial body weight, but this should not replace presentation of the unadjusted (absolute) data.

The guidelines recommend that organs be carefully excised and trimmed, and their fresh (unfixed) weights recorded. It is imperative that the performing laboratory personnel exercise caution to prevent loss of fluids from the organs or desiccation during this process, because this may skew the measurement. Optional organ weights include the liver, kidneys, and adrenals; the reviewer should evaluate any substantial differences in weights of these organs both in the context of the androgen dependent tissue weights and as indicators of systemic toxicity (for additional guidance refer to Owens *et al*., 2006 and Korea FDA, 1999, Gray *et al*., 2005).

6. Hormone Measurements

The Test Guideline states that serum hormone measurements [LH, FSH and testosterone (T)] are optional, but may be useful in interpreting potential findings in the study. If the performing laboratory measured serum hormone levels, the DER will include methods information regarding: hormones measured; method of anesthesia; blood collection; serum preparation; and analytical procedures (e.g., radioimmunoassay). The guideline recommends the following procedures for hormone measurements: 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.

H. Statistical Analysis

The Guideline recommends that the body and organ weight data are statistically analyzed for characteristics such as homogeneity of variance with appropriate data transformations as needed. The Guideline also recommends analyzing treatment groups with homogeneous variances using ANOVA to determine differences among groups followed by pair-wise comparisons with the controls using a post-hoc test that appropriately adjusts for multiple comparisons (e.g., Dunnett's test). If variances are not homogeneous, non-parametric procedures, such as Kruskal-Wallis and Dunn's test, are recommended. The criterion for statistical significance is identified (e.g., $p \leq 0.05$), and groups attaining statistical significance are denoted. In the DER tables, the asterisk (*) is the conventional means of indicating statistical significance. The Guideline recommends that analysis of "relative" organ weights be avoided because there are invalid statistical assumptions underlying these data manipulations. Absolute organ weights are presented and organ weights adjusted using initial body weight (not terminal body weight) as a covariate are included in the DER. Summary data tables, reporting mean,

standard deviation, and CV for each of the organ weights are to be included in the DER for each assay.

IV. STUDY INTERPRETATION

A. Results

1. Mortality

The DER should include mortality and assess whether the animal deaths were due to treatment with the test substance. If treatment-related mortality occurred, the assay would typically have been repeated at lower doses because the maximum tolerated dose was exceeded.

2. Clinical Signs of Toxicity

The DER should include clinical signs of toxicity, including information on the nature, incidence, severity, onset, and duration. As with the mortality data, the reviewer should assess whether the findings are considered adverse and related to treatment and if they had an impact on dose selection.

Table 2. Incidence of Clinical Observations in the Androgen Agonist Assay) [a]

Observation	Dose (mg/kg/day)							
	Vehicle Control		TP (#)		Low (#)		High (#)	
	# Observed	# Examined	# Observed	# Examined	# Observed	# Examined	# Observed	# Examined

[a] Data were obtained from page [#] of the study report.

Table 3. Incidence of Clinical Observations in the Anti-Androgen Assay [a]

Observation	Dose (mg/kg/day)											
	Vehicle Control		TP Negative Control		TP + FT Positive Control		Low (#)		Mid (#)		High (#)	
	# Obs.	# Exam	# Obs.	# Exam	# Obs.	# Exam	# Obs.	# Exam	# Obs.	# Exam	# Obs.	# Exam

[a] Data were obtained from page [#] of the study report.
Obs. = number of observed incidences.
Exam = number of animals examined.

3. Body Weight and Weight Gain

The DER should include data for the first and final days of dose administration, along with body weight gain for the overall study (Days 1-10). Any effects of treatment on body weights including information on statistical significance, magnitude difference from controls, and the onset and duration of the decreases should be reported.

