Secretariat document in support to the Peer Review of the validation of the Hershberger Bioassay

This document was prepared by the Secretariat to present the current status of the validation of the Hershberger Bioassay in rodent within the context of validation criteria outlined in Guidance Document 34. In particular, each of the 8 criteria required for successful validation of a new test method are listed along with details of how these criteria have been addressed during the OECD exercise and how the available information supports the validity of this method. This document is intended to guide the Hershberger Peer Review Panel in their review of supporting material and provide a framework for their discussions.

References to the different documents of the peer review package are identified as presented below and followed by the paragraph number:

- Validation report of Pase-1: VR1
- Validation report of Pase-2: VR2
- Validation report of Pase-3: VR3
- Draft test Guideline: DTG

GD34 - Criteria 1/8

a) **The rationale for the test method should be available.**
This should include a clear statement of the scientific basis, regulatory purpose and need for the test.

Scientific basis

There is widespread concern that ambient environmental levels of chemicals may be causing adverse effects in humans and wildlife due to the interaction of these chemicals with the endocrine system. Initial reviews of existing literature have noted limited evidence for endocrine disruption in humans, but there are several reports indicating that local, high level exposures to environmental pollutants have resulted in endocrine-related effects in wildlife (VR1, ii, 1). In addition, laboratory investigations have suggested that substances to which human populations are exposed have the potential to subtly alter the development of the male reproductive tract by interfering with normal signalling at the androgen receptor (AR). The rodent Hershberger Bioassay is specifically intended to identify chemicals that can inappropriately activate the AR (agonists) or impair the activation of the AR (antagonists; anti-androgens) (VR1, 6). The use of an in vivo model integrates substance metabolism, distribution and excretion while measuring physiologically-relevant responses in the target tissues of the male reproductive tract (DTG, 5).

Regulatory purpose and need for the test

The regulatory need for this assay stems from the absence of existing OECD Test Guidelines which enable the detection of endocrine active substances *in vivo* (VR1, 1).

The OECD initiative to develop and validate *in vitro* and *in vivo* assays for the detection of chemicals that may interfere with the endocrine response was taken following the recommendations of a number of national, regional and international workshops and following a detailed OECD review of the status of existing test and research methods (VR1, 2).

A conceptual framework for the testing and assessment of chemicals has been developed to identify short- and long-term assays of increasing complexity and detail to gather information on potential endocrine disrupters. The assays and techniques include: 1) structure-activity relationships; 2) *in vitro* assays that would identify a chemical based on its ability to bind androgen or estrogen receptors, or to effect transcriptional activation of hormonal-responsive elements *in vitro*; 3) short-term *in vivo* assays to demonstrate relevant activity in the intact animal, (e.g., the uterotrophic assay, and the Hershberger assay); and 4) long-term assays involving exposure to the test substance at different stages of the development of the animal, e.g., the two-generation reproductive assay. The Hershberger assay is intended to be used within framework (VR1, 3).

GD34 - Criteria 2/8

b) **The relationship between the test method's endpoint(s) and the (biological) phenomenon of interest should be described.**

This should include a reference to scientific relevance of the effect(s) measured by the test method in terms of their mechanistic (biological) or empirical (correlative) relationship to the specific type of effect/toxicity of interest. Although the relationship may be mechanistic or correlative, test methods with biological relevance to the effect/toxicity being evaluated are preferred.

The Hershberger assay is intended to identify chemical agents which directly activate or block the activation of the AR by endogenous androgens. The toxicological significance of this activity lies in the fact that androgen signalling is essential for the development and function of the male reproductive tract. Further, exposure to androgen agonists or antagonists during critically sensitive periods of development can result in permanent adverse effects on fertility and increase risk of reproductive tract pathology.

The Hershberger assay is based on the principle that a number of organs and accessory sex tissues in the male reproductive tract require androgens to stimulate and to maintain growth. If the endogenous source of this hormone is not available, either because of immaturity of the animals or because the animals have been surgically castrated, the animal requires an exogenous source to initiate and/or restore the growth of these tissues (VR1, 5; VR2, 4). Chemicals that act as agonists may be identified if they cause a statistically significant increase in the weights of the target androgen-dependent tissues, or chemicals may be identified as antagonists if they cause a statistically significant decrease in target tissues when co-administered with a potent androgen (DTG, 14, 15, 61, 62).

The Hershberger assay was then designed to measure weight increases in androgen-responsive tissues in animals not currently synthesizing endogenous testosterone. The biological relevance of the response evaluated is established.

For the three phases of the validation, the mandatory tissues to be weighed in both the androgen effect procedure and the anti-androgen procedure were the (VP1, 32; VP2, 32; VP3, 20):

- ventral prostate; fresh tissue, and fixed (24-hr) tissue (VP);
- seminal vesicles plus coagulating glands (including fluid) (SV);
- levator ani and bulbocavernous muscle (LABC);
- Cowper's (or bulbourethral) glands (COWS); and
- glans penis (GLANS).

In phase-1, additional mandatory measurements were:

- individual, daily body weights;
- liver weights

Optional measurements were also proposed in phase-1 -2 or -3, such as liver, adrenals or kidney weight or serum hormone concentration.

GD34 - Criteria 3/8

c) **A detailed protocol for the test method should be available.**

The protocol should be sufficiently detailed and should include, *e.g.*, a description of the materials needed, such as specific cell types or construct or animal species that could be used for the test (if applicable), a description of what is measured and how it is measured, a description of how data will be analysed, decision criteria for evaluation of data and what are the criteria for acceptable test performance.

A detailed protocol is provided, including test materials needed: rat (age, weight), castration conditions, administration of doses, clinical observations, necropsy protocol, including what needs to be necropsied and how. Reference is made to a standard operating procedure (SOP) manual for dissection, available from the OECD. The protocol states that careful training according to the SOP guide will minimize a potential source of variation in the study (DTG, 52).

The decision criteria for evaluation of data are also provided in the draft test guideline:
- For androgen agonism, the test substance groups will be compared to the vehicle control. A statistically significant increase (based on the ANOVA model) in tissue weight will be considered a positive androgen agonist result (DTG, 61).
- For androgen antagonism, the test substance with co-administered reference androgen groups will be compared to the reference androgen control. A statistically significant decrease (based on the ANOVA model) in tissue weight will be considered a positive antagonist result (DTG, 62).

GD34 - Criteria 4/8

d) **The intra-, and inter-laboratory reproducibility of the test method should be demonstrated.**

Data should be available revealing the level of reproducibility and variability within and among laboratories over time. The degree to which biological variability affects the test method reproducibility should be addressed.

Intra-laboratory variability as a measure of repeatability

All of the laboratories participating in Phase-3 had participated in Phase-1 and -2 thereby enabling a comparison with the data obtained in the previous independent test series. The data clearly substantiated a good intra-laboratory reproducibility over time, although for each test chemical the number of laboratories available for such a direct comparison was limited (VR3, vii, xiii, xiv).

Inter-laboratory variability as a measure of reproducibility

In Phase -1, all laboratories and all protocols were successful in detecting increases in the weights of the accessory sex organs and tissues in response to testosterone propionate, and in detecting the anti-androgenic effects of flutamide. There was good agreement among laboratories with regard to the dose responses obtained. There was similar agreement in their ability to identify the anti-androgenic effects of flutamide. It could be concluded from this first phase of the work that the protocol is robust, reliable and transferable across laboratories for potent androgen agonists and antagonists (VR1, viii, ix).

In Phase-2, the validation programme successfully achieved the goal of demonstrating the reproducibility of the protocol for detecting weaker androgen agonists and antagonists as well as 5α-reductase.

In phase-3, the results obtained with the coded androgen agonists, androgen antagonists, and negative chemicals also indicated that all of these test chemicals were correctly identified in the majority of the investigated target organs as to their anticipated activity in the Hershberger Bioassay. A high inter-laboratory reproducibility was thus clearly demonstrated for both protocols (VR3, xi, xii). See tables below, provided for trenbolone, linuron and pp'DDE, as well as for the 2 negative chemicals:

Trenbolone – 13 Studies

		VP	SVCG	LABC	GP	Cows
1.5 mg TREN/ kg-bw/d	T-test	0/13	0/13	2/13	0/13	1/13
	Dunnett's	0/13	1/13*	2/13*	0/13	1/13*
40 mg TREN /kg-bw/d	T-test	12/13	13/13	13/13	12/13	12/13
	Dunnett's	10/13	12/13	13/13	12/13	11/13