Table 4. Selected Group Mean (±SD) Body Weights and Cumulative Body Weight Gains (g) in the Androgen Agonist Assay [a]

Study Day #	Dose (mg/kg/day)											
	Vehicle Control			TP (#)			Low (#)			High (#)		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
#												
#												
#												
Body Weight Gain (days #-#)												

[a] Data were obtained from Tables [#]-[#] on pages[#]-[#] of the study report. Percent differences from controls were calculated by the reviewers and included in parentheses.
N = Number of animals in the group
SD= Standard Deviation
* Significantly different from controls at p<0.05
** Significantly different from controls at p<0.01

Table 5. Selected Group Mean (±SD) Body Weights and Cumulative Body Weight Gains (g) in the Anti-Androgen Assay [a]

Study Day #	Dose (mg/kg/day)																	
	Vehicle Control			TP Negative Control			TP + FT Positive Control			Low (#)			Mid (#)			High (#)		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
#																		
#																		
#																		
Body Weight Gain (days #-#)																		

[a] Data were obtained from Tables [#]-[#] on pages[#]-[#] of the study report. Percent differences from controls were calculated by the reviewers and included in parentheses.
N = Number of animals in the group
SD= Standard Deviation
* Significantly different from controls at p<0.05
** Significantly different from controls at p<0.01

4. Food Consumption

Data should be reported as shown in the following tables. Food consumption information is documented, as necessary, to assess any effects on body weight or body weight gain.

Table 6. Food Consumption (g/kg/day) in the Androgen Agonist Assay[a]

Study Day #	Dose (mg/kg/day)											
	Vehicle Control			TP (#)			Low (#)			High (#)		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
#												
#												
#												
(Days #-#)												

[a] Data were obtained from page [#] of the study report.
N = Number of animals in the group
SD= Standard Deviation
* Significantly different from controls at $p<0.05$
** Significantly different from controls at $p<0.01$

Table 7. Food Consumption (g/kg/day) in the Anti-Androgen Assay [a]

Study Day #	Dose (mg/kg/day)																	
	Vehicle Control			TP Negative Control			TP + FT Positive Control			Low (#)			Mid (#)			High (#)		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
#																		
#																		
#																		
(Days #-#)																		

[a] Data were obtained from page [#] of the study report.
N = Number of animals in the group
SD= Standard Deviation
* Significantly different from controls at $p<0.05$
** Significantly different from controls at $p<0.01$

5. Hormone Measurement

If the optional serum hormone measurements (LH, FSH and T) are provided, describe results in text and include a table, as necessary (refer to Tables 8 and 9). The mean, SD, CV, number of animals, and p-value for LH, FSH and T levels for each treatment group, including vehicle control, should be reported. Also, the normal range for each parameter should typically be provided in the study report and appropriate data appended to the DER. The DER should indicate whether these normal values are from the literature (provide reference) or from laboratory historical controls. A comparison will be made by the DER Reviewer for the normal

values to those observed during the study. Dose group parameters that are significantly different from the vehicle control group ($p < 0.05$) should be indicated.

Table 8. Hormone Measurement is in the Androgen Agonist Assay[a]

Hormone	Dose (mg/kg/day)															
	Vehicle Control				TP (#)				Low (#)				High (#)			
	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV
Serum T (ng/mL)																
Serum LH (ng/mL)																
Serum FSH (ng/mL)																

[a] Data were obtained from page [#] of the study report.
N = Number of animals in the group
SD= Standard Deviation
CV= Coefficient of Variation
* Significantly different from controls at $p<0.05$
** Significantly different from controls at $p<0.01$

Table 9. Hormone Measurement in the Anti-Androgen Assay [a]

Hormone	Dose (mg/kg/day)																							
	Vehicle Control				TP Negative Control				TP + FT Positive Control				Low (#)				Mid (#)				High (#)			
	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV
Serum T (ng/mL)																								
Serum LH (ng/mL)																								
Serum FSH (ng/mL)																								

[a] Data were obtained from page [#] of the study report.
N = Number of animals in the group
SD= Standard Deviation
CV= Coefficient of Variation
* Significantly different from controls at $p<0.05$
** Significantly different from controls at $p<0.01$

6. Organ Weight

The DER should contain tables with the mean, SD, CV for both the number of animals (N) and weights for the following organs (refer to Tables 10 and 11): Seminal vesicles; Ventral prostate; LABC; Cowper's glands; and Glans penis. Results of statistical analysis comparing the treated groups relative to the same measures in the reference androgen control group should be reported. Dose group parameters that are significantly different from the vehicle control group ($p < 0.05$) should be indicated.