* 3 of 4 instances of significance in Laboratory 2

Linuron – 11 Studies 0.4 mg TP/kg-bw/d
3 Studies 0.2 mg TP/kg-bw/d

	TP *		VP	SVCG	LABC	GP	Cows
10 mg Linuron/ kg-bw/d	0.2	T	1/3	2/3	0/3	0/3	0/3
		D	0/3	1/3	0/3	0/3	0/3
	0.4	T	1/11	1/11	1/11	0/11	1/11
		D	1/11	1/11	1/11	0/11	1/11
100 mg Linuron/ kg-bw/d	0.2	T	3/3	3/3	3/3	2/3	2/3
		D	3/3	3/3	3/3	3/3	2/3
	0.4	T	10/11	11/11	10/11	5/11**	8/11
		D	9/11	10/11	10/11	5/11**	9/11

* TP mg/kg-bw/d; T = t-test; D=Dunnett's; Laboratory 5 did not dissect the GP, so no data available, not a false negative

p,p'-DDE – 7 Laboratories 0.4 mg TP/kg-bw/d
3 Laboratories 0.2 mg TP/kg-bw/d

	TP *		VP	SVCG	LABC	GP	Cows
16 mg *p,p'*-DDE/ kg-bw/d	0.2	T	2/3	3/3	1/3	1/3	1/3
		D	2/3	2/3	1/3	0/3	1/3
	0.4	T	1/7	1/7	1/7	1/7	1/7
		D	1/7	1/7	1/7	2/7	1/7
160 mg *p,p'*-DDE/ kg-bw/d	0.2	T	3/3	3/3	3/3	3/3	3/3
		D	3/3	3/3	3/3	3/3	3/3
	0.4	T	7/7	7/7	7/7	6/7**	6/7
		D	7/7	7/7	7/7	6/7**	7/7

* TP mg/kg-bw/d; T = t-test; D=Dunnett's; ** Laboratory 5 did not dissect the GP, so no data available, not a false negative

Negative substances – agonist assay

		VP	SVCG	LABC	GP	Cows
Nonyl-phenol	T-test	0/10	0/10	0/10	0/10	0/10
	Dunnett's	0/10	0/10	0/10	0/10	0/10
2,4-Dinitro-phenol	T-test	0/10	0/10	0/10	0/10	0/10
	Dunnett's	0/10	0/10	0/10	0/10	0/10

Negative substances – antagonist assay

	TP *		VP	SVCG	LABC	GP	Cows
Nonyl-phenol	0.2	T	1/3	0/3	0/3	1/3x	0/3
		D	0/3	0/3	0/3	0/3	0/3
	0.4	T	0/7	1/7	1/7	0/7	0/7
		D	0/7	0/7	0/7	0/7	0/7
2,4-Dinitro-phenol	0.2	T	0/3	0/3	0/3	1/3x	0/3
		D	0/3	0/3	0/3	0/3	0/3
	0.4	T	1/7	0/7	0/7	0/7	0/7
		D	1/7	0/7	0/7	0/7	0/7

* TP mg/kg-bw/d; T = t-test; D=Dunnett's; x Significant Increase

For statistical analysis the means, standard errors, standard deviations and CVs were calculated for each endpoint. In addition, the R-square values for different effects were calculated to give an indication of the strength of the association for an effect with an endpoint and to compare the robustness of the effect across endpoints, the variation from laboratory-to-laboratory, or to what degree the dose-responses vary among laboratories (VR3, x).

The responsiveness and the variability of the five mandatory tissues differed. The ability of laboratories to perform the protocol also appeared to differ when measured by the CVs of the tissues. In Phase-2, the fluid-filled tissues (VP, SVCG, and COW) generally had greater and more variable CVs than the solid tissues (LABC and GP). These observations were reinforced by the results obtained in Phase-3 for both test protocols. Notwithstanding the different CVs, the weight changes of all target organs fulfilled the expectations for chemicals acting as androgen agonists or antagonists. Nevertheless, some single "false negative" results were found for both protocols for some of the target organs in some laboratories and

similarly, occasionally there was also a “false positive” organ weight change after treatment with the negative reference chemicals (VR2, xiv, xv; VR3, xvi).

Several factors come into play when assessing the overall variability and sources of error of measurements. One of the major variables in the Hershberger Bioassay is laboratory personnel training and performance; it may be related to the ability to detect weakly acting substances (VR2, xv).

GD34 - Criteria 5/8

e) **Demonstration of the test method's performance should be based on the testing of reference chemicals representative of the types of substances for which the test method will be used.**

A sufficient number of the reference chemicals should have been tested under code to exclude bias (see paragraphs on "Coding and Distribution of Test Samples").

Representativity

The substances tested in the validation exercises were all small aromatic organic molecules with halogens or short functional groups conjugated to the aromatic ring structure as these are the appropriate size and shape to interact with the active pocket of the AR. They range from pharmaceuticals [specifically developed to act as AR agonists (TP, MT and TREN), antagonists (FLU) or blockers of synthesis of endogenous androgens (FIN)], to substances with broad environmental release and human exposure (pesticides PRO, VIN, LIN and industrial chemicals DNP and NP). These substances are broadly representative of the range of substances that would conceivably be tested using the Hershberger Assay.

Chemical	**Mode of action**	**Type of chemical**	**Testing phase of the Hershberger assay validation study**
Testosterone propionate (TP)	Potent androgen	Pharmaceutical	Tested in phase-**1**; reference chemical in phase **-2 and -3**
Flutamide (FLU)	Anti-androgen	Pharmaceutical	Tested in phase-**1**; reference chemical in phase **-2 and -3**
Methyltestosterone (MT)	Potent androgen (aromatisable)	Pharmaceutical	Tested in phase-**2**
Trenbolone (TREN)	Potent androgen (non-aromatisable)	Steroid anabolic	Tested in phase **-2 and -3**
Procymidone (PRO)	Weak anti-androgen	Pesticide	Tested in phase-**2**
Vinclozolin (VIN)	Its metabolites are weak anti-androgens	Pesticide	Tested in phase-**2**
Linuron (LIN)	Weak anti-androgen	Pesticide	Tested in phase **-2 and -3**
p,p'-DDE (DDE)	Weak anti-androgen	Pesticide	Tested in phase-**2 and -3**
Finasteride (FIN)	5α-reductase inhibitor	Pharmaceutical	Tested in phase-**2**
4-Nonylphenol, mixed isomers (NP)	No androgenic or anti-androgenic action	Industrial chemical	Negative reference chemical in phase-**3**
2,4-Dinitrophenol (DNP)	No androgenic or anti-androgenic action	Industrial chemical	Negative reference chemical in phase-**3**

The reference chemicals, Testosterone propionate (TP) and Flutamide (FLU) were the same for Phase-1, -2 and -3. In phase-2, for the agonists, the reference group was the vehicle control group. For the antiandrogens, the positive agonist test substance was TP, and it was the TP group that was the stimulated reference against which antagonists and 5α-reductase inhibitors were statistically evaluated. In some laboratories, FLU was voluntarily used as the control antagonist to confirm the performance of the antagonist assay.

The basis for selecting the two agonists (MT and TREN) was differences in their activation by 5α-reductase and the hypothesis that the tissues may respond differently to these agonists depending upon the activation by 5α-reductase. This would potentially generate a profile for distinguishing agonists activated by 5α-reductase and distinguishing androgen receptor antagonists from 5α-reductase inhibitors.

The basis for selecting the four antagonists was the demonstrated binding affinity of the parent or metabolites to the androgen receptor, the demonstrated activity of the parent or metabolite using *in vitro* assays, activity in previous Hershberger assays, and evidence for biological activity *in utero* inhibiting the development of the male reproductive tract.

The basis for selecting the 5α-reductase inhibitor (FIN) is similar evidence for its inhibition of the Type 2 5α-reductase activity found primarily in the male reproductive tract and evidence for its biological activity *in utero* inhibiting the development of the male reproductive tract (VR2, 33, 37 to 39).

Coding

The purpose of coded substance is to avoid bias in the results due to prior knowledge of the expected responses. In Phase-1 and -2 the substances tested were not coded and thus provided a set of "expected responses". The goal of the Pase-3 test validation study was to further assess the robustness and reproducibility of the Hershberger bioassay in a blinded manner. In Phase-3, coded substances were then tested at one or two predetermined dose levels to exclude possible investigator bias (VR3, 7, 8).

The reference chemicals (TP and FLU) were not coded. NP, DNP, LIN and p,p'DDE were coded and shipped from the centralised chemical repository to the participating laboratories. As for TREN, due to various specific national regulations, it was shipped directly from the producer to the participating laboratory (VR3, 27, 28). Thus it can be anticipated that for TREN, the confidentiality of the code was not kept.

GD34 - Criteria 6/8

f) **The performance of the test method should have been evaluated in relation to relevant information from the species of concern, and existing relevant toxicity testing data.**

In the case of a substitute test method adequate data should be available to permit a reliable analysis of the performance and comparability of the proposed substitute test method with that of the test it is designed to replace.