Table 10. Accessory Sex Organ Weights from Androgen Agonist Assay in [Sprague-Dawley or Wistar] Rats[a]

Organ	Dose (mg/kg/day)															
	Vehicle control				Low (#)				High (#)				TP (#)			
	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV
Seminal vesicles																
Ventral prostate																
LABC																
Cowper's glands																
Glans penis																

[a] Data were obtained from Tables [#]-[#] on pages[#]-[#] of the study report. Percent differences from controls were calculated by the reviewers and included in parentheses.
N = Number of animals in the group
SD= Standard Deviation
CV= Coefficient of Variation
* Significantly different from controls at $p<0.05$
** Significantly different from controls at $p<0.01$

Table 11. Accessory Sex Organ Weights from Anti-Androgen Agonist Assay in [Sprague-Dawley or Wistar] Rats [a]

Organ	Dose (mg/kg/day)																							
	Vehicle Control				TP (negative control)				Low (#) (+TP)				Mid (#) (+TP)				High (#) (+TP)				Flutamide (positive control)			
	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV	N	Mean	SD	CV
Seminal vesicles																								
Ventral prostate																								
LABC																								
Cowper's glands																								
Glans penis																								

[a] Data were obtained from Tables [#]-[#] on pages[#]-[#] of the study report. Percent differences from controls were calculated by the reviewers and included in parentheses.
N = Number of animals in the group
SD= Standard Deviation
CV= Coefficient of Variation
* Significantly different from controls at $p<0.05$
** Significantly different from controls at $p<0.01$

B. Data Interpretation Procedure

The guideline recommends that the highest dose level be at or just below the MTD but need not exceed the Limit Dose (1000 mg/kg/day); however, typically, the Agency also considers the toxicity profile of the chemical (i.e., cholinesterase inhibition, target organ toxicity, etc) in dose selection. EPA recommends that, typically the second dose level be spaced to produce a lesser degree of toxicity relative to the high dose unless justification is provided for testing at a different level.

If significant clinical signs of toxicity are noted in a dose-dependent manner, it is possible that the dose selection was too high; the reviewer should consider whether the assay needs to be repeated using lower doses. Additionally, it is possible to observe clinical signs and/or mortality that are deemed unrelated to treatment due to a lack of dose-relationship and/or an undetermined cause of death. However, incidences of mortality and moribundity reduce sample size which results in a loss of power to detect statistical differences. Therefore, a lower sample size could result in androgens or anti-androgens testing as false negatives.

The DER Reviewer should statistically analyze terminal body and organ weights for characteristics such as homogeneity of variance with appropriate data transformations as needed. The Reviewer should compare treatment groups 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, the DER Reviewer should avoid reliance on "relative organ" weights due to the invalid statistical assumptions underlying this data manipulation.

Serum hormone measurements (T, LH, and FSH) are optional, but may be useful in interpreting potential findings in the study. For example, if the test substance induces liver metabolism of testosterone, serum testosterone levels may be lower. Without the measurement of serum testosterone, such an effect may be misinterpreted as an anti-androgen. LH and FSH levels provide information about the ability of an antiandrogen not only to decrease organ weights, but also to affect the hypothalamic-pituitary function.

The control group is the vehicle-only test group for androgen agonism and the TP-only group for androgen antagonism. The five androgen-dependent tissues included in this Test Guideline are the ventral prostate, seminal vesicle (plus fluids and coagulating glands), LABC, paired Cowper's glands, and the glans penis. A statistically significant increase for androgen agonist or decrease for androgen antagonist ($p \leq 0.05$) in any two or more of the five required androgen-dependent tissue weights is indicative of a positive result**.** When interpreting the results of this assay the reviewer is advised that it is also possible that chemicals that increase metabolism may also cause reductions in these tissue weights. 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.