For the mammalian test methods (indeed for all section 4 –human health– test methods of the OECD TG Programme) the ultimate species of concern is the human. With the exception of the pharmaceutical agents, however, there is very little data to demonstrate the potency of the substances tested in the Hershberger Assay in humans. Even for the pharmaceuticals, data that do exist are unavailable as these are not in the public domain. However, the importance of the AR to the developing reproductive tract in human males has been amply demonstrated by the severe congenital deformations of the reproductive tract seen in male infants with mutant AR and in female infants with genetic defects of steroid production who overproduce androgen (i.e. Congenital Adrenal Hyperplasia). As the pharmacology of these substances are roughly equivalent between rat and human AR it is plausible, and thus appropriate to assume that, responses in rats will likely predict hazards relevant to humans.

To demonstrate that, the Hershberger Bioassay provides relevant information for rodent reproductive toxicity, some comparison may be made with data available from various developmental and reproductive data published in literature. Available studies with the test substances used in Phase-2 have been compiled and extracted for LOAELs and comparison was made with the Hershberger phase-2 results (see tables 1 to 6) (VR2, 186 to 188).

Table 1. Comparison of TREN Hershberger Phase 2 results with developmental and reproductive results

Hershberger Bioassay	**Developmental and/or Reproductive Bioassay**
VP: 8-40 mg/kg/d	By sc administration – increased female anogenital distance ≥ 0.5 mg/kg/d By sc administration – reduced female areolas retention ≥ 2 mg/kg/d By sc administration – reduced female nipple retention ≥ 0.5 mg/kg/d
SVCG: 40 mg/kg/d	
LABC: 8-40 mg/kg/d	
GP: 8-40 mg/kg/d	
COWS: 40 mg/kg/d	
	No frank malformations observed
Comments: The route of administration differences should be taken into account. The authors performed both oral and sc Hershberger studies. By gavage, LABC and SVCG were significant at 10 mg/kg/d trenbolone via po and VP, COWS, GP at 50 mg/kg/d trenbolone via po. By injection, LABC was significant at 0.05 μg/kg/d sc; other tissues at 0.20 μg/kg/d sc.	

Table 2. Comparison of PRO Hershberger Phase 2 results with developmental and reproductive results

Hershberger Bioassay	**Developmental and/or Reproductive Bioassay**
VP: 10-30 mg/kg/d	Anogenital distance: 25 mg/kg/d Retained nipples: 50 mg/kg/d
SVCG: 10-30 mg/kg/d	
LABC: 3-100 mg/kg/d	Malformations (hypospadias) incidence at ≥ 50 mg/kg/d Decreased adult sex accessory tissue wts at ≥ 100 mg/kg/d Histopath lesions in adult sex accessory tissue at ≥ 50 mg/kg/d
GP: 10 mg/kg/d – Non-detect	
COWS: 3-100 mg/kg/d	
Comments: There is general correspondence between the Hershberger results and those studies with *in utero* exposure to procymidone.	

Table 3. Comparison of VIN Hershberger Phase 2 results with developmental and reproductive results

Hershberger Bioassay	Developmental and/or Reproductive Bioassay
VP: 10-100 mg/kg/d	Anogenital distance reduced ≥ 3.125 mg/kg/d
SVCG: 10-30 mg/kg/d	Nipple retention in pnd 14 males increased ≥ 50 mg/kg/d Nipple retention in pubertal males increased ≥ 50 mg/kg/d
LABC: 10-100 mg/kg/d	Malformations (hypospadias) incidence at ≥ 50 mg/kg/d
GP: 10-100 mg/kg/d	Decreased adult VP wts at ≥ 50 mg/kg/d
COWS: 10-100 mg/kg/d	Reduced sperm count at 100 mg/kg/d
Comments: There is a general correspondence between the Hershberger results and the developmental assays. However, a multi-generation study did not reproduce the Anogenital distance findings at 3.125 -15 mg/kg/d.	

Table 4. Comparison of LIN Hershberger Phase 2 results with developmental and reproductive results

Hershberger Bioassay	Developmental and/or Reproductive Bioassay
VP: 30-Nondetect mg/kg/d	Anogenital distance not statistically significant up to 50 mg/kg/d
SVCG: 30-100 mg/kg/d	Nipple retention in pnd 13 males increased 50 mg/kg/d Nipple retention in pubertal males increased ≥ 50 mg/kg/d
LABC: 30-Nondetect mg/kg/d	Malformations (epididymis) incidence observable at ≥ 25 mg/kg/d
GP: 100-Nondetect mg/kg/d	Decreased adult dorsolateral prostate wts at 50 mg/kg/d
COWS: 100-Nondetect mg/kg/d	Histological abnormalities in male repro tract ≥ 25 mg/kg/d
Comments: There is a general correspondence between the Hershberger results and the developmental assays.	