C. Performance Criteria for Organ or Tissue Weights

The OECD validation studies demonstrated that each organ (or tissue) has an "inherent"

coefficient of variation (CV), with some organs having less variability (i.e., lower CV) because the organ is relatively large and therefore more precisely dissected on a consistent basis, while other organs are more variable because they are smaller and/or contain fluids that induce greater variability. Therefore, the Hershberger Bioassay performance criteria are based on maximum CV values for each tissue (Table 12**).** The table below is intended to be used as a worksheet for comparison of the observed results to the performance criteria. Report the relevant findings in the DER.

Table 12. Performance Criteria [Maximum Coefficient of Variation (CV)] for Mandatory Androgen Dependent Tissues for the Castrate Model Based on the OECD Validation studies[a]

Organ/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%

[a]These values were derived from the validation studies conducted by the OECD (OECD, 2006, 2007, 2008b).

D. Laboratory Performance

Prior to the study, a baseline positive control study is typically conducted to demonstrate laboratory proficiency. However, this demonstration of laboratory proficiency prior to initiation of the study is not necessary because concurrent positive controls (TP and flutamide) and negative controls are included in each component of the assay.

V. CHARACTERIZATION OF FINDINGS

On completion of the review of this assay, the Agency will conduct a weight of evidence analysis to consider the potential of the chemical to disrupt the estrogen, androgen, or thyroid hormone systems. Chemicals with demonstrated evidence of a potential to interact with the estrogen, androgen, and/or thyroid hormone systems will be considered as candidates for Tier 2 testing.

VI. DATA EVALUATION REPORT

Once the study has been reviewed using the principles described in the previous sections of this SEP, a DER will be prepared. A DER template is available that provides additional guidance for the preparation of the DER.

VII. REFERENCES

Corrections and Clarifications on Technical Aspects of the EDSP Tier 1 Assays (OCSPP Test Guideline Series 890) in the docket; the link to this document may be found by way of the EDSP web page (http://www.epa.gov/endo/)

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.

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.

Korea Food and Drug Administration (1999). Visual reference guide on Hershberger assay procedure, including a dissection video.

OECD (2006) Series on Testing and Assessment No. 62. Final OECD Report of the Initial Work Towards the Validation of the Rat Hershberger Assay: Phase-1, Androgenic Response to Testosterone Propionate, and Anti-Androgenic Effects of Flutamide. ENV/JM/MONO(2006)30

OECD (2007) Series on Testing and Assessment No. 73. Report of the Validation of the Rat Hershberger Assay: Phase 3: Coded Testing of Androgen Agonists, Androgen Antagonists and Negative Reference Chemicals by Multiple Laboratories. Surgical Castrate Model Protocol. ENV/JM/MONO(2007)20

OECD (2008a) Background Review Document on the Rodent Hershberger Bioassay, Series on Testing and Assessment No. 90, ENV/JM/MONO(2008)17, OECD, Paris.

OECD (2008b). Environmental Health and Safety Publications Series on Testing and Assessment, No. 86. Report of the OECD Validation of the Rodent Hershberger Bioassay: Phase 2: Testing of Androgen Agonists, Androgen Antagonists, and a 5α-Reductase Inhibitor in Dose Response Studies by Multiple Laboratories. ENV/JM/MONO(2008)3, pp. 1-119.

OECD (2009a) Report of the Validation of Hershberger Bioassay (Weanling Model), Series on Testing and Assessment No. 108, ENV/JM/MONO(2009)13, OECD, Paris.

OECD (2009b) Report of the Validation Peer Review for the Weanling Hershberger Bioassay and Agreement of the Working Group of National Coordinators of the Test Guidelines Programme on the Follow-up of this Report, Series on Testing and Assessment No. 110, ENV/JM/MONO(2009)32, OECD, Paris.

OECD (2009c) Guidance Document on the Weanling Hershberger Bioassay in Rats: A Short-term Screening Assay for (Anti)Androgenic Properties, Series on Testing and Assessment No. 115, ENV/JM/MONO(2009)41, OECD, Paris.

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. Supplemental materials for Owens et al. (2006). See section II, The dissection guidance provided to the laboratories. Available at http://www.ehponline.org/docs/2006/8751/suppl.pdf.

U.S. EPA (2009). Endocrine disruptor screening program test guidelines OPPTS 890.1400: Hershberger Bioassay. EPA 740-C-09-008. Washington DC.