Table 5. Comparison of DDE Hershberger Phase 2 results with developmental and reproductive results

Hershberger Bioassay	Developmental and/or Reproductive Bioassay
VP: 30-160 mg/kg/d	Anogenital distance reduced in Long Evans and not SD rats 100 mg/kg/d
SVCG: 30-160 mg/kg/d	Nipple retention in pnd 13 males increased 10 mg/kg/d SD rats; 100 mg/kg/d Long Evans rats
LABC: 30-100 mg/kg/d	Malformations (hypospadias) low incidence at 100 mg/kg/d in one study but not observed in second study
GP: 100-160 mg/kg/d	
COWS: 30-100 mg/kg/d	Decreased adult VP wts at 200 mg/kg/d
Comments: There is a general correspondence between the Hershberger results and the developmental assays.	

Table 6. Comparison of FIN Hershberger Phase 2 results with developmental and reproductive results

Hershberger Bioassay	Developmental and/or Reproductive Bioassay
VP: 0.2-1 mg/kg/d	Anogenital distance reduced pnd 1 ≥ 0.01 mg/kg/d
SVCG: 0.2-1 mg/kg/d	Nipple retention in pnd 13 males increased ≥ 0.01 mg/kg/d Nipple retention in adult males increased ≥ 0.1 mg/kg/d
LABC: 0.2-5 mg/kg/d	Malformations (multiple tissues) significant at ≥ 10 mg/kg/d
GP: 0.2-Nondetect mg/kg/d	Decreased adult LABC wts at ≥ 1 mg/kg/d
COWS: 0.2-5 mg/kg/d	Decreased adult VP and COWS wts at ≥ 10 mg/kg/d
Comments: 0.2 mg/kg/d was the lowest dose; the data of Bowman et al. were not available when the doses were selected.	

This comparison shows a high level of correspondence when the route of administration is similar. In the case of trenbolone (Table 1), there are route of administration considerations. The investigators performed two Hershberger bioassays themselves using very similar protocol conditions. The sc Hershberger corresponded well to the developmental data after exposure *in utero* with sc administration to the dams shown in Table 1. In parallel, the investigators' gavage Hershberger data corresponded well to those generated in Phase-2. There were no developmental or reproductive data for methyl testosterone discovered.

In conclusion, this brief comparison lends support to the toxicological relevance of the Hershberger bioassay to predict possible hazards from *in utero* exposure to (anti)androgens.

GD34 - Criteria 7/8

g)	**Ideally, all data supporting the validity of a test method should have been obtained in accordance with the principles of GLP.** Aspects of data collection not performed according to GLP should be clearly identified and their potential impact on the validation status of the test method should be indicated.

Several laboratories are GLP certified and conducted the studies following the OECD principles of Good Laboratory Practice. The other laboratories have stated that they have performed their work in the spirit of GLP but did not have the resources for full audit and archiving of material. As such, the degree to which the work followed all aspects of GLP in these latter labs is not known.

GD34 - Criteria 8/8

h)	**All data supporting the assessment of the validity of the test method should be available for expert review.** The detailed test method protocol should be readily available and in the public domain. The data supporting the validity of the test method should be organised and easily accessible to allow for independent review(s), as appropriate. The test method description should be sufficiently detailed to permit an independent laboratory to follow the procedures and generate equivalent data. Benchmarks should be available by which an independent laboratory can itself assess its proper adherence to the protocol.

It is intended that the materials documenting the protocol development, validation and associated supporting documents will be made freely available by promulgation through the OECD or by publication in the peer reviewed literature.

Manuscripts of Phase-1 and Phase-2 of the Hershberger Bioassay validation process have been submitted to a scientific journal. Additionally, it is expected that validation reports of phase-1 and phase-2 will be shortly declassified and thus publicly available in the near future.

The validation report of phase-3 has been posted on the protected website. It has been submitted for comments to the VMG mammalian and to the EDTA. A revised version taking into account of these comments should be available to the Peer reviewers.

The draft Test guideline and the Background Review Document will be submitted shortly and will be available to the Peer reviewers.